DOTTORATO DI RICERCA IN

Oncologia e Patologia Sperimentale

Ciclo XXVIII

Settore Concorsuale di afferenza: 06/A2

Settore Scientifico disciplinare: MED/04

CIRCULATING MICRORNAS
DURING HUMAN AGING AND LONGEVITY

Presentata da: Dott.ssa Cristina Morsiani

Coordinatore Dottorato   **Relatore**

Prof. Pier-Luigi Lollini   Prof. Stefano Salvioli

Co-relatore

Dott.ssa Miriam Capri

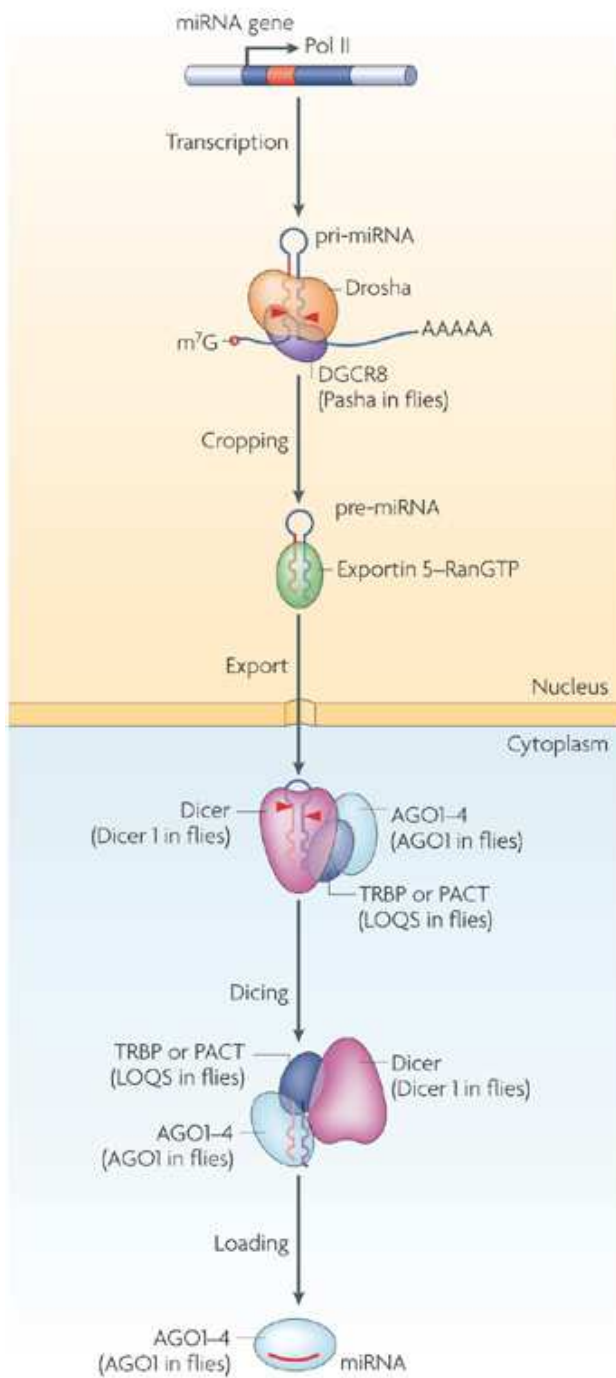**Esame finale anno 2016**

TABLE OF CONTENTS

1. INTRODUCTION

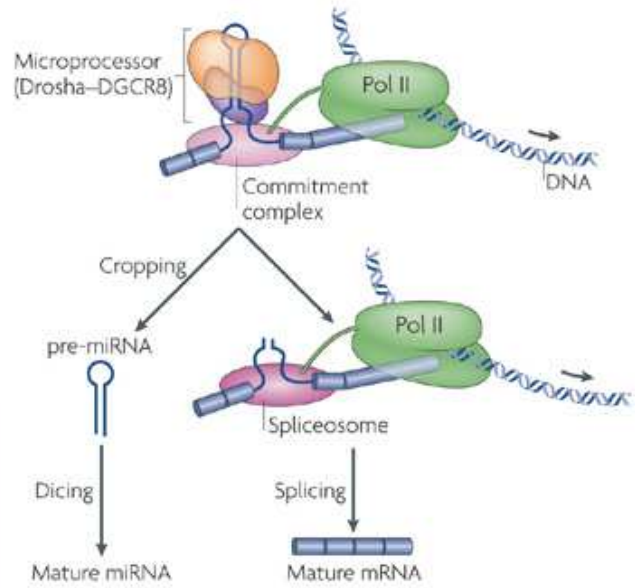## 1.1 MicroRNAs: definition and biogenesis

MicroRNAs (miRs) are a highly conserved class of small non-coding RNAs single-stranded, approximately 18-25 nt in length. These molecules interact with their mRNA targets by base pairing with complementary sequences in the 3' UTR region of mRNA molecules. MiRs have emerged as key regulators of gene expression at a post-transcriptional level since their discovery. The first miR, lin-4, was discovered in C. elegans by Ambros and his group in 1993 (Lee et al., 1993). Ambros succeeded in cloning lin-4 locus and he identified two small transcripts of approximately 22 and 61 nucleotides, where the latter fragment was the precursor of the shorter one. Sequence alignment analysis confirmed that both of them matched in the 3' UTR region of lin-14 mRNA. These experimental evidences supported the hypothesis of lin-4 being a negative regulator of lin-14, with the imperfect binding of 3' UTR and the formation of multiple RNA duplex that down-regulate lin-14 translation. The conclusion of these discoveries was that lin-4 must have a primary role in C. elegans developmental timing during post-embryonic events, controlling lin-14 (Wightman et al., 1993). A second miR, let-7, was identified seven years later in the same nematode model organism. As described for lin-4, let-7 showed an inhibitory effect on various C. elegans heterochronic genes (including lin-14) promoting the progression from late-larval to adult life stage by binding to 3' UTR regions of the target mRNA (Reinhart et al., 2000). Because of their role in regulating developmental stages transition, lin4 and let-7 were named "small temporal RNAs" (stRNAs), guessing they were actually members of a wider class of regulatory small RNAs still unknown at that moment (Pasquinelli et al., 2000). Soon after, many other miRs were isolated in different species and found to be related not only to particular temporal stages, but also to specific cell types. Currently, 35828 mature miR products from 223 different species are listed in "miRBase" database. Among them, 1881 are hairpin precursor expressing 2588 human mature miRs (http://mirbase.org, Release 21, June 2014). As negative gene regulators, miRs mediate protein synthesis repression in different cell types and they affect various biological processes, such as proliferation, differentiation, aging and apoptosis.

MiRs biogenesis is represented schematically in figure 1.1.

a Biogenesis of canonical miRNA

miRNA gene → Pol II

Transcription

pri-miRNA — Drosha

m⁷G — AAAAA

DGCR8 (Pasha in flies)

Cropping

pre-miRNA

Exportin 5–RanGTP

Export

Nucleus

Cytoplasm

Dicer (Dicer 1 in flies) — AGO1–4 (AGO1 in flies)

TRBP or PACT (LOQS in flies)

Dicing

TRBP or PACT (LOQS in flies) — Dicer (Dicer 1 in flies)

AGO1–4 (AGO1 in flies)

Loading

AGO1–4 (AGO1 in flies) — miRNA

b Canonical intronic miRNA

Microprocessor (Drosha–DGCR8) — Pol II

Commitment complex — DNA

Cropping

pre-miRNA — Pol II

Spliceosome

Dicing

Splicing

Mature miRNA — Mature mRNA

c Non-canonical intronic small RNA (mirtron)

pre-mRNA — Spliceosome

Splicing

Mature mRNA — Branched pre-mirtron (excised intron)

Debranching — or — Trimming

pre-miRNA

Dicing

Mature miRNA
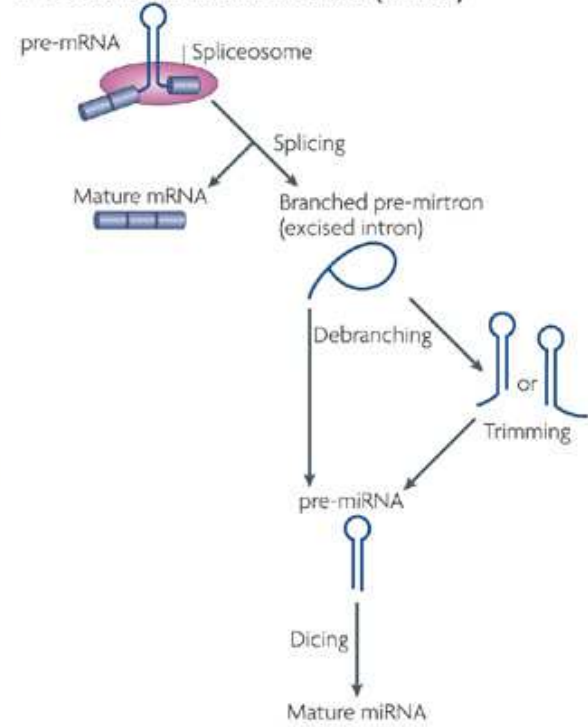
Nature Reviews | Molecular Cell Biology

Figure 1.1. miRs biogenesis pathway (Kim et al., 2009)

Most of miRs generate in the nucleus from RNA polymerase II transcription of miR genes. Only a small class of miR-genes, that are interspersed among Alu repetitive elements, are transcribed by RNA polymerase III. MiR loci are widespread throughout the genome and can be found either inside intronic or exonic regions of non-coding genes, intronic sequences of protein-coding genes, or between independent transcription units (intergenic). The majority of intronic miRs share the same promoter of the host gene, while a smaller number is transcribed from its own promoter enabling separate control of transcription (Monteys et al., 2010). MiRs biogenesis starts with RNA polymerase II or III transcription of pri-miRNA (primary miRNA), a stem-loop sequence which may encode for more than one mature miR. Like mRNAs, pri-miRs are spliced, 7-methylguanosine capped at the 5' terminus and polyadenylated at the 3' terminus. The first step of miR maturation is mediated by a protein complex consisting of the RNase III type enzyme Drosha and its essential cofactor DGCR8 (DiGeorge critical region 8), a dsRNA binding protein also known as Pasha. Besides the core proteins, the microprocessor requires in vivo other accessory proteins, such as the helicases p68 and p72, and the RNA binding protein hnRNP A1 (Liu et al., 2008). DGCR8 anchors pri-miR at the junction point between single and double strand RNA driving the correct positioning of Drosha, which cleaves approximately 11 nucleotides downstream into the stem region. Experimental evidences show that Drosha-mediated cleavage occurs in a co-transcriptionally fashion along with splicing. Pri-miRNA processing generates a 70 nucleotide product called precursor-miRNA (pre-miRNA), whose peculiar structure consists of a 5' phosphate, a hairpin region and a 2-nucleotides hydroxyl overhang at the 3' which is essential for nuclear export factor Exportin 5 (EXP5) recognition. The EXP5-Ran GTP complex recognizes the double-stranded stem as well as the 3' overhang, stabilizes pre-miR preventing its degradation and, finally, actively transports pre-miR intermediate from the nucleus to the cytoplasm. Ran-GTP is then hydrolyzed in Ran-GDP by RanGAP, causing EXP5 to release pre-miR molecule.

Canonical intronic miRs are processed co-transcriptionally before splicing. MiR-containing introns are recognized by the microprocessor and cleaved by Drosha. The resulting hairpin bears a 5' phosphate and a 3' overhang, thus entering the miR processing pathway while the rest of the transcript follows pre-mRNA maturing process. Specific experimental analysis show that miRs "cropping" process does not seem to affect mRNA splicing (Kim et al., 2009).

Non-canonical intronic miRs are a subgroup of miRs whose biogenesis depends totally on splicing machinery, therefore it completely bypasses microprocessor recognition and Drosha cleavage. This

class of miRs (named mirtrons) originates from debranched excised introns, also called "lariat". After debranching, mirtrons fold into hairpin structure that mimics pre-miR molecules and are subsequently exported in the cytoplasm.

Drosha nuclear cleavage defines one end of the mature miR. Once in the cytoplasm, a second RNase III enzyme called Dicer, in combination with cofactors TRBP (Tar RNA binding protein) and PACT (protein activator of dsRNA-dependent protein kinase), removes the loop sequence in pre-miRNA structure producing a 22 nt double stranded RNA having a 5' phosphate and 3' overhang on both ends (Bartel, 2004). The two strands of the duplex are named one antisense or "guide" strand, the other sense or "passenger" strand, and are usually identified respectively with the "-5p" and "-3p" suffix. Once cleaved, the duplex is unwound and generally only the guide strand is loaded into one of the Ago proteins of the Argonaute family as the mature miR, while the passenger strand is degraded. The main parameter to determine which strand will be retained is based on the thermodynamic asymmetry of miR molecules: the strand having the less stable 5' end will be loaded to Ago protein, while the other one will undergo degradation by Ago slicer activity. Some hairpins, however, generate mature miRs from both strands, confirming that strand selection is not a stringent process (Schwarz et al., 2003).

Mature miR and Ago protein are the core components of the effector complex, named miRNP (miR-containing ribonucleoprotein complex) or RISC (miR-induced silencing complex), that mediates miR's activity. Although miRNPs assembly mechanism is still to be clarified in humans, evidences show that Ago protein, Dicer, TRBP and PACT are the main components of the miRNP-loading complex (miRLC) which processes pre-miRs through Dicer cleavage. A yet-unidentified RNA helicase promotes the unwinding of the double strand, then the stable end of the miR duplex, i.e. the passenger strand, binds to TRBP and is soon hydrolyzed by Ago endonucleolytic enzymatic activity, while the mature miR, or guide strand, is loaded into Ago ribonucleoprotein. As miRLC finally disassembles, RISC core complex (miR and Ago) is configured (Maniataki and Mourelatos, 2005).

1.2 MiRs regulate gene expression

1.2.1 MiRs-mediated down-regulation of mRNA

Argonaute proteins are 95kDa proteins containing a PAZ and a PIWI domain. According to aminoacid sequence alignments, the Argonaute family can be clustered in the Ago and Piwi

6

subfamilies: while the former is expressed in all cell types, the latter can be found almost exclusively in stem cells and the germline. Humans have 4 different Ago proteins (Ago 1-4) that bind the 3' 2 nt overhang through their PAZ domain. Archeal Ago protein analysis showed that PIWI domain is structured as a RNaseH-like catalytic domain, which is characterized by the catalytic triad Asp-Asp-Glu (DDE) coordinating one or two Mg^{2+} ions. Argonaute proteins possesses a peculiar catalytic triad that comprises two aspartates and one histidine residue (DDH motif), which are responsible for Ago slicer activity. Among human Ago proteins, Ago2 is the only one competent for mRNA cleaving activity (Liu et al., 2004).

Through RISC complex, miRs mediate gene silencing via base-pairing between miR recognition element (MRE, located in the 3' UTR of target mRNA) and miR 5' "seed" region (nt 2 to 8) (figure 1.2).
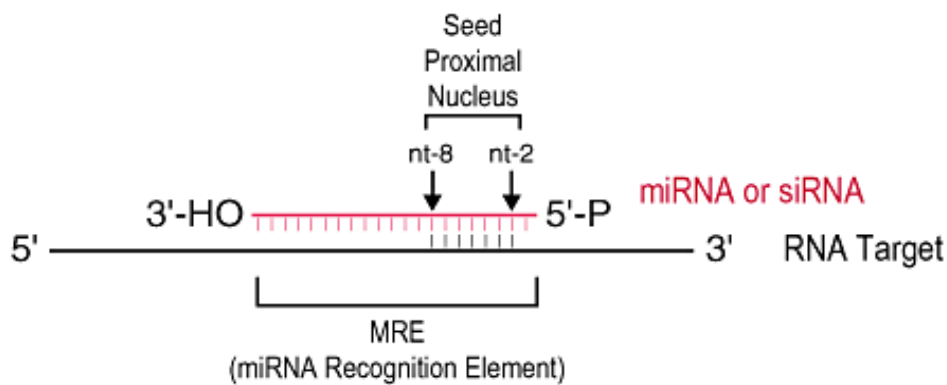
Figure 1.2. miR binding to target mRNA (Liu et al., 2008)

Recent findings, however, confirmed that some miRs actually possess a 3' end interaction site specific to the 5' UTR of the cognate mRNA, and that the most of translational repression was achieved when both 3' UTR and 5' UTR interaction sites were present in miR sequence (Lee et al., 2009).

The biological outcome of Ago-miR interaction strictly depends on the extension of base complementarity and on which Ago protein is deposited on the target mRNA. A non-perfect match between miRs and MRE region results in target mRNA destabilization and translational repression through different mechanisms and at different stages of protein synthesis.

First of all, miRs may impair translation initiation by chromatin remodeling or recruiting decapping and deadenylating enzymes as well as exonucleases and endonucleases. Additionally, Ago proteins compete with CBP (cap binding protein) and initiation factors, such as eIF4E, in binding the 5' methylguanosine cap, thus preventing mRNA circularization and translation initiation.

At a co-transcriptional stage, miRNPs may interfere with ribosome assembly or induce its dissociation due to steric hindrance. Besides, premature termination and cotranslational degradation of the nascent protein may occur, owed respectively to miRNPs competition with elongation factors and specific proteases enzymatic activity.

Apart from translational repression and mRNA destabilization, when a miR bound to Ago2 matches perfectly with the cognate mRNA, the target undergoes enzymatic degradation. As happens with other RNase H-like enzymes, Ago2 slicer activity cuts target mRNA near miR's nucleotides 10-11, producing 5′ phosphate and 3′ hydroxyl ends. While the mature miR remains intact, target mRNA is subsequently degraded via routine exonucleolytic digestion. Target mRNA cleavage by miRs is the major mechanism of regulation by plant miRs. In animals, however, there are very few examples of miRs that regulate their mRNA targets by cleavage; rather, the predominant silencing mode of animal miRs is to repress the translation of their mRNA targets and/or to destabilize them without endonucleolytic cleavage.

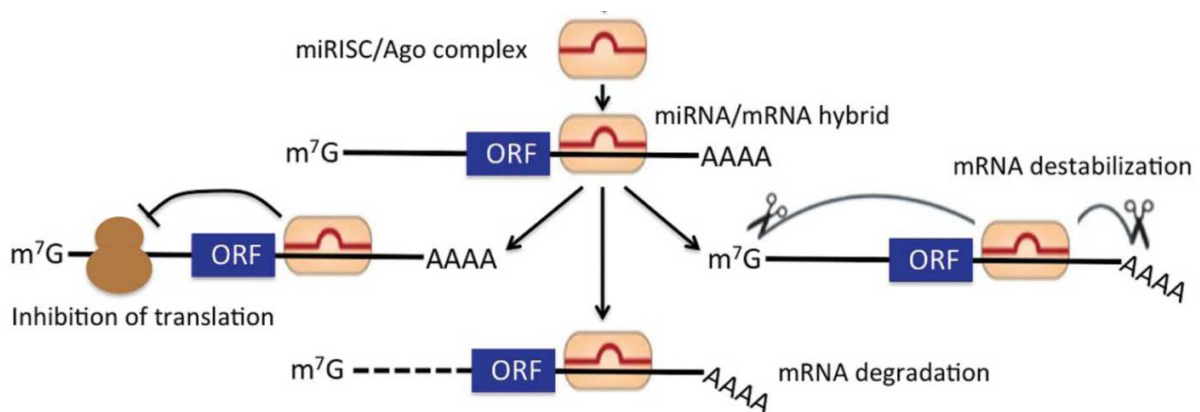All mechanisms of miRs interaction with mRNA target are summarized in figure 1.3.

Figure 1.3. Mechanism of action of miRs (modified from Piva et al., 2013).

1.2.2 MiRs-mediated up-regulation of mRNA

Since the last decade, many studies have revealed that not only miR-mediated down-regulation is a reversible process, but also that some miRs may act as positive regulators of gene expression. Translational enhancement may derive from a direct activation by miRNPs or from the relieving of a preexisting translational repression state.

One of the first cases reported in literature is miR-122, a liver-specific miR that binds either to the 5' UTR or the 3' UTR of Hepatitis C virus (HCV) genome. Recent studies have proved that the sequestration of miR-122 by miR-122 ASO (antisense oligonucleotide) led to a 80% decrease of HCV RNA and viral proteins. Indeed, when the binding site resided in the 5' NCR region, both viral gene replication and translation resulted up-regulated (Jopling et al., 2005). Since miR-122 binds two adjacent sites (seed-match-site 1 and 2) located upstream the internal ribosome entry site (IRES), the subsequent recruitment of miRNPs on 5' NCR could scaffold many other proteins required for genome replication and translation, e.g. RNA polymerase. In addition, these multiprotein complexes could conceal the uncapped 5' NRC of viral RNA from cytoplasmic exonucleases as well as mimic a cellular mRNA cap structure, therefore enhancing ribosome association and viral mRNA translation (Valinezhad Orang et al., 2014). Figure 1.4 represents the different mechanisms of miR-122-mediated HCV gene upregulation.
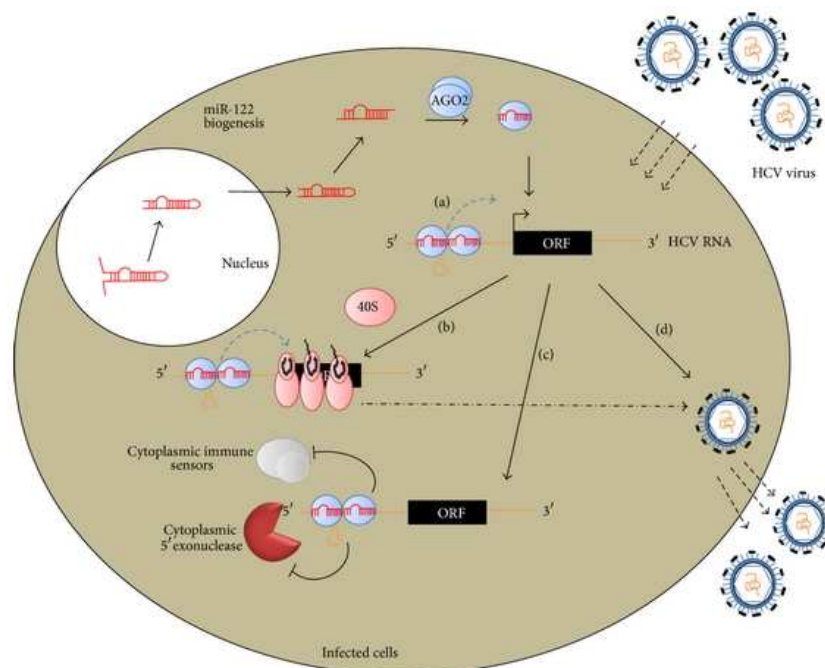

Figure 1.4. Activation of hepatitis C virus expression by miR-122, a liver-specific miR (Valinezhad Orang et al., 2014).

Another example of miRs targeting the 5' UTR is miR-10a, which positively regulates ribosomal proteins (RP) translation. A peculiar characteristic of several mRNAs encoding for RP, elongation factors and poly-A binding proteins, is the 5' terminal oligopyrimidine tract (5' TOP). This highly conserved motif acts as a regulatory sequence which suppresses RP synthesis in response to cellular stress or nutrient deprivation. miR-10a was found to bind to a sequence immediately downstream the 5' TOP motif, thus counteracting 5' TOP-induced down-regulation and leading to an increase in RP mRNA translation and protein synthesis (Ørom et al., 2008).

Recent studies revealed that miR-346 is produced from the second intron of glutamate receptor ionotropic delta 1 (GRID1) mostly in brain tissues and is capable of upregulating RIP140 (receptor-interacting protein 140) gene. RIP140 is a transcription coregulator and modulates many metabolism-related pathways; miR-364 accelerate mRNA interaction with polysomes via binding to 5′ UTR of the target RIP140 mRNA. Nevertheless, miR-346 does not require AGO2 for its activity; therefore, it possibly applies an AGO-independent pathway to control the protein yield of RIP140 without altering its mRNA levels (Tsai et al., 2009).

Translational activation can be also related to cell cycle phase. AU-rich elements (ARE) are cis-acting regulatory sequences of gene expression located in the 3' UTR of several mRNAs, encoding for oncogenes, growth factors, cytokines, etc. AREs are recognized by specific ARE-binding proteins (ABPs), which are generally mediate mRNA decay via interaction with exosomes, exonucleases and decapping enzymes. However, different ABPs can either perturb mRNA stability or mediate its translational activation through specific interactions with miRNP effector complexes, which are deposited on target mRNA by miR-mRNA matching.

An excellent example is FXR1 (Fragile X mental retardation syndrome-related protein 1), one of the ARE-binding proteins involved in mRNA translational up-regulation. FXR1 is generally associated with Ago2 in RNA-induced silencing complexes. This association was deeply investigated in a 2007 work in which it was reported that, even though FXR-1 and Ago2 are generally thought to switch off mRNA translation, FXR1-Ago2 complex promoted TNF-α upregulation after cell cycle arrest induced by serum starving, hence proving that AU-rich elements can act as translation activation signals in case of cell cycle arrest (Vasudevan and Steitz, 2007).

Furthermore, Zhang and colleagues published a 2014 study where they reported the ability of many cellular miRs to up-regulate gene promoters' activity by targeting the TATA-box region. TATA-box is a cis-regulating element located approximately 25 bp upstream the TSS (transcription start site)

within the core promoter. The first step of DNA transcription requires the binding of TBP (TATA-box binding protein) and the pre-initiation complex (PIC) assembly on TATA-box consensus sequence. Many endogenous miR, complexed with Ago proteins, were found to bind within 50 bp upstream the TSS in a sequence-specific fashion. The direct miR-TATA box interaction possibly facilitates TBP recruitment and PICs assembly on the core promoter (Zhang et al., 2014).

Finally, the same miR can mediate both down- and up-regulation. One of them is miR-145, which on one hand positively regulates the expression of myocardin and other co-activators of vascular smooth muscle cells differentiation, on the other hand inhibits ROCK1 expression in osteosarcoma cells, therefore acting as a tumor suppressor (Cordes et al., 2009; Li et al., 2014). As another example, KLF-4 is upregulated by miR-206 in confluent and nontumor cells, while it is downregulated by miR-344 in proliferating and normal cells (Lin et al., 2011).

1.3 Tissue-associated and circulating miRs

MiRs can be divided into two categories: tissue specific and circulating miRs. Some miRs show restricted tissue distribution, for example, miR-122 is highly enriched in liver, whereas miR-124 is preferentially expressed in neurological tissues. It has been shown that changes in the spectrum of cellular miRs correlate with various physiopathological conditions, including differentiation, inflammation, diabetes, and several types of cancers. Recently, some of the miRs previously identified in cells and tissues have also been found in extracellular fluids, show that miRs mediate long- and short-range cell communications through their secretion in the systemic circulation. These circulating miRs (c-miRs) can be found both in serum and plasma as well as in other bodily fluids such as urine, breast milk, saliva and interstitial fluids. Plasma miRs were found to be remarkably stable even under conditions as harsh as boiling, low or high pH, long-time storage at room temperature, and multiple freeze-thaw cycles (Chen et al., 2008). In contrast to the stability of c-miRs endogenous, when synthetic miRs were added exogenously, they were quickly degraded by the high level of RNase activity in plasma (Mitchell et al., 2008). This suggests that endogenous plasma miRs are protected in some manner to prevent their degradation. As shown in figure 1.5, c-miRs are protected by plasma/serum RNase activity by the formation of RNA-protein complexes with HDL, LDL or Ago2 proteins (Arroyo et al., 2011; Vickers et al., 2011), as well as through their

inclusion inside extracellular vesicles. These vesicles are generally referred to as microvesicles, apoptotic bodies, or exosomes. Exosomes are 40–100-nm vesicles released during reticulocyte differentiation as a consequence of multivesicular endosome (MVE) fusion with the plasma membrane and differs from microvesicles that are 50–1,000 nm vesicles and shed directly from the plasma membrane. Apoptotic bodies, the largest vesicles (800–5,000 nm diameter), are released by cells undergoing programmed cell death. Ten years ago it was demonstrated that the cargo of extracellular vesicles included miRs and that vesicles-associated RNAs could be transfer into target cells (Ratajczak et al., 2006; Valadi et al., 2007).
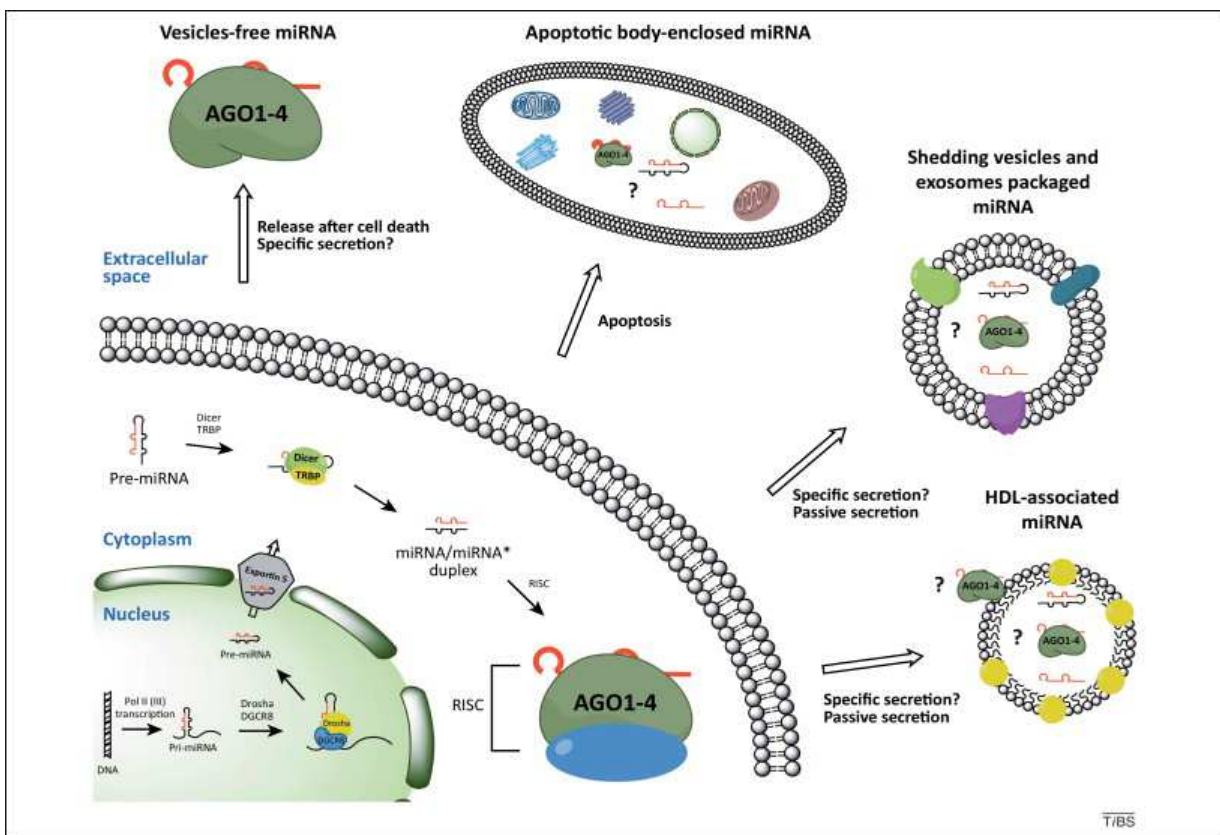
Figure 1.5. Biogenesis of miRs in the cell and the modes of extracellular miR packaging (Turchinovich et al., 2012)

Several groups investigated the exosome content in cultured cells in greater detail. Valadi and his group revealed that mast cell–derived exosomes carry about 121 different miRs. Some of these were found at relatively higher levels in exosomes than in their donor cells, which implies an active mechanism by which selected miRs are promoted toward exosomes. Others studies confirmed that cells can indeed select some miRs for cellular release while others are retained. Specifically,

although 66% of the miRs were released in quantities that reflected their intracellular level, 13% of the miR species were selectively retained by the cell (and thus released at lower levels), whereas on the other hand, 21% of the miRs seemed to be actively released and appeared at disproportionately higher levels (Pigati et al., 2010).

The proportion of miRs in the different cell-derived compartments is not yet settled. Currently, evidence is accumulating that the majority of miRs are not found inside vesicle but rather are bound to RNA-binding proteins. 2 populations of c-miRs exist: vesicle-associated and non–vesicle-associated miRs, and evidences support the hypothesis that vesicle-associated plasma miRs represent the minority, whereas potentially up to 90% of c-miRs are present in a non–membrane-bound form. These non–vesicle-associated miRs were specifically destabilized by proteinase K digestion of plasma, which indicates the existence of an miR-protein complex as a mechanism for their stability in the RNase-rich circulation (Arroyo et al., 2011). Interestingly, some miRs (e.g., let-7a) were exclusively associated with vesicles, whereas others (e.g., miR-122) were exclusively present in nonvesicle Ago2 complexes. This may reflect cell type–specific miR release mechanisms. For instance, the liver-specific miR-122 was detected only in protein-associated fractions, which suggests that hepatocytes release this miR through a protein carrier pathway. In contrast, miRs that are mainly associated with vesicles might originate from cell types that are known to generate vesicles, such as reticulocytes or platelets. Secretory miRs act as signaling molecules that modulate gene expression of recipient cell similarly to intra-cellular miRs. It was showed that T cells receive small RNAs from B cells that can affect the expression of target genes in the recipient T cells upon cell contact (Rechavi et al., 2009). Furthermore, it was demonstrated that miRs secreted by Epstein–Barr virus (EBV)-infected cells are transferred to and act in uninfected recipient cells through exosomes (Pegtel et al., 2010). In addition, it was reported that embryonic stem (ES) cell microvesicles contain abundant miRs and that they can transfer a subset of miRs to mouse embryonic fibroblasts in vitro, suggesting that stem cells can affect the expression of genes in neighboring cells by transferring miRs contained in microvesicles (Yuan et al., 2009).

Though it is well established the existence of c-miRs and their gene silencing activity, it is still to clarify the combination of factors and mechanisms behind their release and uptake. Only few reports show the mechanism of secretion of miRs. Recently, it was proved that miRs are released through a ceramide-dependent secretory machinery and the secretory miRs are transferable and functional in the recipient cells (Kosaka et al., 2010). Ceramide, a bioactive sphingolipid whose biosynthesis is

tightly controlled by neutral sphingomyelinase 2, triggers the secretion of exosomes. Blockade of neutral sphingomyelinase 2 by either a chemical inhibitor or by small interfering RNAs reduced secretion of miRs, whereas overexpression of neutral sphingomyelinase 2 increased the amount of extracellular miRs.

Among RNA-binding proteins identified, there were nucleophosmin 1 (NPM1), which has been implicated in the nuclear export of the ribosome, and nucleolin, a known NPM1-interacting protein (Borer et al., 1989). Subsequent experiments showed that NPM1 can fully protect synthetic miR degradation by RNase A. It has been suggested that NPM1 may be involved in shuttling RNAs from the nucleus to the cytosol, and independent studies have also shown that NPM1 can be released into the extracellular space (Nawa et al., 2009). Together, this suggests that this mechanism may be relevant for miR export and stability.

It has been proposed that the Ago2-miR complexes are passively released by death or apoptotic cells and remain in the extracellular space because of the high stability of the Ago2 protein. It is also possible that cell membrane–associated channels or receptors exist that allow for the specific release of these Ago2-miR complexes.

Concerning how HDL is loaded with miRs is not known exactly; however, biophysical studies suggest that HDL simply binds to extracellular plasma miR through divalent cation bridging. Furthermore, it was reported that ceramide pathway repress miR export to HDL, indicating that the export of specific miRs through the exosomal pathway and the HDL pathway may be distinct mechanisms, possibly opposing, although both pathways are probably regulated by nSMase2 activity and ceramide synthesis (Vickers et al., 2011).

In figure 1.6 are summarized the mechanisms for c-miRs transfer in different shuttle and their uptake in the target cells.

During "plasma membrane budding," ectosomes containing cytoplasmic components like miRs are released into the extracellular surrounding (Figure 1.6 - A). The fusion of exocytic multivesicular bodies (MVBs) with the plasma membrane releases miR-containing intraluminal vesicles, exosomes (B). The ABCA1 (ATP-binding cassette transporter A1) mediates the release of HDL-complexed miRs (C). Apoptotic bodies are released from a cell that undergoes apoptosis. These large vesicles contain fragmented DNA and cytoplasmic components including miRs (D).

General mechanisms of vesicle uptake in recipient cells involve endocytosis (E), phagocytosis (F), and fusion (G) with the plasma membrane. The uptake of HDL-complexed miRs in recipient cells is mediated by SR-B1 receptors (H). The uptake mechanisms of extravesicular AGO or NPM-1 complexed miRs are not described yet.
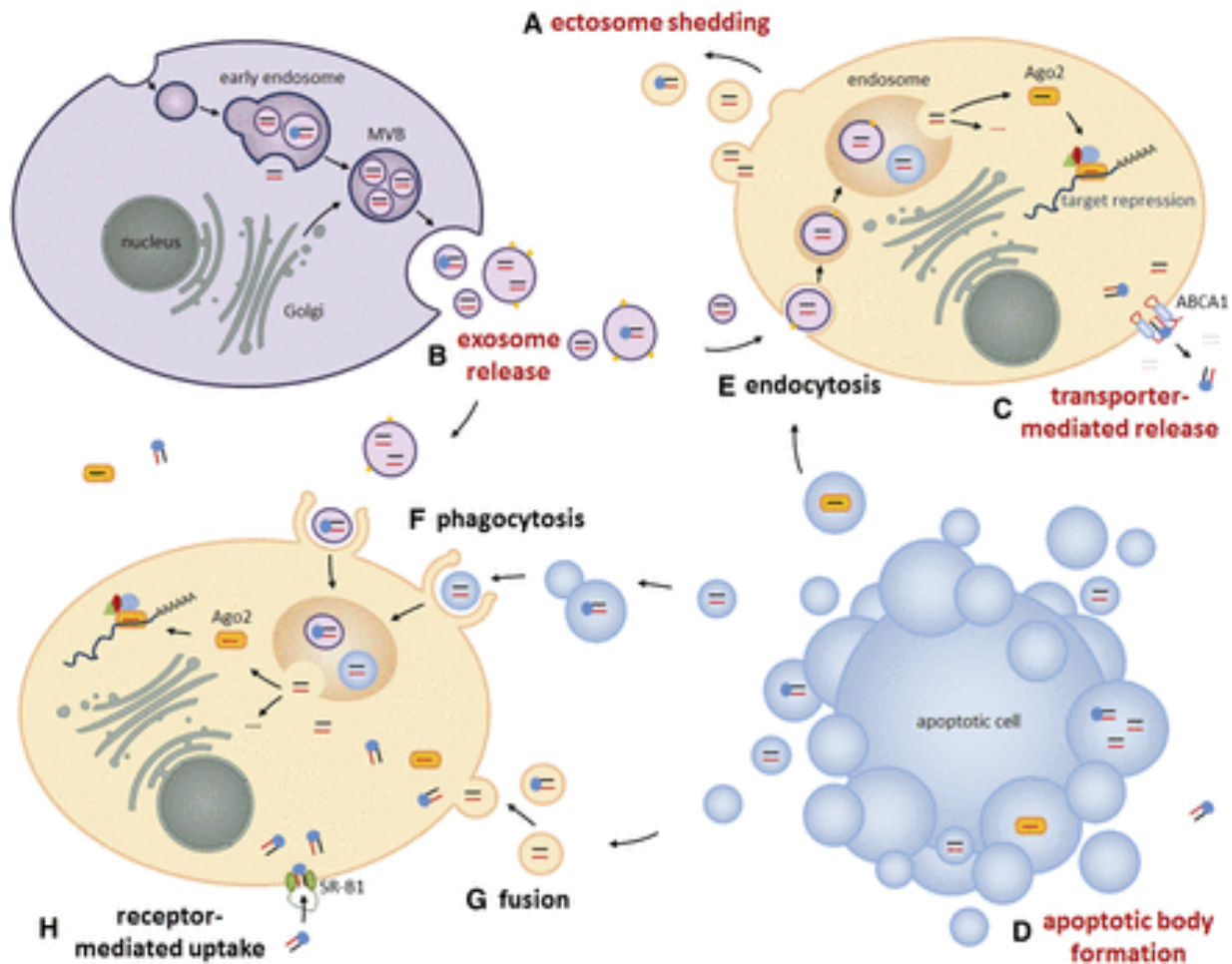
Figure 1.6. Long-distance intercellular miR transfer (Grasedieck et al., 2013).

1.4 C-miRs as potential biomarkers

The level and composition of c-miRs show changes that correlate well with diseases or injurious conditions. These observations suggest that extracellular miRs can be used as informative biomarkers to assess and monitor the body's physiopathological status (Etheridge et al., 2011). The ideal biomarker must be accessible using noninvasive protocols, inexpensive to quantify, specific to

the disease of interest, translatable from model systems to humans, and a reliable early indication of disease before clinical symptoms appear (early detection). Biomarkers that can be used to stratify disease and assess response to therapeutics are also medically valuable. Indeed most current biomarkers are protein based, such as troponin for cardiovascular conditions, carcinoembryonic antigen (CAE) for various cancers, prostate specific antigen (PSA) for prostate cancer, and aminotransferases (alanine aminotransferase, ALT and aspartate aminotransferase, AST) for liver function. Challenges for developing new protein-based biomarkers include the complexity of protein composition in most biological samples (especially blood), the diversity posttranslational modifications of proteins, the low abundance of many proteins of interest in serum and plasma, and the difficulty of reliably developing suitable high-affinity capture agents. These intricacies make the discovery and development of additional protein-based biomarkers with the proper diagnostic specificity and sensitivity an expensive, time-consuming, and difficult task. On the other hand, c-miRs have many requisite features of good biomarkers. Detecting specific miR species, while not trivial, is generally much easier. Stability in various bodily fluids, lower complexity, no known post-processing modifications, simple detection and amplification methods, tissue-restricted expression profiles, and sequence conservation between humans and model organisms make extracellular miRs ideal candidates for noninvasive biomarkers to reflect and study various physiopathological conditions in the body (Weber et al., 2010).

The changes of several miR levels in plasma, serum, urine, and saliva have already been associated with different diseases. For example, serum levels of miR-141, have been used to discriminate patients with advanced prostate cancer from healthy individuals (Mitchell et al. 2008), the ratio of miR-126 and miR-182 in urine samples can be used to detect bladder cancer (Hanke et al., 2010) and decreased levels of miR-125a and miR-200a in saliva is associated with oral squamous cell carcinoma (Park et al., 2009). In addition to these potential uses in detection of various cancers, another intriguing possibility is the use of levels of organ-specific miRs in body fluids to monitor the physiopathological conditions of specific organs. The severity of liver injury in this well-established model system can be precisely detected and monitored by measuring the levels of miR-122 in plasma (Wang et al., 2009). Others studies have demonstrated that the plasma level of miR-499, a heart specific miR, shows a perfect correlation with blood troponin levels in patients with myocardial infarction (Adachi et al., 2010).

All these evidences clearly demonstrates the possibility of using miR in circulation as strong and powerful biomarker to reflect the healthy and unhealthy status.

Moreover different strategies for overexpression or downregulation of specific miRs are being pursued for therapeutic intervention for various diseases. On the one hand, modulation of a single miR offers the opportunity to target multiple genes and regulatory networks simultaneously. However, for the same reason, careful design is needed to prevent unwanted off-target effects. Another reason to pay attention is the amount of delivered miR needed to affect the knockdown efficacy and target selection. MiR-based therapy can be considered as a powerful strategy for therapeutic purpose, but some issues are still to be bring under discussion. Nuclease-mediated degradation before achieving target modulation is a major issue in achieving the desired outcome. Systemic administration of miR using adeno-associated virus or cytoplasmic viruses of negative polarity capable of producing functional miRs have been described. MiR inhibitors termed miR sponges, antagomirs, locked-nucleic-acid-modified oligonucleotides, and reconstituted high-density lipoprotein nanoparticles are some of the approaches that have been pursued. Other strategies are used for the delivery, development of novel nanomaterials to pass barriers and antibodies against various cell surface receptors to be taken up by specific cells via receptor-mediated endocytosis. Vesicular structures that include exosomes and shedding vesicles are also being explored for delivery of exogenous therapeutic cargoes (Ajit, 2012). A better understanding of miR biology and the development of safe and effective delivery strategies can greatly enhance the therapeutic potential of miRs.

1.5 Human aging

Demographic projections report a world-wide dramatic increase in human life expectancy over the last few decades. The proportion of elderly people over 65 years in Europe (EU-28 Countries) is predicted to increase from 18.5% in 2014 to 28.7% by 2080. The share of those aged 80 years or above in the EU-28's population is projected to more than double between 2014 and 2080 (Eurostat source), these data are reported in figure 1.7. Italy in 2014 was one of the country with most high percentage of aged people in Europe. The share of persons in our country aged 65 or older was 21.4 % in the total population, people aged 80 or over were 6.4%. Also life expectancy at birth will

increase in the future year, from 70.5 years in 2015 to 77.1 in 2050 according to the World Population Ageing report (2015).
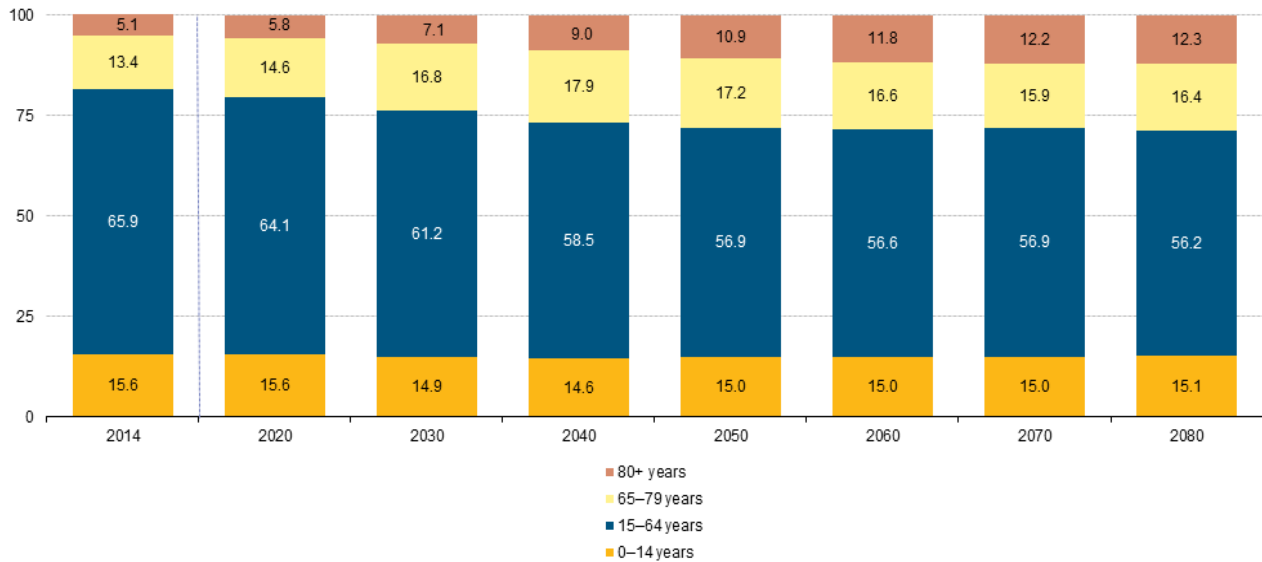
Figure 1.7. Population structure by major age groups, EU-28, 2014–2080 (% of total population).

This demographic remodeling emphasizes the critical importance of identifying new strategies able to counteract or slow down aging and the onset of age-related diseases and disabilities. These new strategies can contribute to increase the number of elderly citizens in good health, and reducing age-related medical and social costs. From this perspective, a thorough comprehension of the molecular and cellular aspects underlying aging is of primary importance to understand and, possibly, prevent age-related diseases, in order to improve quality of life as well as to promote healthy longevity.

Aging has been defined as the time-dependent decline of functional capacity and stress resistance, associated with increased risk of morbidity and mortality. It's a process based on the accumulation of unrepaired and deleterious changes occurring in molecules, cells, tissues and organs of the body with advancing age generated by internal and external sources. An integral part of the aging process is represented by the adaptive mechanisms that the body set up to compensate and neutralize the adverse effects of such damages and that lead to a progressive change of body composition and microenvironments.

18

It is also well established that the aging process is associated with a low-grade chronic inflammatory status, named "inflamm-aging" (Franceschi et al., 2000). In this view, healthy aging and longevity are likely the result not only of a lower propensity to mount inflammatory responses, but also of efficient anti-inflammatory networks. Studies performed on centenarians compared to old subjects have evidenced that centenarians escape the major age-related diseases, and a minority of them is still in quite good health (25% of the total centenarians, according to IMUSCE criteria). Centenarians represent a cohort of selected survivors, able to counterbalance the damaging effects of inflammaging by activating a series of anti-inflammatory networks and, at the same time, characterized by the capacity to mount an appropriate immune response counteracting infectious diseases and delaying all pathologies that normally cause mortality. Thus, centenarians can be considered a model of people capable of avoiding or postponing major age-related diseases (Franceschi and Bonafè, 2003). Accordingly, these exceptional long-living individuals have been found to be equipped with gene variants that allow them to optimize the balance between pro- and anti-inflammatory cytokines and other mediators of inflammation (Franceschi et al., 2007). Therefore, it is becoming evident that a genetic predisposition to produce excessive amounts of inflammatory molecules counteracted by an inadequate anti-inflammatory response might play a central role in the development of frailty and age-related pathologies in late life. Centenarians are a good choice for the study of human longevity, because they represent an extreme phenotype.

Because a number of age-related chronic diseases share a chronic inflammatory background, inflammaging represents a risk factor, or, more realistically, one of the major causes for age-related pathologies development such as type 2 diabetes, osteoporosis, Alzheimer's diseases, atherosclerosis and syndromes like sarcopenia (Cevenini et al., 2013). Sarcopenia is defined as the decrease in skeletal muscle mass and functionality associated with age. It's the progressive atrophy of skeletal muscles occurs during aging, leading to general frailty, disability and difficulties in everyday basic activities. Sarcopenia is a syndrome associated with many adverse clinical outcomes in elders, ranging from functional impairment to increased morbidity and mortality, which leads simultaneously to a remarkable rise in healthcare costs (Malafarina et al., 2012). Muscle mass loss is a gradual process that begins around the fifth decade. As sarcopenia progresses, muscle mass loss increases by 8% every decade until the age of 70 years old, and over this age the loss is estimated at 15% per decade (Kim and Choi, 2013). Muscle mass atrophy associated with physical inactivity and immobilization represents a major risk factor for secondary injuries, falls, functional impairment as

well as metabolic diseases and inability to perform daily living tasks, especially in the elderly. This is the case of bedridden and physically inactive centenarians, that show a huge difference respect to centenarians that are in good shape, able to walk and live a functional life.

The mechanisms underlying the aging process are beginning to be unraveled at the molecular level and some common denominators are discovered: genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication (López-Otín et al., 2013). Even if there is a global common background, it's clear the evidence that the rate of aging differs significantly between members of the same animal species, including humans. In other words, the biological age may differ from the chronological age. This variability is due to the complexity of genetics and epigenetics (Capri et al., 2014), which interact with environment (Garm et al., 2013) and stochasticity with diverse weights at different phases of life. Further, these main factors could differently affect the rate of aging at the levels of cells, tissues or body systems within the same organism according to the hypothesis of the "mosaic of aging" (Cevenini et al., 2008). This complexity makes more difficult the identification of a unique comprehensive mechanism of aging and related biomarkers (Deelen et al., 2013).

Many candidate biomarkers of human aging have been proposed in the scientific literature but in all cases their variability in cross-sectional studies is considerable, and therefore no single measurement has proven to serve a useful marker to determine biological age. A plausible reason for this is the intrinsic multi-causal and multi-system nature of the aging process.

Different immunological markers, as immunoglobulins, cytokines, autoantibodies, antibodies and cellular immunity, can be used as indicator of the biological state. It has been observed that a typical feature of the aging process is a general increase in plasmatic levels and cell capability to produce pro-inflammatory cytokines. The most involved cytokines are IL6, IL1β, TNFα and C-reactive protein (CRP). The loss of physiological control of the inflammatory reactions is likely due to an imbalance of the finely tuned equilibrium between the levels of pro and anti-inflammatory cytokines, or in the capability to restore the equilibrium once the inflammatory stimulus has been relieved, and it can lead to a chronic pro-inflammatory status (Salvioli et al., 2006).

Furthermore a wide group of classical clinical chemistry parameters have been proposed as potential biomarkers of aging. Examples are parameters of carbohydrate and lipid metabolism or hormones.

The integrity of the nuclear genome and the epigenome is of vital importance for the proper function of the body system. However, there is a constant attack by exogenous and endogenous compounds and agents that can damage DNA. Possible consequences are mutation and dysregulation of gene expression, which either can lead to cell death or cellular senescence or to malignant transformation of the cells ultimately resulting in cancer. Furthermore telomeric DNA undergoes attrition with each replication cycle and also as a consequence of DNA damage. A critical loss of telomere has been shown to prevent further cell proliferation and in some cell types induces cellular senescence. Mitochondrial DNA is an especially vulnerable target for mutagenesis, due to the high local levels of endogenous reactive oxygen species ROS and also because of a limited capability of their repair systems. Damage and mutation of mitochondrial DNA is viewed as a major mechanism driving the aging process (Altilia et al., 2012). Nevertheless D-loop region contains level of heteroplasmy associated with longevity, suggesting also mtDNA variants-based mechanisms of protection (Raule et al., 2014).

One important physiological posttranslational modification of secreted proteins is addition of N-linked oligosaccharides (N-glycans). N-glycans can modulate or mediate a wide variety of events in cell-cell and matrix interactions crucial for the development and function of organisms. Because the biosynthesis of glycans is controlled by the complicated concerted action of glycosyltransferases, the structures of glycans are much more variable and they can be easily altered by the physiological conditions of the cells. Accordingly, studying age-related alterations of the glycans could be relevant to understanding the complex physiological changes in aging individuals (Vanhooren et al., 2010).

During normal aging, there is accumulation of AGEs (Advanced Glycation Endproducts) of long-lived proteins such as collagens and cartilage, resulting from reactions between glucose and amino groups on proteins. AGEs directly or through interactions with their receptors, are involved in the pathophysiology of numerous age-related diseases (Simm et al., 2014), such as cardiovascular and renal diseases and neurodegeneration. It is well known that levels of oxidised proteins increase with age, due to the increase of protein damage and the decrease of their elimination. Since the proteasome is involved in the general protein turnover, it has received considerable attention during aging and evidence has been provided for an impairment of proteasome function with age in different cellular systems (Baraibar and Friguet, 2013). Certain oxidised proteins can be also repaired, the repair is limited to the reversion of a few oxidative modifications of sulphur-containing amino acids, such as the reduction of methionine sulfoxide by the methionine sulfoxide reductase

(Msr) system. Msr activity is known to be impaired during aging and replicative senescence (Picot et al., 2007).

Several recent studies propose biomarkers of aging based on DNA methylation levels (Garagnani et al., 2012; Zampieri et al., 2015). Recently it was developed the "epigenetic clock", based on 353 dinucleotide markers known as Cytosine phosphate Guanines or CpGs, applied to estimate the biological age of the most of human cell types, tissues, and organs (Horvath, 2013). Predicted age, referred to as "DNA methylation age" (DNAm age), correlates with chronological age in sorted cell types (CD4 T cells, monocytes, B cells, glial cells, neurons), tissues and organs including whole blood, brain, breast, kidney, liver, lung, saliva and even prenatal brain samples.

Looking at epigenetic regulators, many studies have shown that miRs levels change during the lifespan of several species and can be associated with age-related disorders. Taken together all these evidences, miRs can be propose to become powerful biomarkers of the aging process.

1.6 MiRs and aging

From the earliest evidences miRs were deeply involved in the aging process of several species. From nematode to mice it's now clear the extremely important function of these molecules in all cell types and organs. There are lot of miRs that regulate genes involved in known pathways of aging, as reported in figure 1.8.
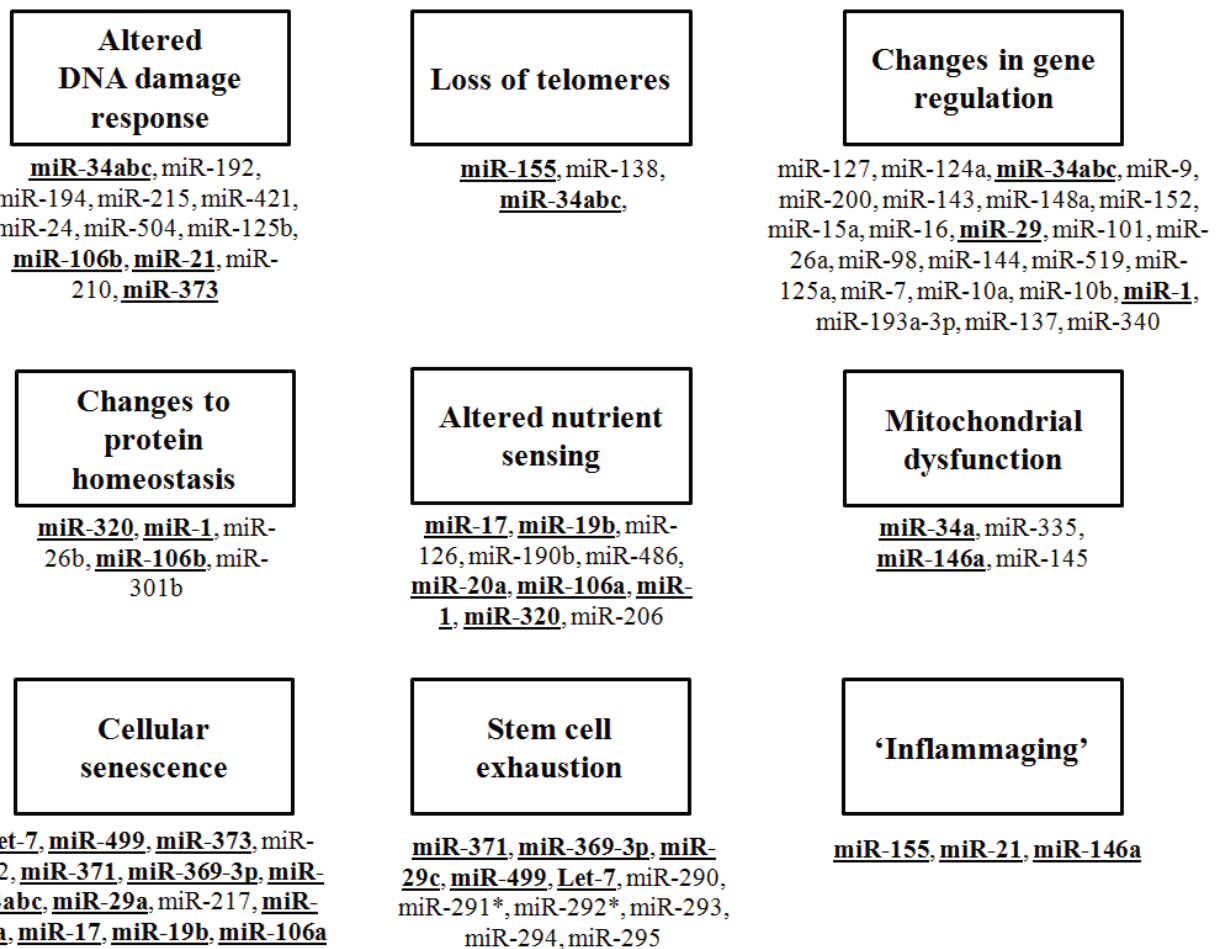
Altered DNA damage response

miR-34abc, miR-192, miR-194, miR-215, miR-421, miR-24, miR-504, miR-125b, miR-106b, miR-21, miR-210, miR-373

Loss of telomeres

miR-155, miR-138, miR-34abc,

Changes in gene regulation

miR-127, miR-124a, miR-34abc, miR-9, miR-200, miR-143, miR-148a, miR-152, miR-15a, miR-16, miR-29, miR-101, miR-26a, miR-98, miR-144, miR-519, miR-125a, miR-7, miR-10a, miR-10b, miR-1, miR-193a-3p, miR-137, miR-340

Changes to protein homeostasis

miR-320, miR-1, miR-26b, miR-106b, miR-301b

Altered nutrient sensing

miR-17, miR-19b, miR-126, miR-190b, miR-486, miR-20a, miR-106a, miR-1, miR-320, miR-206

Mitochondrial dysfunction

miR-34a, miR-335, miR-146a, miR-145

Cellular senescence

Let-7, miR-499, miR-373, miR-372, miR-371, miR-369-3p, miR-34abc, miR-29a, miR-217, miR-20a, miR-17, miR-19b, miR-106a

Stem cell exhaustion

miR-371, miR-369-3p, miR-29c, miR-499, Let-7, miR-290, miR-291*, miR-292*, miR-293, miR-294, miR-295

'Inflammaging'

miR-155, miR-21, miR-146a

Figure 1.8. The hallmarks of eukaryotic aging and the miRs that may interact with these pathways. Those miRs that are known to be involved in more than one feature are given in bold underlined text (Harries, 2014).

As previously described, c-miRs are already used for biomarkers in pathology conditions and many of them appear to be promising biomarkers for major aging-related diseases such as Alzheimer's disease (AD), type 2 diabetes mellitus (T2DM) and cardiovascular disease (CVD). Recent data have found an altered expression of c-miRs, including miR-9, miR-29a, miR-29b, miR-101, miR-125b, and miR-181c in serum of 105 AD patients compared with 150 age- and gender-matched healthy individuals (Geekiyanage et al., 2012). miR-125b and miR-181c levels were decreased but miR-9 level was increased in serum of AD patients compared with that of control individuals. Importantly, afterwards it was found that miR-125b level in AD patients was correlated with the Mini Mental

State Examination (MMSE), a test that is used to screen cognitive impairment in AD patients, suggesting that circulating miR-125b may act as a novel biomarker for AD (Tan et al., 2014).

MiRs were profiled in blood samples of over 800 individuals randomly selected from the Bruneck population (Bolzano Province, Italy) and identified five miRs (miR-15a, miR-28-3p, miR-29b, miR-126 and miR-223) downregulated in 80 participants with either prediabetes or T2DM. Interestingly, the level of these miRs is already reported to be changed 10 years before the onset of the disease, suggesting the potential role of c-miRs as early diagnostic biomarkers of T2DM (Zampetaki et al., 2010). Seven diabetes-related miRs (miR-9, miR-29a, miR-30d, miR-34a, miR-124a, miR-146a and miR-375) were found to be upregulated in T2DM patients compared with patients who had pre-diabetes or were susceptible to T2DM definite as Body Mass Index (BMI) ≥ 25 and/or with a family history of diabetes (Kong et al. 2011). MiR expression in the blood and exosomes of 265 patients with T2DM and metabolic syndrome were measured. They found miR-27a, miR-150, miR-192, miR-320a and miR-375 were upregulated in patients with T2DM. The expression levels were highly correlated with higher fasting glucose levels, suggesting the potential of the miRs as biomarkers for T2DM (Karolina et al., 2012).

MiR-1, miR-133a, miR-499 and miR-208a, were identified possible biomarkers and confirmed their expression in rats and humans with acute myocardial infarction (AMI) (Wang et al., 2010). AMI patients show higher abundance of all these miRs compared with healthy individuals. MiR-208a revealed the highest sensitivity and specificity in AMI diagnoses. Remarkably, increased level of miR-208a was detected at 1–4 h after acute injury, implicating the potential of miR-208 as an early diagnostic biomarker of AMI. miR-1 has been also suggested as a potential biomarker of AMI, the serum level was significantly increased with a 200-fold at six hours of AMI onset (Cheng et al., 2010).

MiR-1 and miR-133a are expressed in both cardiac and skeletal muscle tissue, and together with miR-206, expressed only in skeletal muscle, are the most studied myomiRs. MyomiRs are muscle-specific miRs, which have a primary role in muscle cells proliferation and differentiation (Kirby and McCarthy, 2013). miR-1 and miR-206 are also involved in skeletal muscle hypertrophy and atrophy (Güller and Russell, 2010) and miR-206 has been found up-regulated with age in mice biopsies (Hamrick et al., 2010). Generally, the involvement of myomiRs in aging and their presence in circulation it's clear (Zacharewicz et al 2013), now it should be better investigated the role of these

c-miRs during the progressive decrease in skeletal muscle mass and functionality associated with age.

Recent studies have reported miRs that are differentially expressed in serum or plasma during aging in mammalian, but only a few studies have shown significant alterations in c-miR levels during aging in human population. Olivieri and her group have measured plasma levels of miRs in healthy young and old humans, including centenarians, and in older patients with cardiovascular disease using an array of 365 miRs. miR-21 level was higher in the cardiovascular disease (CVD) patients and lower in the centenarian offspring compared to the age-matched healthy adults. They also suggested that transforming growth factor-beta (TGF-b) signaling can be one of the key pathway possibly modulated by the differentially expressed c-miRs (Olivieri et al., 2012).

A genome-wide miR screen for differential expression between long-lived individuals and controls revealed that 10 % of the miR microarray (863 miRs) demonstrated significant alterations in expression in blood sample. Most of these differentially expressed miRs have been associated with genes linked to major age-associated diseases, suggesting that regulation of key genes by miRs could promote longevity in humans (ElSharawy et al., 2012).

A cross-sectional analysis of individuals aged 50–90 and centenarians was used to identify that miR-363* decline with age but is preserved at youthful levels in the centenarians (Gombar et al., 2012).

The following year Noren Hooten has shown aging-related alterations in miR levels in human serum. MiR expression level was tested in sera from young (mean age 30 years) and old (mean age 64 years) individuals using next generation sequencing technology and real-time quantitative PCR. Among miRs that they found, miR-151a-5p, miR-181a-5p, and miR-1248 were significantly decreased in 20 older individuals compared to 20 younger individuals (Noren Hooten et al., 2013).

Solexa sequencing technology was used for a miR screening of serum samples in Chinese subjects of different aged, from 22 years to 70 years. It demonstrated 10 markedly altered miRs in the aging process, RT-qPCR identified five downregulated miRs (miR-29b, miR-106b, miR-130b, miR-142-5p, and miR-340) and three upregulated miRs (miR-92a, miR-222, and miR-375) with age (Zhang et al., 2015).

The multitude of important roles played by miRs suggest to improve methods and produce data on c-miRs expression in physiological and pathological conditions in human. However, quantification of miR has always been technically challenging due to small size, low copy number, interference from other small RNAs, and contamination by degradation products of mRNAs or other RNA

species. The increased availability and affordability of massively parallel sequencing offers a dramatically improved method to gain high-resolution views of miR expression.

2. AIM OF THE STUDY AND EXPERIMENTAL DESIGN

The main purpose of the current thesis is to investigate the role of blood circulating (c)-miRs and their expression profile during human aging and longevity by the means of the most advanced technology, i.e. smallRNA sequencing, with Next Generation Sequencing (NGS) platform.

Thus, primary objective was to optimize the protocol for smallRNA-seq (library preparation) using plasma samples and then to answer the following questions:

1. May c-miRs profile distinguish between healthy and unhealthy longevity? The answer of this question is towering the crucial aspect of EU aging population accordingly with recent World Health Organization (WHO) declaration that "increased longevity without quality of life is an empty prize", thus human aging research field has to face the problem to slow down the onset of age-related pathologies and promote healthy longevity.

2. May c-miRs profile specifically characterize aging and longevity trajectories? The answer of this question has been addressed since some years by Claudio Franceschi & Miriam Capri's team with the final objective to identify biological markers useful to monitor aging process trajectories.

The experimental design was established in two phases as follows:

1. Discovery phase on a few samples but at highest resolution using smallRNA-seq. miRs profile was conducted on samples from 12 donors (9 healthy subjects: 3 young, 3 old, 3 centenarians and 3 unhealthy centenarians). Subjects were enrolled in Bologna and centenarians were well characterized for their physical and cognitive status, the healthy centenarians had good cognitive performance, ability to walk and an high ADL (Activities of Daily Living) score. The unhealthy centenarians were demented and bedridden.

2. Validation phase on selected miRs and using RT-qPCR on a larger cohorts of subjects (48 total subjects), i.e. 16 healthy young donors, 16 healthy old donors and 16 centenarians, divided in 10 healthy and 6 unhealthy centenarians.

3. MATERIALS AND METHODS

3.1 Patients recruitment

Young, old and centenarians subjects included in the study were recruited in Bologna and characterized for the baseline information. Centenarians were also tested for hemato-biochemical analyses, such as white blood cell count (WBC), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution (% - RDWCV), albumin (ALB), total protein (PROT), C-reactive protein (CRP). Further, a detailed questionnaire describing their life style, physical and cognitive status was administrated. The healthy centenarians were defined by good cognitive performance, i.e. MMSE (Mini-Mental State Examination) > 24 scores, ability to walk and an high ADL (Activities of Daily Living) score. The unhealthy centenarians were demented without the possibility to perform the MMSE and bedridden. Young, old subjects were defined in health condition by general practitioners.

3.1.1 Subjects included in the discovery phase

12 donors were included in the discovery phase (smallRNA sequencing) and their characteristics are reported in the table I.

Table I. Age and gender characteristics of the 12 subjects chosen for the smallRNA-seq.

	Code	Age	Sex
Young	M002	25	F
	M003	26	F
	M004	26	M
Old	10M032	73	F
	10M064	69	F
	10M107	71	M
Healthy centenarians	1139M	102	M
	1145M	101	F
	1146M	102	F
Unhealthy centenarians	1141M	101	F
	1142M	104	F
	1143M	101	F

3.1.2 Subjects included in the validation phase

48 patients (including those previously selected for discovery phase) were included in validation of the sequencing results, in order to confirm previously selected miRs. The subjects were divided in the following groups (average age ± standard deviation):

- 16 healthy young donors (8 female and 8 male), 30.1 ± 3.4
- 16 healthy old donors (8 female and 8 male), 70.8 ± 1.9
- 16 centenarians: 10 healthy centenarians (7 female and 3 male) and 6 unhealthy centenarians (only females), 101.4 ± 1.3

3.1.3 Sample processing

Blood was collected in vacuum tubes following standard procedures and processed within 1 hour. The tubes were centrifuged at 2000 xg for 20 minutes at 4°C, separated plasma was collected and aliquoted in cryotubes for long term storage at -80°C.

3.2 SmallRNA sequencing

The protocol is composed by different steps, each has been optimized to reach the best results in sequencing: RNA extraction and evaluation of the RNA yield, cDNA library preparation and quantification, library size selection. When the library is ready, the samples can be finally run in the sequencer.

3.2.1 RNA extraction

Different kits were tested for RNA isolation, and in particular two kits were applied for the comparison of RNA yield, one kit from the Qiagen company and one from the Norgen. RNA extraction was assessed also testing different starting volume of EDTA-plasma: 0.2 – 0.5 – 1 – 2 – 5 ml, and at the end 2 ml of plasma was chosen as best solution. The protocols described below refer to the 2 ml of plasma. UniSp2, UniSp4, UniSp5 were introduced at the beginning of RNA extraction to evaluate at the end the quality of RNA isolation.

Total RNA was isolated from 2 ml of plasma samples with the miRNeasy Mini kit (Qiagen), adapted for total RNA extraction from plasma. Here is reported the protocol used:

- ➢ Centrifuge at 4°C for 5 minutes at 12,000 xg to clean samples from impurities
- ➢ Transfer cleaned plasma to a new 1.5 ml tube
- ➢ Add to sample 5 ml Qiazol + 1 µl RNA Spike-Ins kit (Exiqon)
- ➢ Mix by pipetting and vortex vigorously for 10 seconds
- ➢ Incubate 5 minutes at room temperature (RT)
- ➢ Add 2 ml chloroform, vortex vigorously for 10 seconds
- ➢ Incubate for 3 minutes at RT
- ➢ Centrifuge at 4°C for 15 minutes at 12,000 xg
- ➢ Transfer upper aqueous phase and add glycogen (5 mg/ml) to a final concentration of 50 mg/ml, mix by vortexing
- ➢ Add 1.5 volumes of 100% ethanol, mix by pipetting and inverting the tube
- ➢ Transfer all volume to an RNeasy Mini Spin Column attached to a vacuum system and apply 500 mbar to drawn through the column the lysate by vacuum pressure
- ➢ Wash the column by adding 700 µl RWT and using vacuum pressure
- ➢ Wash three times by adding 500 µl RPE
- ➢ Transfer the column to a new collection tube and spin at 13,000 xg for 2 minutes at RT
- ➢ Leave the tube open for 1 minute
- ➢ Transfer the column to a new 1.5 ml tube and add 30 µl RNase-free water
- ➢ Incubate 1 minute and spin at 13,000 xg for 1 minute at RT
- ➢ Store RNA at -80°C

Plasma/Serum circulating and exosomal RNA purification kit was chosen as the best among Norgen kits and compared in more detail with the Qiagen kit. The protocol adapted for this procedure is reported below:

- ➢ Warm up PS Solution A, PS Solution B and PS Solution C for 20 minutes at 60°C and mix well until the solutions become clear again if precipitates are present
- ➢ In a 50 ml tube, add to 2 ml of plasma, 0.2 ml PS Solution A and 3.8 ml PS Solution B after the addition of β-mercaptoethanol (the use of β-mercaptoethanol in lysis is highly

recommended to isolate RNA for sensitive downstream applications, 10 µl of β-mercaptoethanol were added to each 1 mL of PS Solution B)

➢ Mix well by vortexing for 15 seconds

➢ Incubate the mixture for 10 minutes at 60°C

➢ After incubation add 6 ml of 100% ethanol and mix well by vortexing for 15 seconds

➢ Centrifuge for 30 seconds at 1,000 rpm at RT, then carefully decant the supernatant in order to ensure that the pellet is not dislodged

➢ To the pellet add 0.3 ml PS Solution C, and mix well by vortexing for 15 seconds

➢ Incubate the mixture for 10 minutes at 60°C

➢ After incubation add 0.3 ml 100% ethanol and mix well by vortexing for 15 seconds

➢ Transfer 650 µl from the mixture into a Mini Filter Spin column and centrifuge for 1 minute at 14,000 rpm at RT. Discard the flowthrough and reassemble the spin column with its collection tube

➢ Repeat step until all the mixture has been transferred to the Mini Filter Spin column

➢ Apply 400 µl of Wash Solution to the column and centrifuge for 1 minute at 14,000 rpm at RT. Discard the flowthrough and reassemble the spin column with its collection tube

➢ Repeat the wash step two more times, for a total of three washes

➢ Spin the column empty, for 3 minutes at 14,000 rpm at RT. After the centrifugation discard the collection tube

➢ Transfer the spin column to a fresh 1.5 ml elution tube. Apply 100 µl of Elution solution to the column and centrifuge for 2 minutes at 2,000 rpm at RT, followed by 3 minute at 14,000 rpm.

➢ RNA is now ready for downstream applications or to be store at -80°C.

RNA is normally scarcely present in circulation, thus 100 µl of elution solution carrying the RNA extracted in a very low concentration due to the high dilution. Different elution volumes and concentration protocols were tested, the best result was reached with an overnight precipitation, applying the following protocol:

➢ Add 10 µl of 3M NaOAc (pH 5.2), 1 µl of Glycogen Solution and 300 µl of 100% ethanol. Mix gently but thoroughly

➢ Incubate the mixture at -80°C for an overnight precipitation

- The day below, vortex the mixture and centrifuge for 30 minutes at 16,000 xg at 4°C. Discard the supernatant without disturbing the pellet
- Wash the pellet with 70% ethanol
- Centrifuge 15 minutes at 16,000 xg. Discard the supernatant carefully to avoid disturbing the pellet
- Air dry the pellet, being careful to not over-dry, which renders resuspension more difficult
- Dissolve RNA pellet in 15 µl nuclease-free water

3.2.2 RT-qPCR LNA technology

In Grillari's lab, it's usually applied the LNA system from Exiqon to perform RT-qPCR (Reverse Transcription quantitative PCR). LNA, locked nucleic acid, are a class of high-affinity RNA analogs in which the ribose ring is "locked" in the ideal conformation for Watson-Crick binding. As a result, LNA oligonucleotides exhibit thermal stability when hybridized to a complementary DNA or RNA strand, this is important studying small molecules and widely varying of GC-content, such as miRs. In addition, LNA oligonucleotides can be made shorter than traditional oligonucleotides and this is important when the oligonucleotide is used to detect small or highly similar targets. Some miR family members vary by a single nucleotide, LNA can be used to allow excellent discrimination of closely related miR sequences.

The miR assay expect a first retrotranscription (RT) step and then the qPCR.

cDNA synthesis was performed using miRCURY LNA Universal RT microRNA PCR (Exiqon). During the RT a poly-A tail is added to the mature miR, the synthesis is done using a poly-T primer with a 3' degenerate anchor and a 5' universal tag. Here is reported the protocol used:

- Add to 2 µl RNA, 8 µl Universal cDNA Synthesis Mastermix composed by
    - 2 µl 5x Buffer
    - 0.5 µl cDNA Spike-In (UniSp6)
    - 1 µl Enzyme Mix
    - 4.5 µl nuclease-free water
- Incubate at 42°C for 60 minutes with heated lid set to 99°C
- Heat-inactivate at 95°C for 5 minutes

qPCR uses miR-specific LNA enhanced forward and reverse primer to amplify the cDNA template. SYBR Green (ExiLENT SYBR Green master mix – Exiqon), which only binds to double stranded DNA, was used for detection as described below:

➢ Dilute 2 µl cDNA in 78 µl nuclease-free water (40x dilution)

➢ Distribute 4 µl of diluted cDNA into all wells, for NTC use 4 µl nuclease-free water

➢ Add 6 µl PCR Mix composed by 5 µl PCR Master Mix and 1 µl PCR Primer Mix

➢ Run qPCR using this program

    - Activation/Denaturation: 95°C 10 minutes

    - Amplification 40 cycles:

    95°C 10 seconds

    60°C 60 seconds

    - Melting curve analysis

Real time PCR allow the investigation of gene expression patterns based on the assessment of Cycling threshold (Ct) value, which is the number of cycles required for the fluorescence signal to cross the threshold line. The Ct value is inversely proportional to the initial copy number of the miR in the sample.

3.2.3 cDNA library preparation

For cDNA library preparation 3 commercial kits were tested: XRNA Exosome RNA-seq NGS library prep kit (System Biosciences), TruSeq Small RNA library prep kit (Illumina), CleanTag Ligation Kit (TriLink BioTechnologies). The three kits apply different procedures and strategies to perform cDNA library preparation.

1) The XRNA Exosome RNA-seq library prep kit from System Biosciences adopted the following protocol:

STEP 1: 3' Adapter Ligation

➢ Allow Mix C300 to equilibrate to RT for 30 minutes before use

➢ Denature the RNA Sample by assembling in a sterile 200 µl PCR tube 8 µl of this reaction mix:

- 6 μl RNA sample
- 2 μl Mix A300

➢ Gently pipette mix thoroughly and incubate at 70°C for 1 minute and then place the tube on ice

➢ Set up the following 3' Adapter Ligation reaction for a total volume of 16.5 μl:
  - 8 μl Denatured RNA mix
  - 2 μl Mix B300
  - 6.5 μl Mix C300

➢ Gently pipette mix thoroughly and incubate at 25°C for 1 hour

STEP 2: Ligation Product Clean Up

➢ Vortex the AMPure XP beads until they are evenly resuspended

➢ Add 30 μl of AMPure XP beads to each sample. Gently pipette mix thoroughly and incubate at RT for 15 minutes

➢ Place the sample tube on the magnetic stand at RT for 5 minutes

➢ Remove and discard 40 μl of the supernatant

➢ Keep sample tube on the magnetic stand. Gently add 100 μl of 80% ethanol into each sample tube without disrupting the beads. Incubate at RT for 30 seconds

➢ Remove and discard 95 μl of the supernatant

➢ Repeat washes steps. Remove and discard all residual supernatant after the second 80% ethanol wash

➢ Air dry sample tube at RT for 15 minute or until the AMPure XP beads are dry

➢ Remove sample tube from the magnetic stand. Resuspend the dried AMPure XP beads in 8.5 μl of nuclease free water and incubate resuspension at RT for 2 minutes

➢ Place the sample tube on the magnetic stand at RT for 5 minutes

➢ Transfer 7 μl of the supernatant into a fresh 200 μl PCR tube

STEP 3: 5' Adapter Ligation

➢ Set up the following 5' Adapter Ligation reaction for a 12 μl total reaction:
  - 7 μl 3' Adapter Ligated RNA

- 3 μl Mix D300
- 2 μl Mix E300

➢ Gently pipette mix thoroughly and incubate at 25°C for 1 hour, then place the tube on ice

STEP 4: cDNA Synthesis

➢ Set up the following 15 μl cDNA Synthesis reaction on ice:
- 12 μl 3' and 5' Adapter Ligated RNA
- 2 μl Mix F300
- 1 μl Mix G300

➢ Gently pipette mix thoroughly and incubate at 50°C for 1 hour, then place the tube on ice

STEP 5: PCR Amplification

➢ Set up the following PCR reaction in a fresh sterile 200 μl PCR tube on ice. Only one of the barcode primers is used for each sample. The reaction of 25 μl is composed by:
- 5 μl cDNA
- 18 μl Mix H300
- 1 μl PCR Primer
- 1 μl Barcode Primer

➢ Gently pipette mix thoroughly and amplify the samples in the thermal cycler using the following PCR cycling conditions:

98°C for 30 seconds

18-20 cycles of: 98°C for 15 seconds, 60°C for 15 seconds, 72°C for 1 minute

72°C for 5 minutes

Hold at 4°C

2) TruSeq Small RNA library prep kit from Illumina applied the protocol below:

STEP 1: 3' Adapter ligation

➢ Set up the ligation reaction in a sterile nuclease-free 200 μl PCR tube on ice, the reagent volume is 6 μl in total:

- – 1 μl RNA 3' Adapter (RA3)
- – 5 μl 1 μg Total RNA in nuclease-free water

➢ Gently pipette the entire volume up and down to mix thoroughly, and then centrifuge briefly

➢ Place the tube on the preheated thermal cycler. Close the lid and incubate the tube at 70°C for 2 minutes and then immediately place the tube on ice (to prevent secondary structure formation)

➢ Prepare 4 μl mix in a fresh 200 μl PCR tube on ice for each sample
- – 2 μl Ligation Buffer (HML)
- – 1 μl RNase Inhibitor
- – 1 μl T4 RNA Ligase 2, Deletion Mutant

➢ Gently pipette the entire volume up and down to mix thoroughly, and then centrifuge briefly

➢ Add 4 μl of the mix to the reaction tube from the previous step and gently pipette the entire volume up and down to mix thoroughly. The total volume of the reaction is 10 μl

➢ Place the tube on the preheated thermal cycler. Close the lid and incubate the tube at 28°C for 1 hour

➢ With the reaction tube on the thermal cycler, add 1 μl Stop Solution and gently pipette the entire volume up and down to mix thoroughly. Continue to incubate the reaction tube on the thermal cycler at 28°C for 15 minutes and then place the tube on ice

STEP 2: 5' Adapter ligation

➢ Add 1.1 μl RNA 5' Adapter into a new 200 μl PCR tube, for a total equal to the number of samples being prepared

➢ Place the tube on the preheated thermal cycler, close the lid and incubate the tube at 70°C for 2 minutes and then immediately place the tube on ice (to prevent secondary structure formation)

➢ Add 1.1 μl 10mM ATP to the RNA 5' Adapter aliquot tube, with an equal number to the number of samples being prepared. Gently pipette the entire volume up and down to mix thoroughly

➢ Add 1.1 μl T4 RNA Ligase to the RNA 5' Adapter aliquot tube, equal to the number of samples being prepared. Gently pipette the entire volume up and down to mix thoroughly

➤ Add 3 μl of the mix from the RNA 5' Adapter aliquot tube to the reaction from the step of Ligate 3' Adapter. Gently pipette the entire volume up and down to mix thoroughly. The total volume of the reaction is 14 μl

➤ Incubate the reaction tube at 28°C for 1 hour and then place the tube on ice

STEP 3: Reverse Transcription

➤ Dilute the 25 mM dNTPs by premixing the following reagents in a fresh 200 μl PCR tube. Multiply each reagent volume by the number of samples being prepared.
  - 0.5 μl 25 mM dNTP Mix
  - 0.5 μl Ultra Pure Water

➤ Gently pipette the entire volume up and down to mix thoroughly, then centrifuge briefly and place it on ice.

➤ Transfer 6 μl of each 5' and 3' adapter-ligated RNA to a new 200 μl PCR tube

➤ Add 1 μl RNA RT Primer to each tube containing adapter-ligated RNA

➤ Gently pipette the entire volume up and down to mix thoroughly, and then centrifuge briefly

➤ Place the tube on the preheated thermal cycler and incubate the tube at 70°C for 2 minutes, then immediately place the tube on ice

➤ Prepare the following mix in a separate 200 μl PCR tube placed on ice. Multiply each reagent volume by the number of samples being prepared. The total volume per sample is 5.5 μl
  - 2 μl 5X First Strand Buffer
  - 0.5 μl 12.5 mM dNTP mix
  - 1 μl 100 mM DTT
  - 1 μl RNase Inhibitor
  - 1 μl SuperScript II Reverse Transcriptase

➤ Gently pipette the entire volume up and down to mix thoroughly, and then centrifuge briefly

➤ Add 5.5 μl of the mix to the reaction tube with the retrotranscribed. Gently pipette the entire volume to mix thoroughly, and then centrifuge briefly. The total volume is 12.5 μl

➤ Incubate the tube at 50°C for 1 hour and then place the tube on ice.

STEP 4: PCR Amplification

➤ Prepare a separate PCR tube for each index used, only 1 for each reaction. Combine the following reagents, multiply each reagent volume by the number of samples being prepared. 37.5 μl is the total volume per sample.
  – 8.5 μl Ultra Pure Water
  – 25 μl PCR Mix (PML)
  – 2 μl RNA PCR Primer (RP1)
  – 2 μl RNA PCR Primer Index (RPIX)

➤ Gently pipette the entire volume up and down, centrifuge briefly, and then place the tube on ice

➤ Add 37.5 μl of PCR master mix to the reaction tube from step of the reverse transcription

➤ Gently pipette the entire volume, centrifuge briefly, and then place the tube on ice. The total volume is 50 μl

➤ Place the tube on the preheated thermal cycler and incubate the tube using the following PCR cycling conditions (thermal cycler preheat lid set to 100°C)

98°C for 30 seconds

11 cycles of: 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 15 seconds

72°C for 10 minutes

Hold at 4°C

3) The protocol used for CleanTag™ Ligation Kit from TriLink BioTechnologies is composed by different steps and here is reported:

STEP 1: 3' Adapter Ligation to RNA Template

➤ For RNA inputs less than 1000 ng, CleanTag™ 3' Adapter has to be diluted. For the very low amount of RNA extracted, less than 1 ng, adapters were diluted 1:12

➤ Heat 2.5 μl RNA template at 70°C for 2 minutes and place immediately on ice

➤ Prepare reaction containing all components for a total volume of 7.5 μl:
  – 0.5 μl CleanTag™ 3' Adapter
  – 1 μl RNase Inhibitor

38

- 1 µl Enzyme 1
- 5 µl Buffer 1

➢ Add 7.5 µl reaction to 2.5 RNA µl and mix gently

➢ Place the tubes into a thermal cycler with a heated lid and perform the cycling conditions:

28°C for 1 hour

65°C for 20 minutes

➢ Place 3' tagged RNA template on ice.

STEP 2: 5' Adapter Ligation to Tagged RNA Template

➢ For RNA inputs less than 1000 ng, CleanTag™ 5' Adapter has to be diluted. For the very low amount of RNA extracted, less than 1 ng, adapters were diluted 1:12

➢ Heat CleanTag™ 5' Adapter at 70°C for 2 minutes and place immediately on ice

➢ Add CleanTag™ 5' Adapter to each reaction tube, mix by pipetting up and down

➢ Place the tubes into a thermal cycler with a heated lid and perform the cycling conditions:

28°C for 1 hour

65°C for 20 minutes

➢ Place tagged RNA library on ice

STEP 3: Reverse Transcription (RT) Reaction of Tagged RNA Library

➢ Set up the reaction as follow (some reagents from SuperScript III Reverse Transcriptase, ThermoFischer):

- 2 µL RT Primer (14.4 µM)
- 1.92 µL De-ionized Water
- 5.76 µL RT Buffer (5X)
- 1.44 µL dNTPs (10 mM)
- 2.88 µL DTT (100 mM)
- 1 µL RNase Inhibitor
- 1 µL RT Enzyme (200 U/µL)
- 20 µL Tagged Library

➢ Place in thermocycler with a heated lid at 50°C for 1 hour

STEP 4: PCR Amplification of RT Product

Index Primers Set1 were from Illumina (Primers 1-12 with RT and Forward PCR Primer) and Master Mix from New England Biolabs.

- ➢ Set up reaction as follow
    - – 40 μl Q5 High Fidelity 2X Master Mix
    - – 2 μl Forward PCR Primer (20 μM)
    - – 2 μl Index PCR Primer (20 μM)
    - – 36 μl RT Reaction Product
- ➢ Place into a thermal cycler with a heated lid. Perform the following cycling conditions using the recommended number of PCR cycles, for RNA inputs less than 1 ng 21 cycles are recommended:

  98°C for 30 seconds

  21 cycles: 98°C for 10 seconds; 60°C for 30 seconds; 72°C for 15 seconds

  72°C for 10 minutes

  hold at 4°C

The CleanTag™ Ligation was chosen for the preparation of the cDNA library for the smallRNA sequencing study. For each sample a different Index Primer from the set 1 of Illumina was used (Primers #1-12) to run all 12 samples in one lane for the sequencing.

3.2.4 PCR products purification

A purification step after PCR reaction was performed using QIAquick PCR Purification Kit (Qiagen) to remove primers, nucleotides, enzymes, mineral oil, salts, and other impurities from DNA samples. The protocol applied is described below:

- ➢ Add 1:250 volume pH indicator I to Buffer PB. The yellow color of Buffer PB with pH indicator I indicates a pH of ≤7.5
- ➢ Add 5 volumes Buffer PB to 1 volume of the PCR reaction and mix. If the color of the mixture is orange or violet, add 10 μl 3 M sodium acetate, pH 5.0, and mix. The color of the mixture will turn yellow
- ➢ Place a QIAquick column in a 2 ml collection tube

- To bind DNA, apply the sample to the QIAquick column and centrifuge for 30-60 seconds at 17,900 xg at RT
- Discard flow-through and place the QIAquick column back in the same tube
- To wash, add 0.75 ml Buffer PE to the QIAquick column and centrifuge for 30-60 seconds. Place the QIAquick column back in the same tube
- Centrifuge the QIAquick column once more for 1 minute to remove residual wash buffer. Place each QIAquick column in a clean 1.5 ml microcentrifuge tube
- To elute DNA, add 30 μl water to the center of the QIAquick membrane. For increased DNA concentration, let the column stand for 1 minute, and then centrifuge for 1 minute.
- Store DNA at -20°C

3.2.5 cDNA library evaluation

cDNA libraries were quantified by Agilent DNA 1000 Kit using Agilent 2100 Bioanalyzer. The following protocol was applied:

- Setting up the Chip Priming Station replacing the syringe
- Adjust the base plate in position C
- Adjust the syringe clip releasing the lever of the clip and slide it down to the lowest position
- Preparing the Gel-Dye Mix
  - Allow DNA dye concentrate and DNA gel matrix to equilibrate to RT for 30 minutes
  - Vortex DNA dye concentrate and add 25 μl of the dye to a DNA gel matrix vial
  - Vortex solution well and spin down. Transfer to spin filter
  - Centrifuge at 2,240 g for 15 minutes. Protect solution from light. Store at 4 °C
- Loading the Gel-Dye Mix
  - Allow the gel-dye mix equilibrate to RT for 30 minutes before use
  - Put a new DNA chip on the chip priming station
  - Pipette 9 μl of gel-dye mix in the well marked Ⓖ
  - Make sure that the plunger is positioned at 1 ml and then close the chip priming station
  - Press plunger until it is held by the clip
  - Wait for exactly 60 seconds then release clip
  - Wait for 5 seconds and slowly pull back plunger to 1ml position

- Open the chip priming station and pipette 9 μl of gel-dye mix in the wells marked G

➢ Loading the Markers
- Pipette 5 μl of marker in all 12 sample wells and ladder well. Do not leave any wells empty

➢ Loading the Ladder and the Samples
- Pipette 1 μl of DNA ladder in the well marked
- In each of the 12 sample wells pipette 1 μl of sample (used wells) or 1 μl of de-ionized water (unused wells)
- Put the chip horizontally in the adapter and vortex for 1 minute at 2,400 rpm
- Run the chip in the Agilent 2100 Bioanalyzer within 5 minutes

3.2.6 Gel purification

A gel purification step is required to run amplified cDNA library and select only miRs library according to the size. Here is reported the protocol used for the library size selection:

➢ Use Novex TBE Gels, 10% (Invitrogen)
➢ Load 5 μl Quick Load pBR322 DNA-MspI Digest (New England Biolabs) on the first lane as ladder
➢ Mix 8 μl 140bp control with 2 μl Novex Hi-Density TBE Sample Buffer (Invitrogen) and load 10 μl on the second lane
➢ Add 7 μl Novex Hi-Density TBE Sample Buffer to 28 μl sample and load 17 μl in two lanes nearby (leave one lane free between marker and different samples)
➢ Prepare, each time freshly, Running Buffer 0.5X TBE buffer diluted from 5X in nuclease-free water
➢ Run at 120V 2 hours with 0.5 TBE Running Buffer
➢ Remove the gel from the apparatus and stain the gel with 1:10,000 diluted SYBR Gold nucleic acid gel stain (Invitrogen) in TBE 1X, in a clean container for 15 minutes, protected from light on shaker
➢ View the gel on a UV transiluminator and cut corresponding miR band (~130-140bp). Place the gel slice in a 1.5 ml tube with an hole in the bottom

- Centrifuge the tube in a 2 ml tube at 13,200 rpm for 2 minutes. The centrifugation force will make pass through the hole the gel disrupted
- Add 250 μl DNA Gel Elution Buffer and rotate for 2 hours at RT
- Transfer the eluate and the gel debris to the top of a gel filtration column and centrifuge for 2 minutes at 13,200 rpm
- Collect the flow-through, add 1 μl Linear Acrylamide to enhance precipitation, 25 μl 3M sodium acetate, 750 μl 100% ethanol and vortex well
- Precipitate at -80°C overnight
- Vortex and centrifuge for 3 minutes at 14,000 xg at 4°C
- Remove the supernatant taking care not to disturb/remove the pellet
- Add 900 μl 80% Ethanol and centrifuge at 14,000 xg for 30 minutes
- Air dry pellet for up to 10 minutes at RT to remove residual ethanol
- Resuspend pellet in 12 μl TE Buffer
- Load 1 μl on an Agilent Technologies 2100 Bioanalyzer using an Agilent DNA-1000 chip to quantify the DNA concentration after the purification

3.2.7 Sequencing and data analysis

The sequencing was done by Exiqon company on the NextSeq500 platform, using single-end read and 50 nt as number of sequencing cycles. The analysis were performed in collaboration with the same company.

Quality control was done at first calculating the Q-score, that is a prediction of the probability of an incorrect base-call. A Q-score of 30 equals an accuracy of 99.9% for the base-calling. After the sequencing adapters were trimmed off as part of the base calling, showing the profile of the distribution of read length for each donor.

The sequencing reads were mapped to the reference genome reported in miRbase 20 (http://www.mirbase.org/). Mapping of the sequencing reads is the first part of the data analysis but it also represents a useful quality control step in the NGS data analysis pipeline as it can help to evaluate the quality of the samples.

Exploratory analysis was performed on data normalized with tag per million (TPM) method and converted to a log2 scale. TPM is a unit used to measure expression in NGS experiments. The number of reads for a particular miR is divided by the total number of mapped reads and multiplied by 1 million (Tags Per Million). At first the analysis was done by Principal Component Analysis (PCA) and heat map. PCA is a method used to reduce the dimension of large data sets and thereby a useful way to explore the naturally arising sample classes based on the expression profile. The heat map is composed by two-way hierarchical clustering of miRs and samples. For these investigations, top 100 miRs that have the largest variation across all samples were used.

The differential expression analysis is done using the EdgeR statistical software package (Bioconductor, http://www.bioconductor.org/). Differential expression analysis investigates the relative change in expression (i.e. counts) between different samples. The miRs with p values ≤ 0.05 should be considered as statistically significant difference in the expression.

Normalization process was performed by the trimmed mean of M-values, based on log-fold and absolute gene wise changes in expression levels between samples (TMM normalization). This normalization is primarily concerned with compensating for sample specific effects (generally caused by the variation in sequencing depth between samples). Additionally, the normalization step offsets under-sampling effects (due to highly expressed miRs dominating the read set) by identifying scaling factors that minimize log fold changes between samples across the majority of miRs.

3.3 Validation in RT-qPCR

To validate sequencing results, RT-qPCR was performed in Bologna on a bigger cohorts of subjects. In the laboratory in Bologna, miR analysis was developed using different methods respect to Vienna laboratory. RNA is extracted with Total RNA Purification Kit (Norgen), cel-miR-39 (Qiagen) is used as spike-in to monitor RNA isolation, and RT-qPCR is performed through TaqMan technologies.

3.3.1 RNA extraction

Total RNA was isolated from 100 μl of plasma samples using the Total RNA purification kit (Norgen Biotek Corporation). Here is reported the protocol used:

➢ add 350 μl of Lysis solution to 100 μL of plasma and mix by vortexing for 15 seconds

44

- add 5 μl of a 20 fmol/μl cel-miR-39 in water solution (Qiagen)
- add 200 μl of 100% ethanol and mix by vortexing for 10 seconds
- assemble a column with one of the provided collection tubes
- apply up to 600 μl of the lysate with the ethanol onto the column and centrifuge for 1 minute at 14,000 xg
- discard the flow-through and reassemble the column with its collection tube
- apply 400 μl of wash solution to the column and centrifuge for 1 minute at 14,000 xg
- discard the flow-through and reassemble the column with its collection tube
- repeat the washing step twice
- spin the column for 2 minutes at 14000 xg in order to thoroughly dry the resin and discard the collection tube
- place the column in a new 1.5 ml RNase-free tube
- add 50 μl elution solution to the column
- centrifuge for 2 minutes at 200 xg, followed by 1 minute at 14,000 xg
- store RNA at -80°C

3.3.2 RT-qPCR TaqMan technologies

RT-qPCR is based on miR-specific TaqMan probes (Applied Biosystems by Life Technologies) which are oligonucleotides labeled with a fluorophore covalently attached to the 5'-end (e.g. FAM, 6-carboxyfluorescein) and a quencher at the 3'-end. The fluorophore and quencher dyes form a donor-acceptor FRET pair which hybridize to the complementary sequence located downstream one of the PCR primers. When the probe is intact, the proximity of the reporter dye to the quencher dye suppresses the reporter fluorescence. As the primer is extended, the 3'→5'exonucleasic activity of DNA polymerase hydrolyzes the probe from the 5' end separating the reporter from the quencher, therefore allowing detection of the reporter dye fluorescence.

RT-qPCR consists of two parts: RT reaction and PCR reaction.

The RT reaction is the reverse transcription of total RNA into cDNA using a primer specific for each miR. 15 μl of total reaction is composed by:

- 5 μl RNA
- 3 μl specific primer for each miR

- ➢ 7 μL master mix
  - – 0.15 μl dNTP mix (100 mM total)
  - – 1 μl Multiscribe RT enzyme (50U/μl)
  - – 1.5 μl 10x RT Buffer
  - – 0.19 μl RNase Inhibitor (20U/μl)
  - – 4.16 μl Nuclease free water
- ➢ Spin the mixture and apply the following thermal cycling conditions:

  30 minutes at 16°C

  30 minutes at 42°C

    5 minutes at 85°C

    hold at 4°C
- ➢ cDNA is now ready for downstream applications or to be stored at -20°C

In the PCR reaction, cDNA is amplified using specific probes in a total 20 μl reaction contains:

- ➢ 10 μl TaqMan 2X FAST Universal PCR Master Mix
- ➢ 7.7 μl Nuclease free water
- ➢ 1 μl TaqMan MicroRNA Assay 20X
- ➢ 1.33 μl RT reaction product

Thermal cycling conditions (Rotor Gene - Qiagen):

  20 seconds at 95°C

  40 cycles: 1 second at 95°C, 20 seconds at 60°C

3.3.3 Analysis of RT-qPCR results

Data acquired by the instrument is in form of Ct values. These values were normalized with cel-mir-39 measured in each sample, resulting ΔCt values. The relative expression was calculated with the following formula: $2^{-\Delta Ct}$. Relative expression means and standard deviations were calculated for each group of donors. In some cases, one or two samples were excluded when the miR measured in that sample showed a value completely out of range, and it was considered as an outlier in that

evaluation. Since the values were not normally distributed, Kruskal-Wallis test was used to perform statistic evaluation, using SPSS system. P values ≤ 0.05 were considered statistically significant.

miRs expressions and hemato-biochemical parameters from healthy and unhealthy centenarians were evaluated together in a Partial Least Squares regression (PLS-DA) to analyze discrimination parameters of the two phenotypes of centenarians. For this analysis SIMCA-P software was used.

Finally, validated miRs were analyzed using mirPath v2.0, a miR pathway analysis web-server belonging to DIANA TOOLS (http://diana.imis.athena-innovation.gr/DianaTools/index.php). MirPath was used to find experimentally validated targets (DIANA-microT-CDS algorithm) and/or putative targets (DIANA-TarBase v6.0) of validated miRs, as well as to identify relevant common pathways among them.

4. RESULTS

The objectives of the current thesis were two as follows:

1. Optimization of the smallRNA sequencing protocol for the discovery phase. In fact, my personal effort was to acquire new knowledge on NGS technologies to be applied in the study of c-miRs performed with high resolution and technological level in the biological process of human aging and longevity. This work has been done in collaboration with the BOKU University of Natural Resources and Life Sciences in Vienna. For that purpose I have spent 6 months of my PhD in the laboratory of Biotechnology of Skin Aging directed by Professor Johannes Grillari and I have worked with the academic startup TAmiRNA, which is specialized in c-miRs studies. A total of 12 subjects, recruited in Bologna for a discovery phase, were included, i.e. 3 healthy young donors, 3 healthy old donors, 3 healthy centenarians and 3 unhealthy centenarians. These experimental phases will be described in 4.1, 4.2 and 4.3.

2. The identification of c-miRs characterizing the process of human aging and longevity and likely a signature able to distinguish healthy and unhealthy longevity. A total of 48 subjects, recruited in Bologna for a validation phase, were included, i.e. 16 healthy young donors, 16 healthy old donors, 10 healthy centenarians and 6 unhealthy centenarians. This experimental phase will be described in 4.4.

4.1 Optimization of the sequencing protocol

4.1.1 RNA extraction and quantification

RNA extraction was assessed by testing different volumes of EDTA-plasma, i.e.: $0.2 – 0.5 – 1 – 2 – 5$ ml using a plasma pool of healthy donors (test samples). These volumes were processed with 3 kits: i. Plasma/Serum RNA Purification Kit (one for small volumes and another for large volumes); ii. Plasma/serum circulating and exosomal RNA purification kit (both commercialized by the Norgen company); iii. kit of Qiagen, miRNeasy Mini kit, readapted for RNA extraction from plasma. They are all column based methods. After some RNA extraction, Plasma/Serum circulating and exosomal RNA purification kit was chosen as the best among Norgen kits (data not shown) and compared in more detail with Qiagen kit.

RNA extracted was quantified by Bioanalyzer (platform 2100 - Agilent Technologies, Inc.) using the chip for smallRNA analysis. Direct quantification of smallRNA was extremely difficult probably due to a very low concentration of material. Therefore, RT-qPCR was used to estimate RNA extraction. Spike-ins added at the beginning of the extraction, i.e. UniSp2, UniSp4, UniSp5, were analyzed to evaluate RNA extraction, while UniSp6 was added during cDNA preparation to evaluate retrotranscription. Some endogenous miRs (i.e. miR-21; -23a) were assessed too. Following the protocol of different kits, final elution volumes are different, thus RNA was eluted in 30 µl applying Qiagen kit, and it was eluted in 100 µl by Norgen kit.

Results, reported in figure 4.1, show an expected linearity between miR Ct values and plasma volume, except in the case of 5 ml of plasma processed by Norgen kit (figure 4.1, lower panel), where RNA extraction efficiency was affected and RNA quantity was not evaluable.
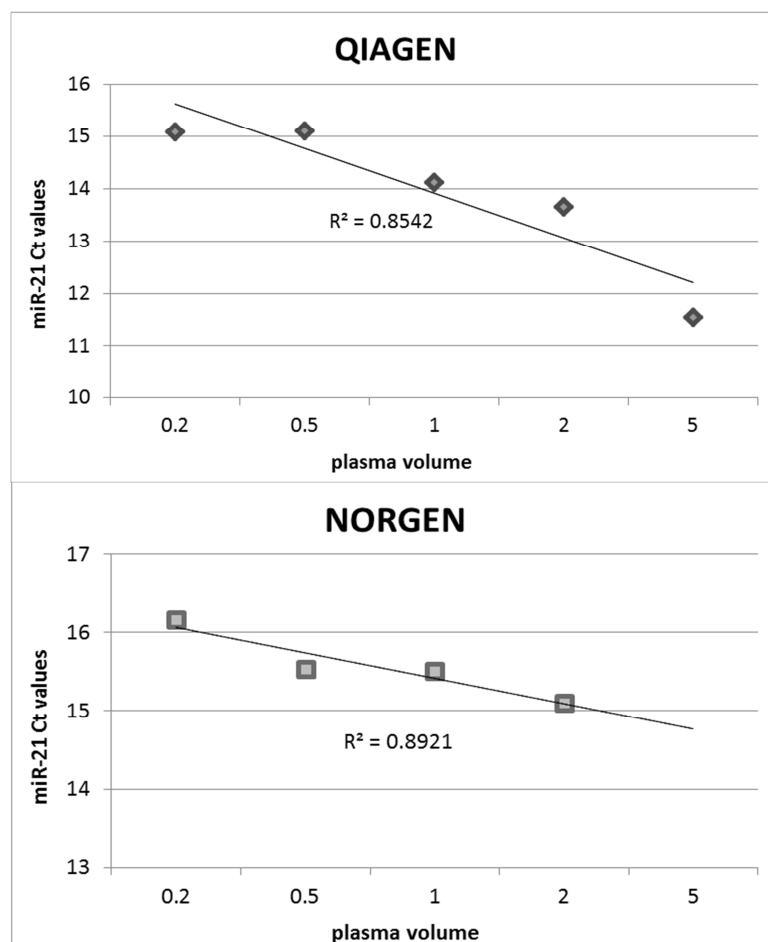

Figure 4.1. RT-qPCR evaluation of extracted RNA samples by two kits (Qiagen, Norgen). Ct values of miR-21 are reported, in different plasma volumes (0.2 – 0.5 – 1 – 2 – 5 ml). The coefficient of determination of the regression analysis is reported in both conditions.

This result suggested that 5 ml plasma gives technical problems, therefore it was excluded for further analysis and 2 ml was chosen as the best volume to achieve the maximum quantity of RNA.

A further optimization of the protocol was focused on reducing RNA elution volume with Norgen kit, expected to be 100 µl. Different elution volumes and precipitation/concentration protocols were tested and the best result was achieved by an overnight precipitation with sodium acetate, glycogen and ethanol, at final elution of RNA in 15 µl of solution.

On the whole, both kits showed good results in terms of RNA extraction and for that reason it was decided to test both to prepare cDNA library.

4.1.2 cDNA library preparation

Three commercial kits for cDNA library preparation were tested as follows:
- XRNA Exosome RNA-seq NGS library prep kit (System Biosciences)
- TruSeq Small RNA library prep kit (Illumina)
- CleanTag Ligation Kit (TriLink BioTechnologies)

The three kits differ on the strategy applied to prevent adapter dimers formation. Specifically, System Biosciences kit adopt the method of cleaning up with magnetic beads the product after the 3' ligation step. Illumina applies a T4 RNA ligase deletion mutant for the 3' ligation and the full-length enzyme for 5' ligation. TriLink BioTechnologies kit applies chemically modified adapters to reduce adapter dimers formation.

In all experiments, amplified products after PCR were purified with QIAquick PCR Purification Kit (Qiagen) to removes impurities, such as primers, nucleotides, enzymes, mineral oil, salts. Agilent DNA 1000 Kit was used at Bioanalyzer to quantify library preparation.

Results of cDNA library preparation are reported in the electropherograms as figure 4.2. on the bases of different kits. Illumina kit produced only adapter dimers, which can be identified by 121 bp peak. Better results were obtained with System Biosciences kit and a little quantity of library was reached (140-150 bp) in the sample of RNA isolated with Norgen kit. No library was found with Qiagen RNA. Some attempts were done on the optimization of PCR step in library preparation, but any significant improvement were obtained in terms of cDNA library quantity. On the contrary,

TriLink kit showed a good quality in library preparation, and electropherogram shows a clear peak of the library at 132 bp, while only a little quantity of adapter dimers (peak at 116 bp) is evidenced.
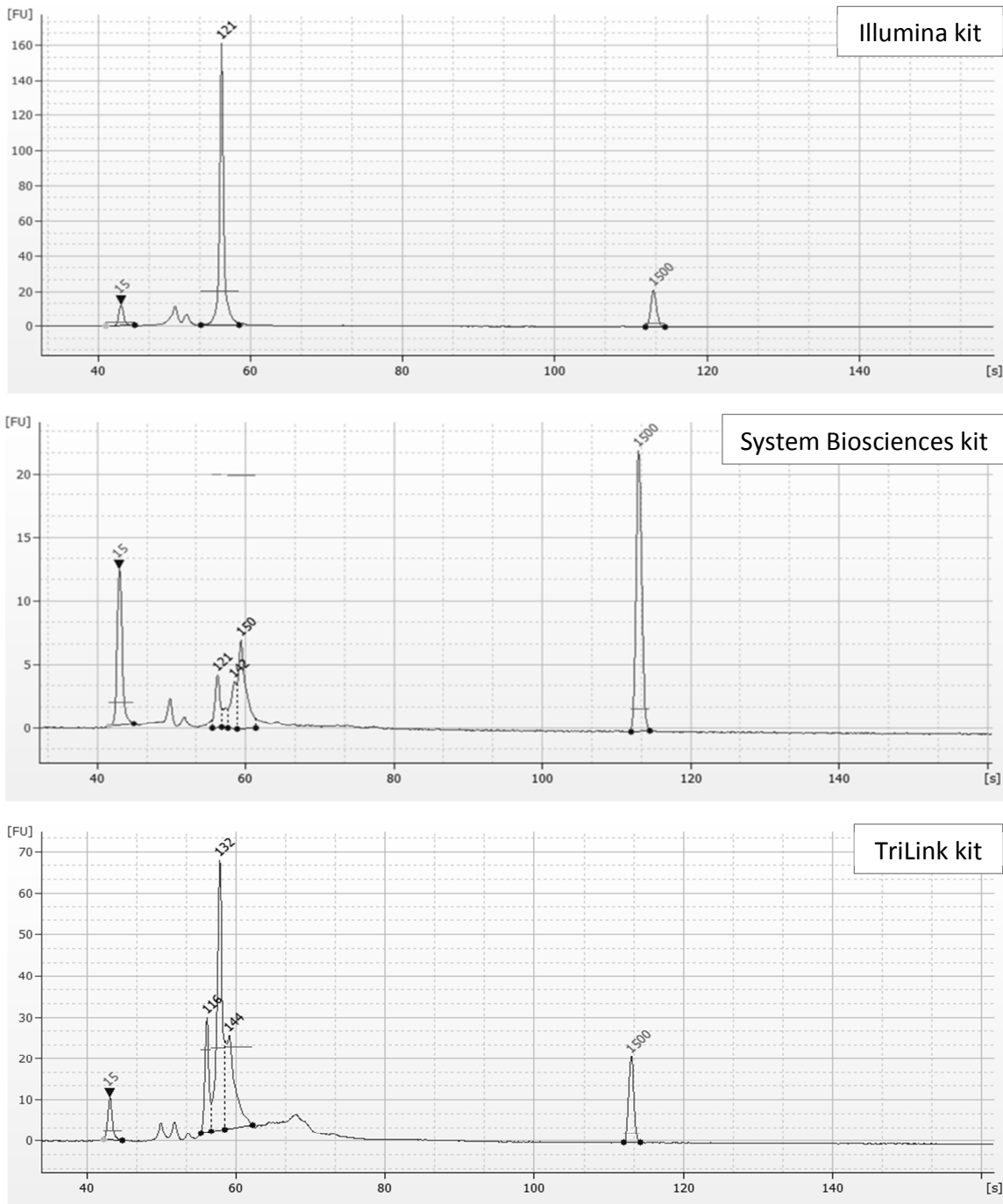

Figure 4.2. Library preparation evaluated by Bioanalyzer with three different kits. TruSeq Small RNA library prep kit (Illumina) in the upper panel, XRNA Exosome RNA-seq NGS library prep kit (System Biosciences) in the middle panel, CleanTag Ligation Kit (TriLink BioTechnologies) in the lower panel. cDNA library is about at 130-140 bp peak and it depends on adapters length.

4.1.3 Library size selection

Accordingly to the best results obtained, libraries from TriLink kit were processed and purified on gel. After a run on agarose gel, bands corresponding to cDNA library were cut with a sterile scalpel blade, then DNA trapped in gel fibers was recovered and analyzed with Bioanalyzer. TriLink kit gave a plentiful recovery of libraries, as reported in figure 4.3, especially using RNA sample extracted by Qiagen kit, resulting the best molarity of 360 nM, instead of 300 nM obtained by Norgen sample.
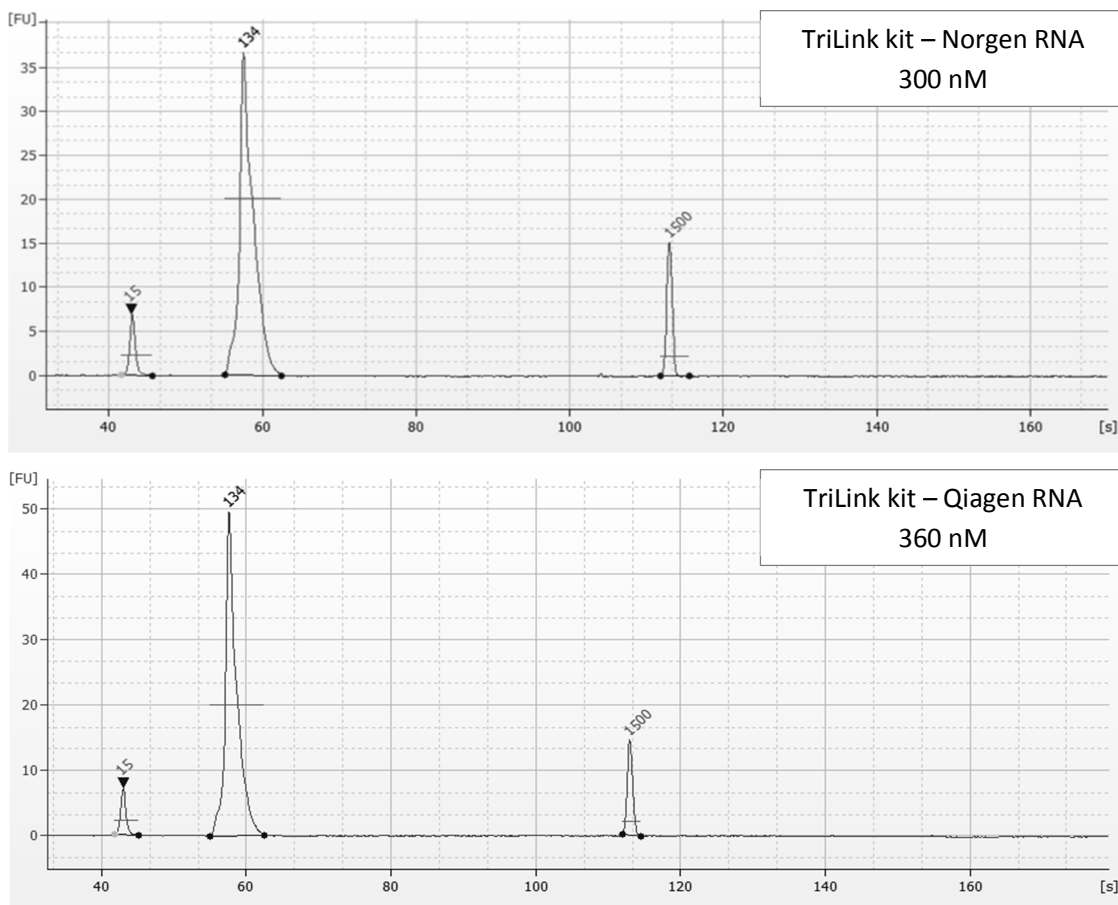
Figure 4.3. TriLink cDNA libraries analyzed by Bioanalyzer. Upper panel: library prepared with RNA extracted by Norgen kit. Lower panel: library prepared with RNA extracted by Qiagen kit. Libraries are evidenced by the peak at 134 bp.

At this stage, the optimization phase was over and the final protocol was established as follows:

- 2 ml of EDTA-plasma sample

- miRNeasy Mini kit (Qiagen) for RNA extraction

- RT-qPCR to evaluate RNA isolation

- CleanTag Ligation Kit (TriLink BioTechnologies) for cDNA library preparation

- Libraries evaluated by Bioanalyzer before and after their size selection on agarose gel

4.2 Processing samples for smallRNA sequencing

The above protocol was applied to analyze miR profiling in three different aged groups using a discovery phase design (12 subjects as described previously). Plasma of the donors was collected in Bologna and processed in Vienna.

4.2.1 RNA extraction evaluation

2 ml plasma from each donor were used to extract RNA with miRNeasy Mini kit (Qiagen). RNA was evaluated by RT-qPCR, assessing the following molecules: i. 2 spike-ins added during the extraction, i.e. UniSp4 and UniSp5; ii. 1 spike-in added in retrotranscription phase, i.e. UniSp6; iii. the endogenous miR-21. Results, reported in figure 4.4, show that 1145M sample and mostly 1141M and 1142M samples were not properly processed, as miR-21 Ct values resulted very high (meaning low amount of materials) and no amplification was found of the RNA extraction-related spike-ins. UniSp6 does not show any variation therefore the conclusion is that the problem of low RNA level has come out during RNA extraction .
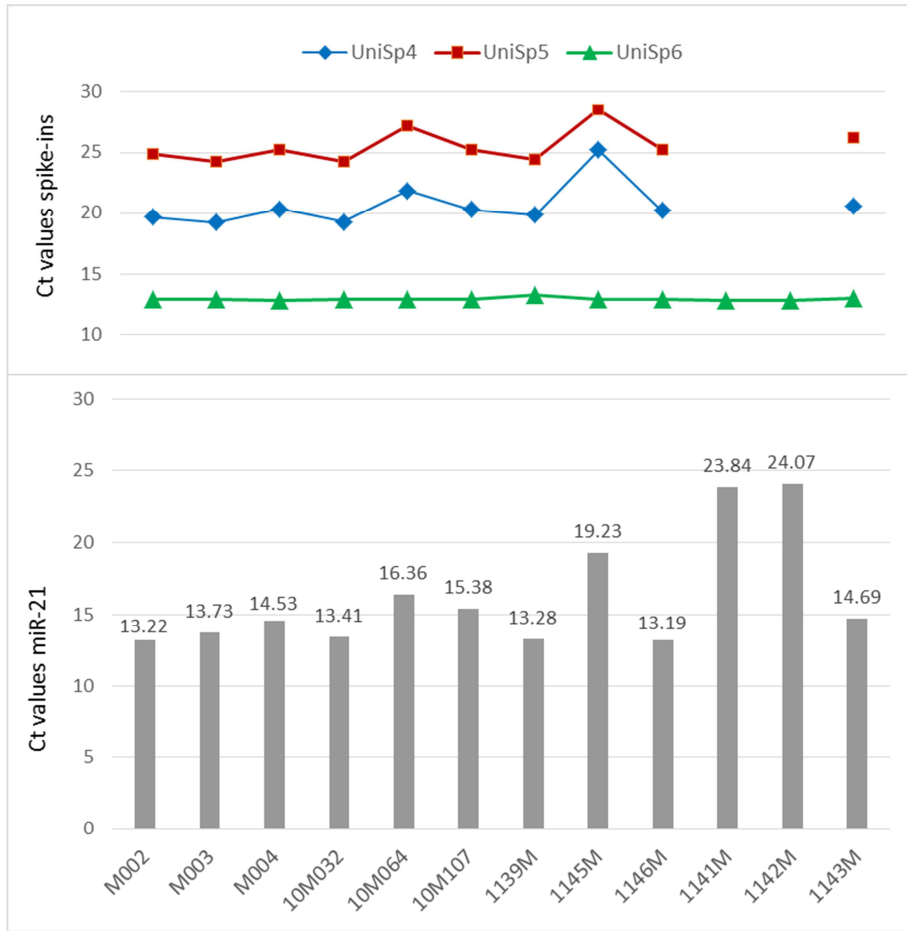
Figure 4.4. RT-qPCR analysis on RNA extracted from the 12 donors samples. Ct values of the spike-ins are shown in the upper panel, Ct values of the endogenous miR-21 is shown in the lower panel.

As two out of three failed samples, i.e. 1141M and 1142M are unhealthy centenarians while 1145M sample is from healthy centenarian, it could be hypothesized that something present in centenarian samples, such as proteins or lipids, interfered with RNA isolation. To further investigate this problem, one smaller aliquot for each subjects, 200 µl of plasma, was used to perform RNA isolation twice. As showed in figure 4.5, miR-23a was further detected and compared in two different RNA extracted samples (200 µl and 2 mL volumes).
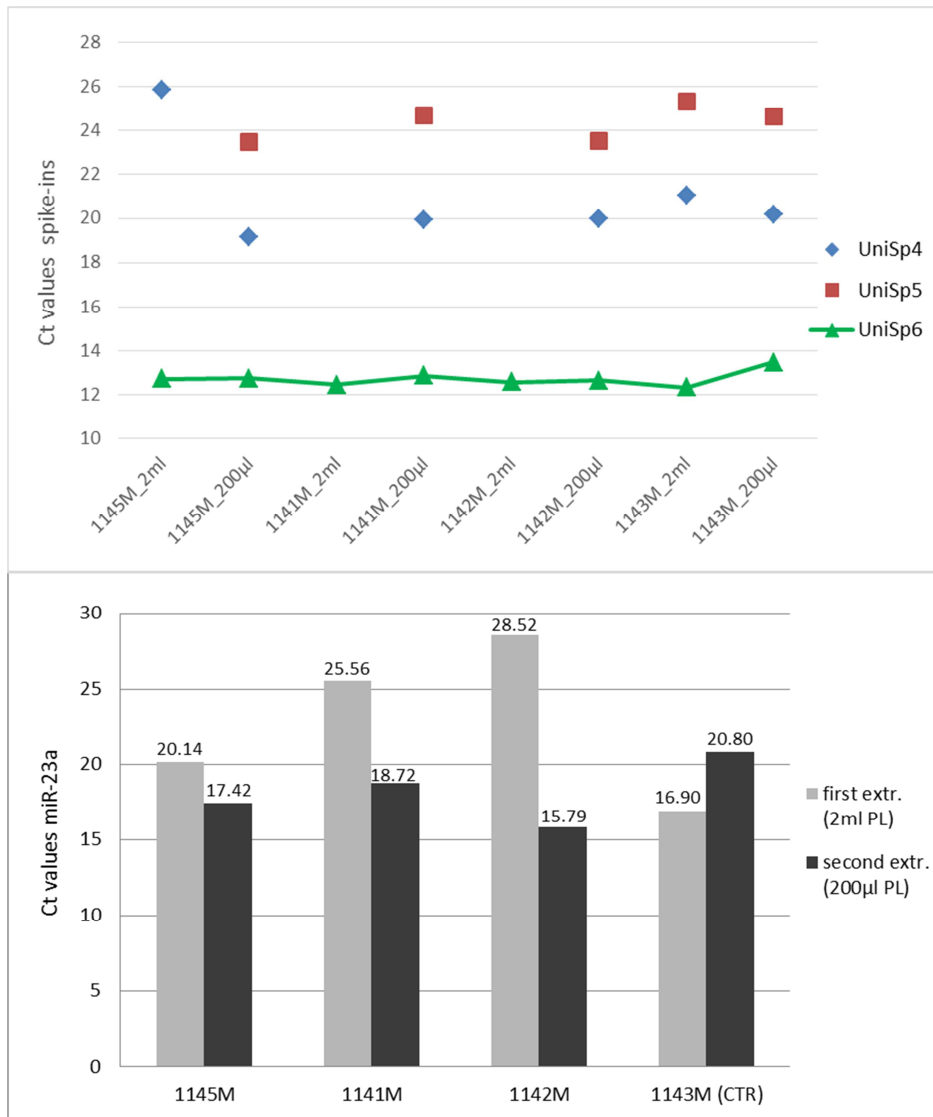
Figure 4.5. RT-qPCR analysis on RNA obtained from the first (2 ml plasma, PL) and the second extraction (200 µl plasma). Ct values of the spike-ins are shown in the upper panel, those of the endogenous miR-23a are reported in the lower.

RNA isolated from 200 µl of plasma resulted the best condition for higher efficiency of extraction (smaller Ct values) and 1143M sample, used as reference for the best extraction, confirmed that the quantity of the endogenous miR is proportional to the starting material. To explain why the extraction was successfully completed starting from less quantity of plasma is not easy, but results suggest that something in the samples (membranes, proteins, antibodies, etc.) might interfere with the extraction mostly if a large amount of material is used.

4.2.2 cDNA library preparation

RNA isolated from the first (2 ml) samples handling and the second one (200 µL, only for samples: 1145M, 1141M, 1142M) were processed with CleanTag Ligation Kit (TriLink BioTechnologies) to prepare cDNA libraries. Results obtained by running the libraries on Bioanalyzer, showed a good production of cDNA libraries which are evidenced in figure 4.6, as peaks (with circles) related to bp length).
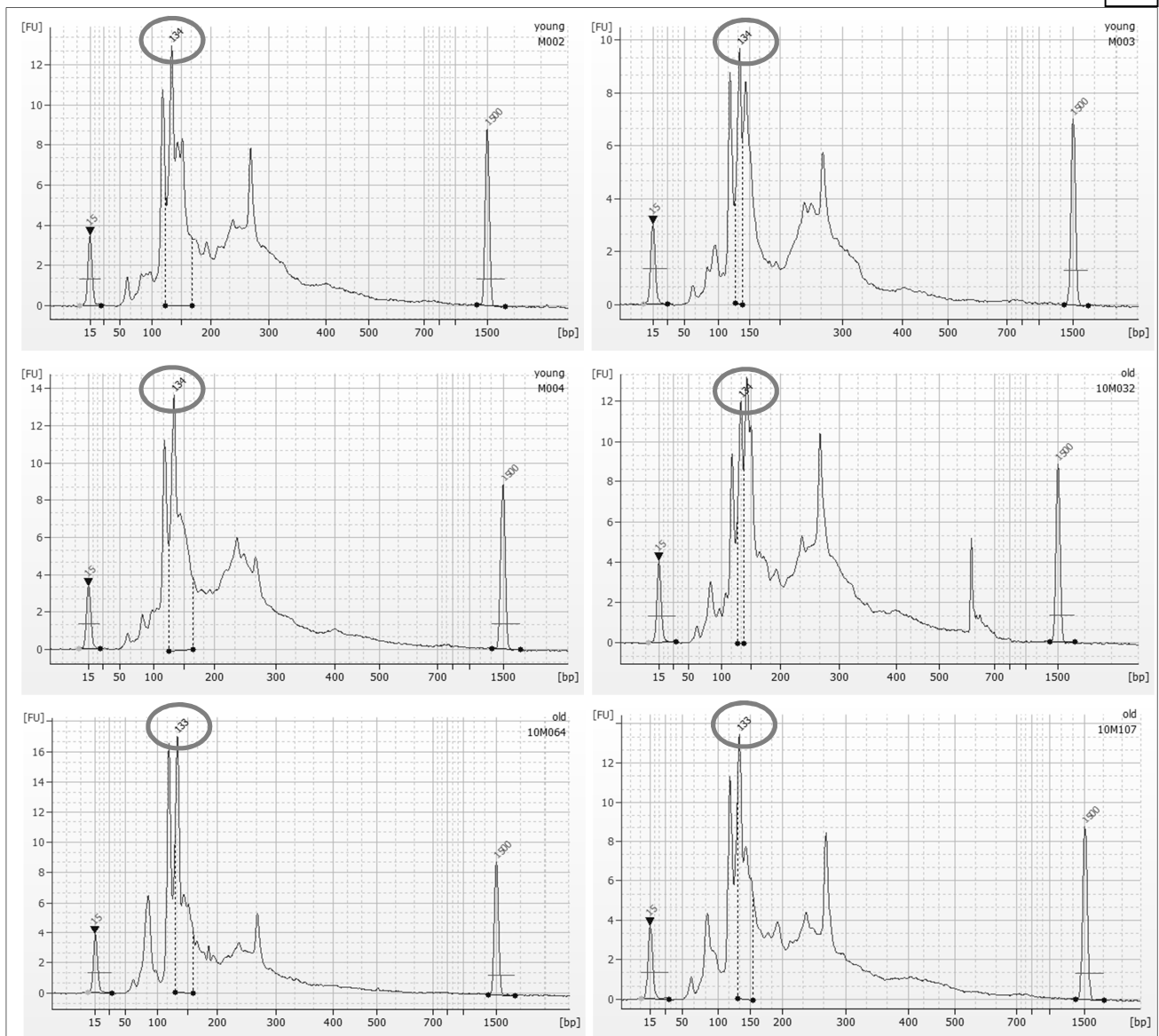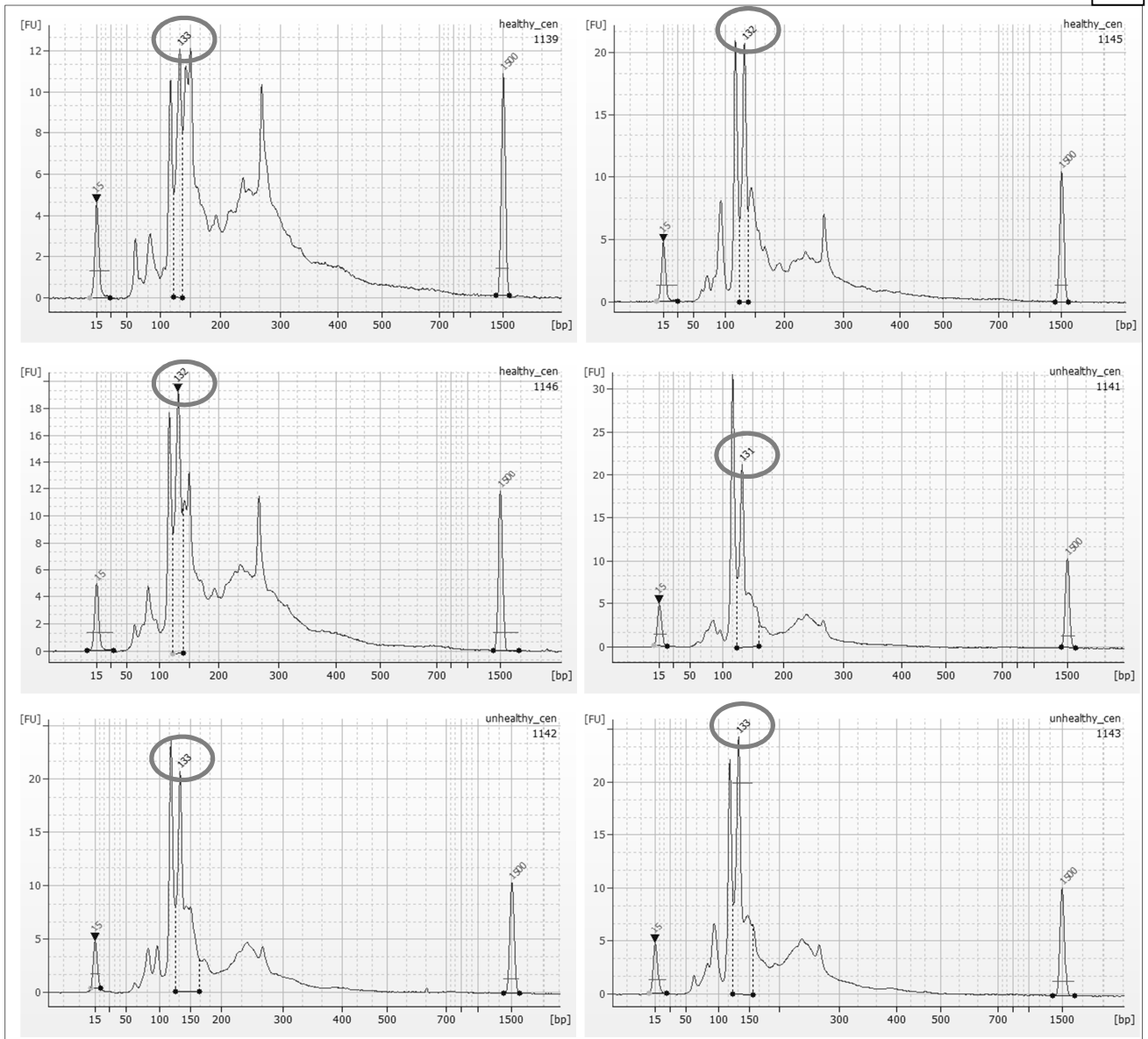
Figure 4.6. cDNA libraries after detection by Bioanalyzer. Electropherogram for each subject is reported, 3 young donors and 3 old donors are in panel A, 3 healthy centenarians and 3 unhealthy centenarians are in the panel B. The cDNA library is circled and it correspond to the peak at about 130 bp.

4.2.3 Library size selection

Libraries were loaded on agarose gel and were selected for purification according to their size. After gel purification, the libraries were recovered with the protocol previously described, and their concentrations are reported in the Table I.

Table I. Concentration and molarity for each library. After the size selection on agarose gel, quantification of the recovered libraries was calculated with Bioanalyzer.

	Concentration [ng/µl]	Molarity [nmol/l]
M002	18.17	205
M003	4.1	43.2
M004	20.97	235.5
10M032	29.17	308.5
10M064	17.17	195.8
10M107	18.41	208.3
1139M	16.98	177.2
1145M	17.45	195.6
1146M	15.08	162.5
1141M	21.96	249.6
1142M	17.6	198.6
1143M	26.95	304.3

Library obtained from M003 sample showed the lowest molarity, likely due to a problem of loading on agarose gel. Anyhow 43.2 mM were sufficient to run the sequencing.

Libraries were pooled in an equimolar ratio according to the lowest concentration, i.e. 43.2 nM, in order to load all the samples into a single lane. The pool was further tested with Bioanalyzer to verify the final molarity, which resulted to be 38.1 nM, as reported in figure 4.7.
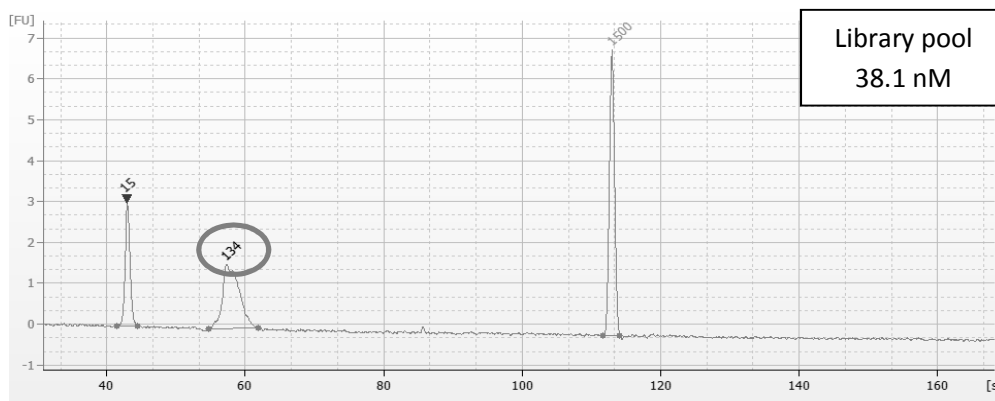

Figure 4.7. Library pool analyzed on Bioanalyzer. The 134 bp peak correspond to the library pool, and the final molarity was 38.1 nM.

The pool was sent in Denmark to Exiqon company where the sequencing was performed by Illumina technology.

4.3 Sequencing and data analysis

SmallRNA sequencing was performed by Exiqon company using Illumina NextSeq500 platform, adopting single-end read and 50 nt lenght. Sequencing data analysis was also done in collaboration with Exiqon.

4.3.1 Data quality control

First of all, smallRNA sequencing was evaluated calculating the Q-score (or quality score). The samples showed overall good data quality, presenting Q-score higher than Q30, meaning an accuracy of 99.9% of base-calling.

After sequencing, adapters were trimmed off as part of the base calling. Trimming of adapters from the dataset revealed one distinct peak (~9-11 nt) representing probably degraded RNA and a minor miR (~18-22 nt) peak. The abundance of miRs varies from sample to sample. Some reads were longer sequences likely due to other origin (i.e. rRNA, tRNA and mRNA fragments, ~30+ nt). In figure 4.8, the profiles of the distribution of read length from one donor of each age group are reported, since they were really similar in according to each age group.
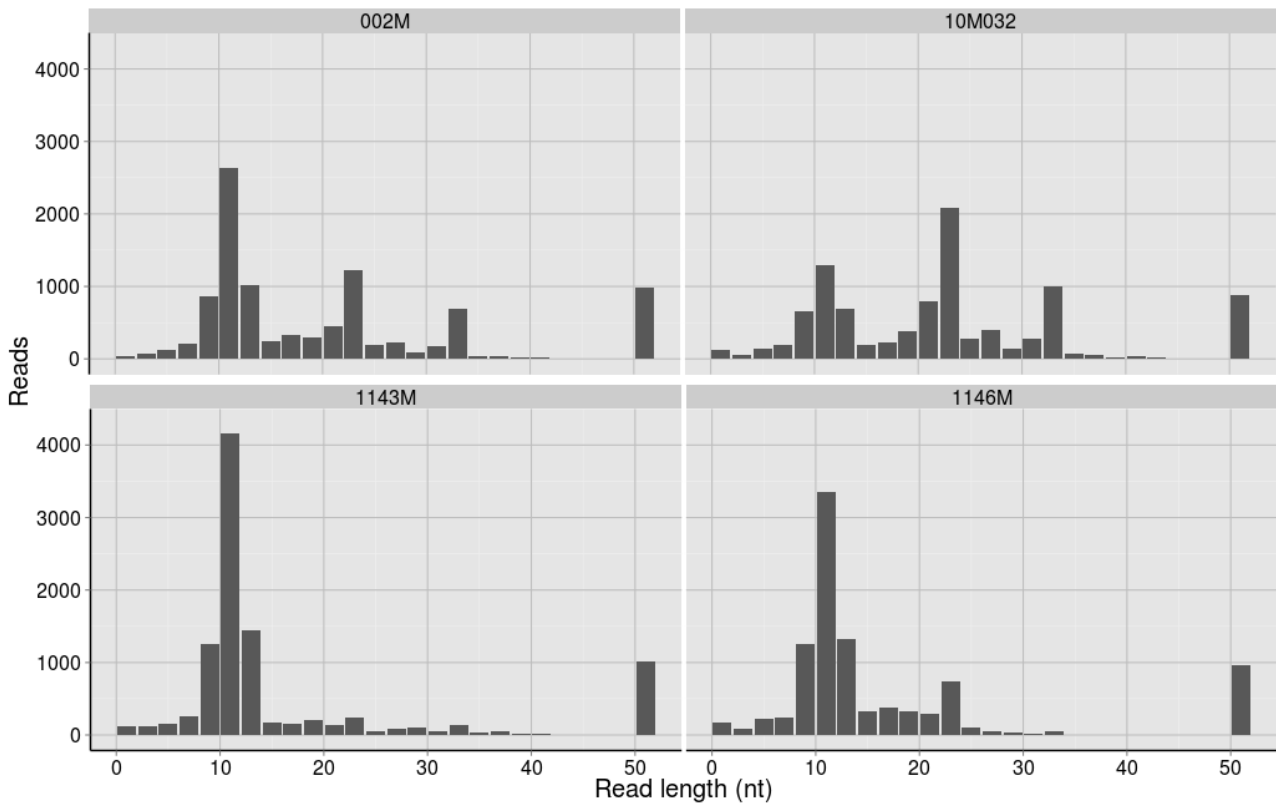
Figure 4.8. Read length distribution after adapters filtering. All the 12 samples have a main peak around 10 nt, representing probably degraded RNA. One sample for each group are reported, M002 young, 10M032 old, 1143M unhealthy centenarians, 1146M healthy centenarians.

Mapping of the reads was done as first part of the data analysis but it also represents a useful quality control step. For this purpose, the reads were classified in the following classes:

- Outmapped reads: For example polyA and polyC homopolymers
- Unmapped reads: no alignment possible to host genome
- Genome reads: aligning to reference genome, but not to smallRNA or miR
- miRNA: Maps to used version of mirBase
- SmallRNA: Maps to smallRNA database
- Predicted (pred) miRs: two different kinds of predicted miRs are present, i.e sequences belonging to another genome (other organism) or sequences predicted to be human hsa-miR (mirBase).

A typical miR sequencing experiment yields approximately 10-60% miRs mapping to the reference genome. However, this is dependent on the quality of the sample and how well of the reference genome is characterized, as well as the miR annotation in miRBase 20. If the sample is degraded fewer reads will be miR specific and more material will be assigned as rRNA/tRNA. The plot showed in figure 4.9, summarizes the overall mapping results and it evidences that the fraction of mappable reads (=counts) is within the normal range but the overall distribution varies due to the 10 nt peak content.
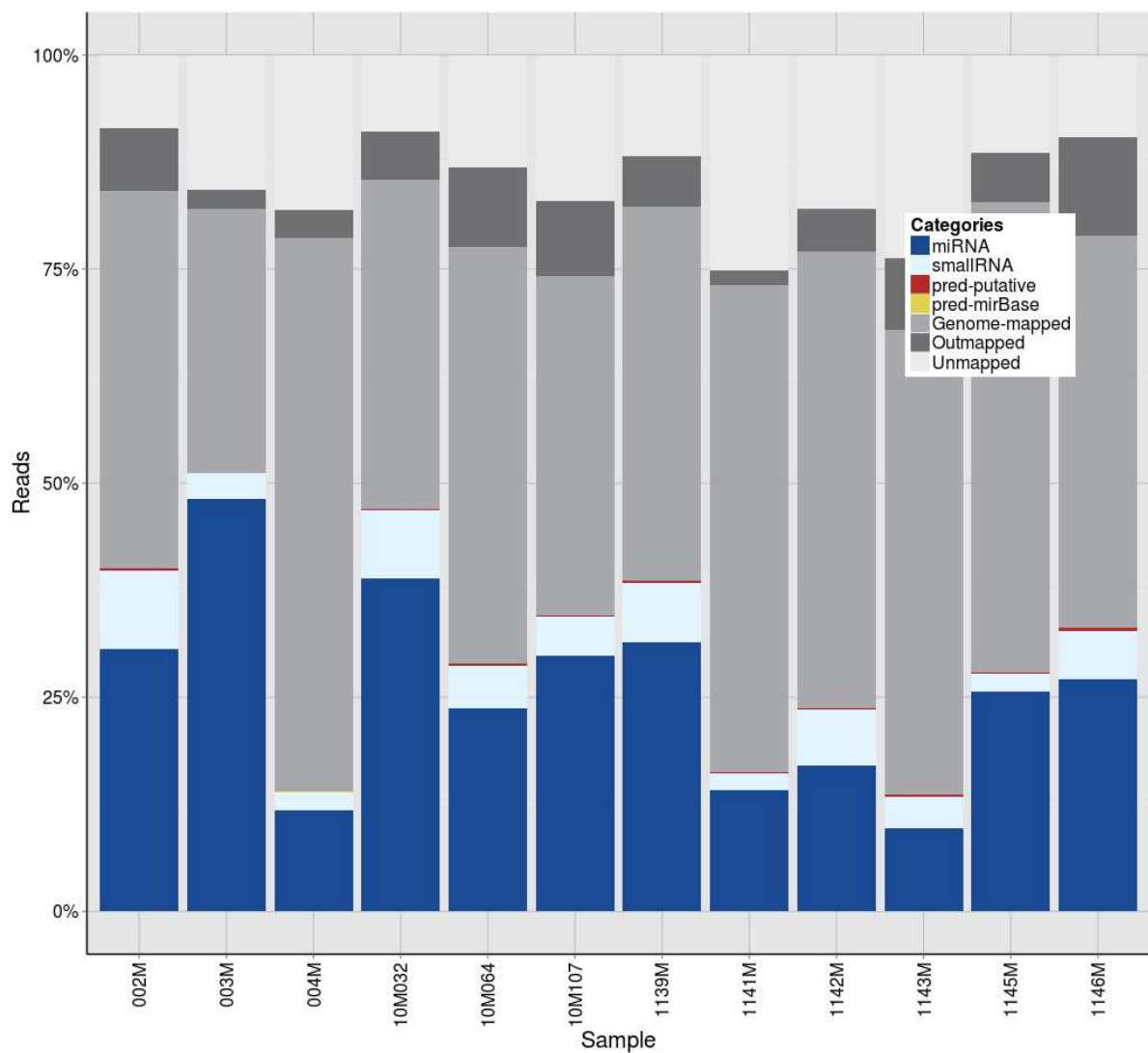

Figure 4.9. Summary of mapping of sequencing read counts by sample. Each category is reported in a different color.

4.3.2 Results of sequencing

SmallRNA sequencing profiling was successfully completed. An average of 9.6 million reads were obtained per sample and the total number of reads per sample is shown in the figure 4.10.
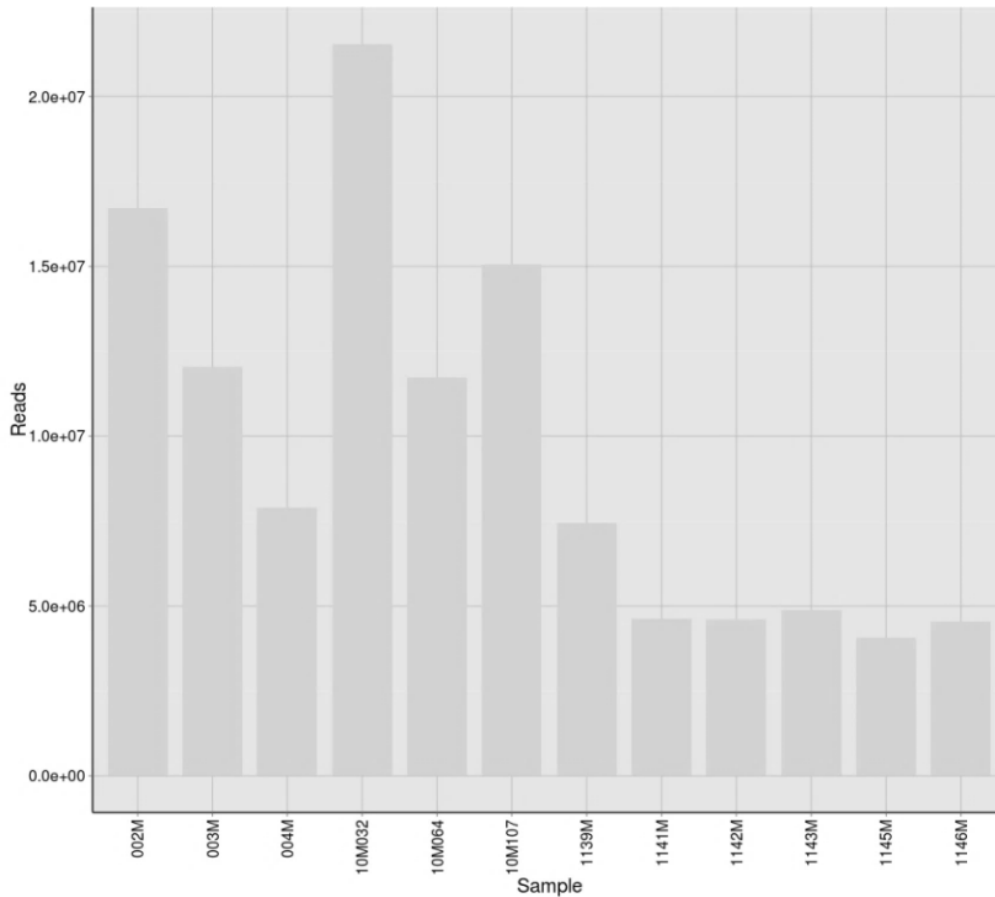

Figure 4.10. Total number of reads sequenced for each sample. An average of 9.6 million reads was obtained.

After mapping and counting reads the numbers of known miRs, in respect to relevant entries in miRBase 20, was calculated. The reliability of the identified miRs increased with the number of identified fragments. Expression levels were measured as Tags Per Million (TPM) which is an unit used to measure miR expression in NGS experiments. The number of reads for a particular miR is divided by the total number of mapped reads and multiplied by 1 million (Tags Per Million), and the number of identified and known miRs is reported in figure 4.11 for each sample, taking into account two different level of expression (1 or 10 TPM). Among all samples, the number of shared miRs having ≥ 1 TPM was 261 and having ≥ 10 TPM was 123 miRs.
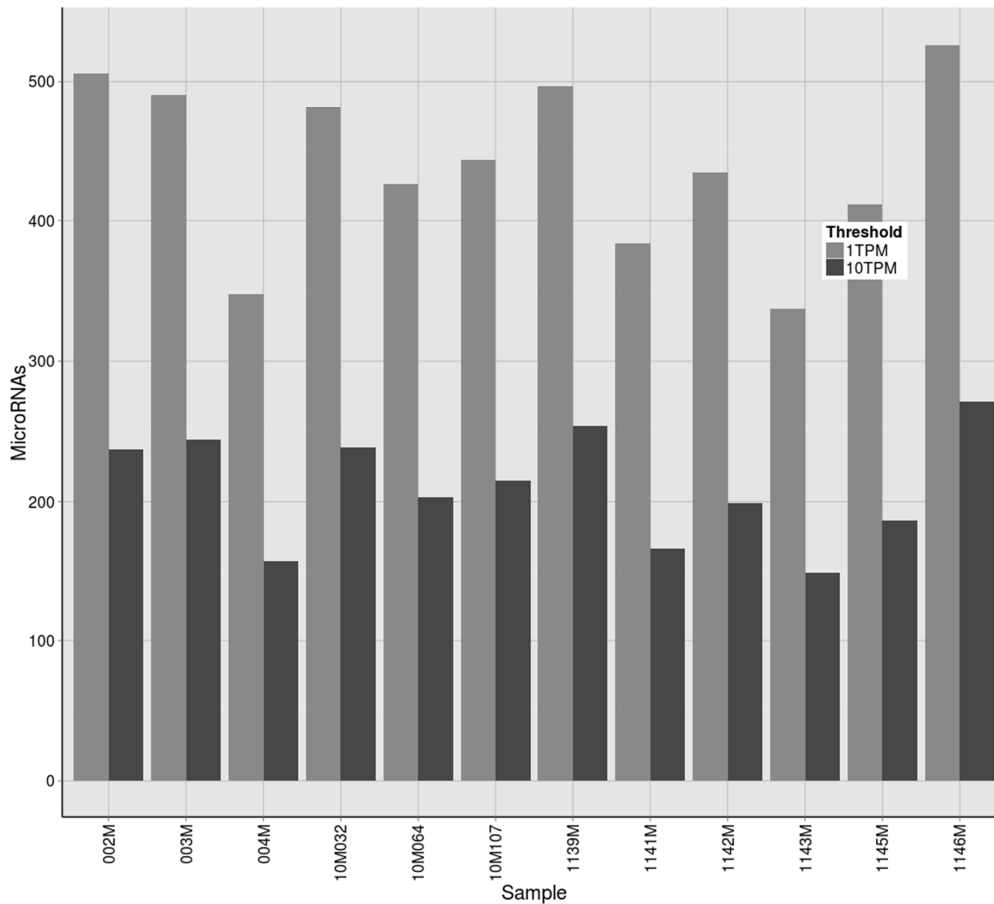
Figure 4.11. Number of identified and known miRs for each donor, with number of counts. >1TPM per sample in grey bars and >10TPM per sample in black bars.

All data were normalized with tag per million (TPM) method and converted to a log2 scale. Exploratory analysis was performed by PCA and heat map. To obtain the best optimization of the statistical analysis the top 100 miRs with the largest variation across all samples were selected, and the components of sample variance (PC1; PC2) were obtained, as shown in figure 4.12. The largest component of variation (PC1) is plotted along the X-axis and the second largest is plotted on the Y-axis (PC2).
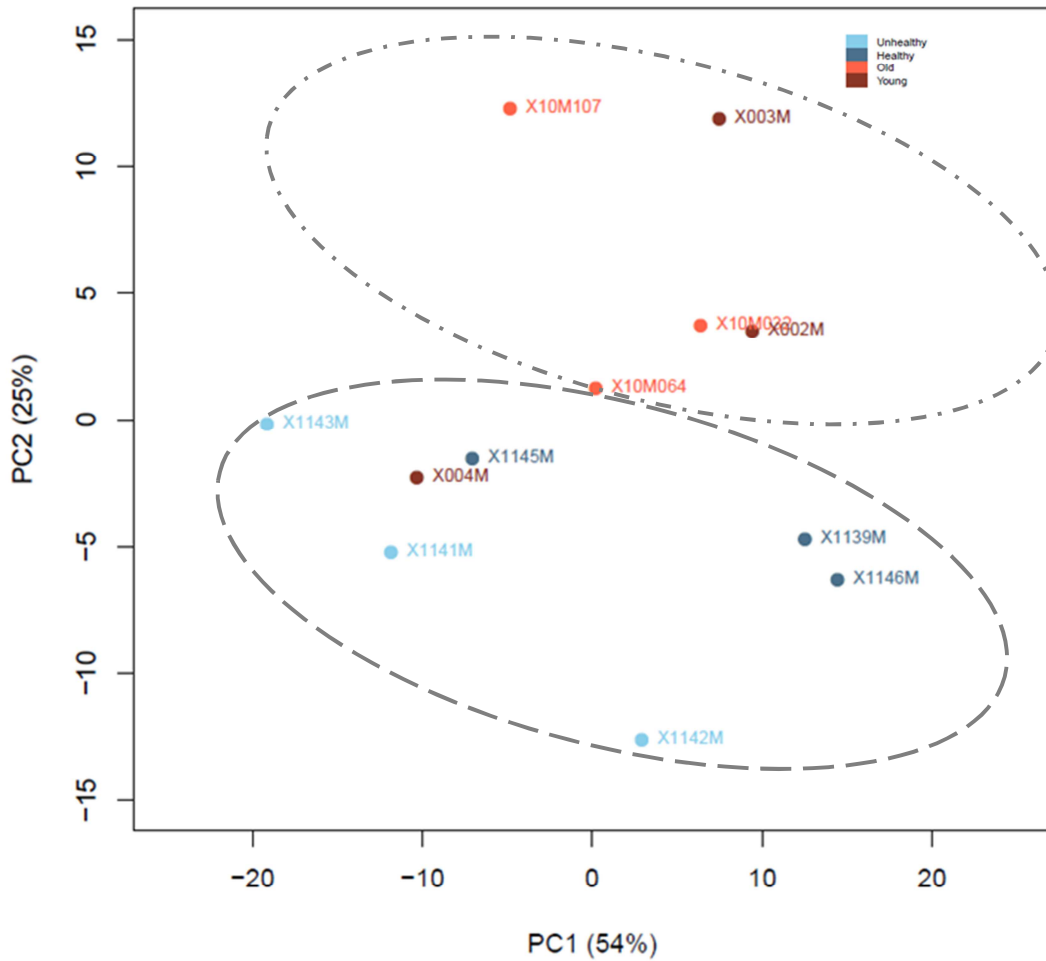
Figure 4.12. Principal component analysis (PCA) plot. PCA was performed on all samples using the top 100 miRs with highest Coefficient of Variation (CV, based on TPM normalized reads). Two circled clusters are shown .

Two clusters raise up suggesting that samples can correspond to two different groups, i.e young-old subjects and healthy-unhealthy centenarians, except for sample 004M (young subject) which does not cluster with its related group of young-old subjects. The most important conclusion is that 004M is likely an outlier and it should not be considered in the further analysis.

The figure 4.13 shows the heat map diagram and two-way hierarchical clustering of miRs and samples. Each row represents a specific miR and each column represents a sample. Matrix colors represent the relative expression level of the miRs across all samples. Three subclusters emerge from the dendogram analysis, one of them corresponds to young-old subjects while the others two

subclusters correspond to unhealthy-healthy centenarians, except 004M which is confirmed to be an outlier sample.
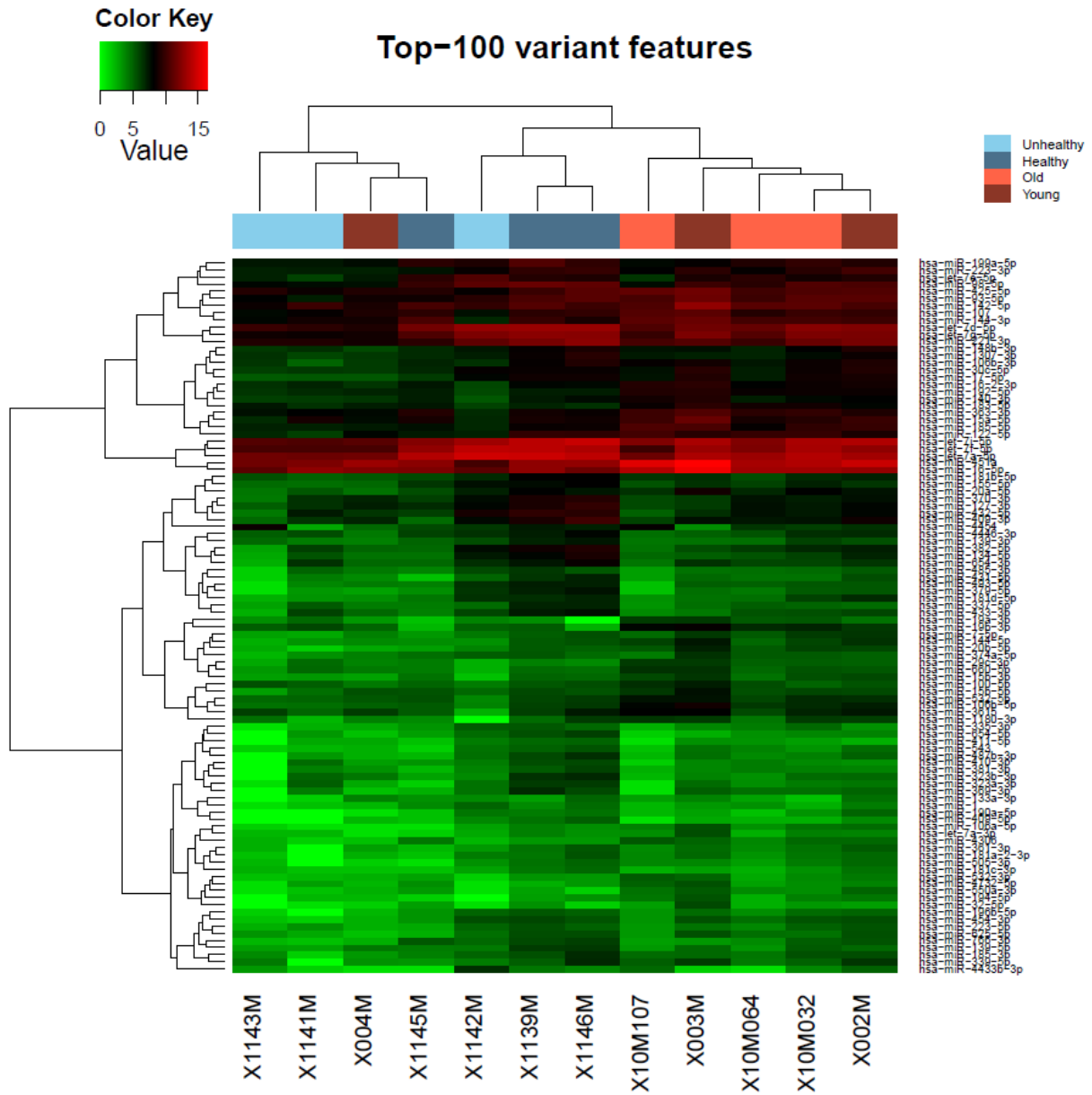

Figure 13. Heat Map and unsupervised hierarchical clustering by sample and miR. The clustering was performed on all samples, and on the top 100 miRs with highest CV based on TPM normalized counts.

Alignment and mapping allowed to identify known or novel (not annotated) human hsa-miRs. To discover novel miRs, the bioinformatic analysis was done searching in mirBase in different genomes, such as human, mouse and rat and other organisms. The result was that only one miR sequence corresponded to *Ornithorhynchus anatinus* species, oan-miR-1386. This interesting sequence needs to be further investigated, but currently it is out of the scope of the present thesis.

Further, putative novel miRs were checked mapping all organisms found in miRbase, or other known RNA sequences (mRNA, tRNA etc). miRPara was used to analyze the potential folding of these last sequences. All these analyses were iterated to identify putative novel miRs and a total of 91 putative-miRs were recognized.

Another type of analysis was performed to identify eventually IsomiR, i.e. variations of miRs annotated in miRBase at the 3' and 5' UTRs and, with minor frequencies, variations of nucleotide substitution along the miR length. Only isomiRs counted at 5% level of total reads for a specific miR were merged to generate a single count file. For example, 15 isomiRs were found for hsa-let-7b-3p, reaching 81 total counts for this miR. On the whole, miRs reads contains also isomiR counts.

4.3.3 Differentially expressed miRs in discovery phase

P-values for significantly differentially expressed miRs were estimated by the exact test on the negative binomial distribution.

As first approach, the differences between the two groups of healthy and unhealthy centenarians were considered with the purpose to study the longevity process and likely to identify a signature. The table II shows miRs differentially expressed between healthy and unhealthy centenarians, only 13 miRs reached the statistical significance threshold (p value ≤ 0.05).

Table II. The 13 significantly differentially expressed miRs between healthy and unhealthy centenarians. MiRs list with log fold change between the two groups, p-values and the average count values per group.

	logFC	PValue	HC avg reads	UHC avg reads
miR-19b-3p	2.74	0.0011	16.67	44.67
miR-19a-3p	2.78	0.0014	6.00	19.67
miR-1226-3p	-3.15	0.0031	3.33	0.00
miR-598-3p	-1.98	0.0069	13.33	1.33
miR-887-3p	2.84	0.0094	0.33	1.67
miR-145-5p	2.11	0.0102	40.33	66.00
miR-4433b-3p	2.43	0.0112	13.00	39.67
miR-296-5p	2.78	0.0117	0.33	1.67
miR-30a-5p	2.07	0.0167	1511.33	3411.33
miR-10b-5p	1.89	0.0217	220.33	375.33
miR-873-3p	1.99	0.0331	1.00	2.00
miR-487a-3p	-1.91	0.0502	3.67	0.00
miR-766-3p	-1.38	0.0539	58.00	13.00

LogFC: log Fold Change. pvalue: p value. HC: healthy centenarians. UHC: unhealthy centenarians. AVG reads: average of the reads count in each group.

After the identification of differentially expressed miRs between the two groups of centenarians a further selection of miRs was performed taking into account the read counts and excluding miRs with lowest values (near 0) meaning an extremely low expression (difficult to be validated).

A further selection was based on the miR trend along the different age groups in order to identify only those interesting for aging trajectories

Specifically, miR-598-3p and miR-766-3p were selected for a decreasing trajectory during normal aging (across young, old and unhealthy centenarians) while in healthy centenarians miR levels restored at the level of young subjects. MiR-30a-5p showed an increasing trend in the different aged groups but not in healthy longevity. In the table III the three most interesting miRs are reported and they were selected for validation phase. Accordingly with pre-processed data analysis sample 004M in the young group was considered an outlier and it was excluded.

Table III. miRs selected for the validation phase. Three miRs were selected among the significantly differently expressed miR between healthy and unhealthy centenarians, and taking into account the aging trajectory.

	HC avg reads±SD	UHC avg reads±SD	O avg reads±SD	Y avg reads±SD
miR-598-3p	13.33 ± 4.64	1.33 ± 1.25	7.66 ± 3.09	7.5 ± 3.50
miR-30a-5p	1511.33 ± 751.27	3411.33 ± 974.13	1448.33 ± 633.87	804 ± 208
miR-766-3p	58 ± 11.43	13 ± 12.03	19.67 ± 11.73	26.5 ± 20.5

HC: healthy centenarians. UHC: unhealthy centenarians. O: old. Y: young. AVG reads: average of the reads count in each group. SD: standard deviation

4.4 Validation phase

To validate the previously selected miRs a larger cohort of donors was included in the study, for a total of 48 subjects, including 16 healthy young donors (average 30 years old), 16 healthy old donors (average 71 years old), 16 centenarians of which 10 healthy and 6 unhealthy (average 101 years). Healthy and unhealthy centenarians were also studied by hemato-clinical parameters.

RNA extraction, starting from 100 µl of EDTA-plasma, was performed using the Total RNA purification kit (Norgen) and analysis were done using the RT-qPCR with TaqMan technologies, as standardized protocol. The spike-in, cel-miR-39, was evaluated in each sample to detect RNA extraction quality. Results showed a good consistence of RNA isolation, since Ct value average of this exogenous miR was expressed at 18.3 ± 0.6 level. Cel-miR-39 was also used to normalize the Ct values of each miR.

4.4.1 RT-qPCR validation results

TaqMan technologies was applied to analyze miRs in RT-qPCR using specific probe for each miR. MiRs level was calculated as relative expression compared to cel-miR-39 normalizer.

MiR-598-3p validation results are reported in figure 4.15. MiR-598-3p increase in old subjects and decrease in centenarians, significantly in healthy conditions. The difference between young and old donors has an almost significant trend (p value=0.063). Since no difference was relieved between the two groups of centenarians, miR expression was also analyzed joining miR data of the two groups as reported in lower panel of the figure 4.15. The decreased expression in the oldest group was confirmed.
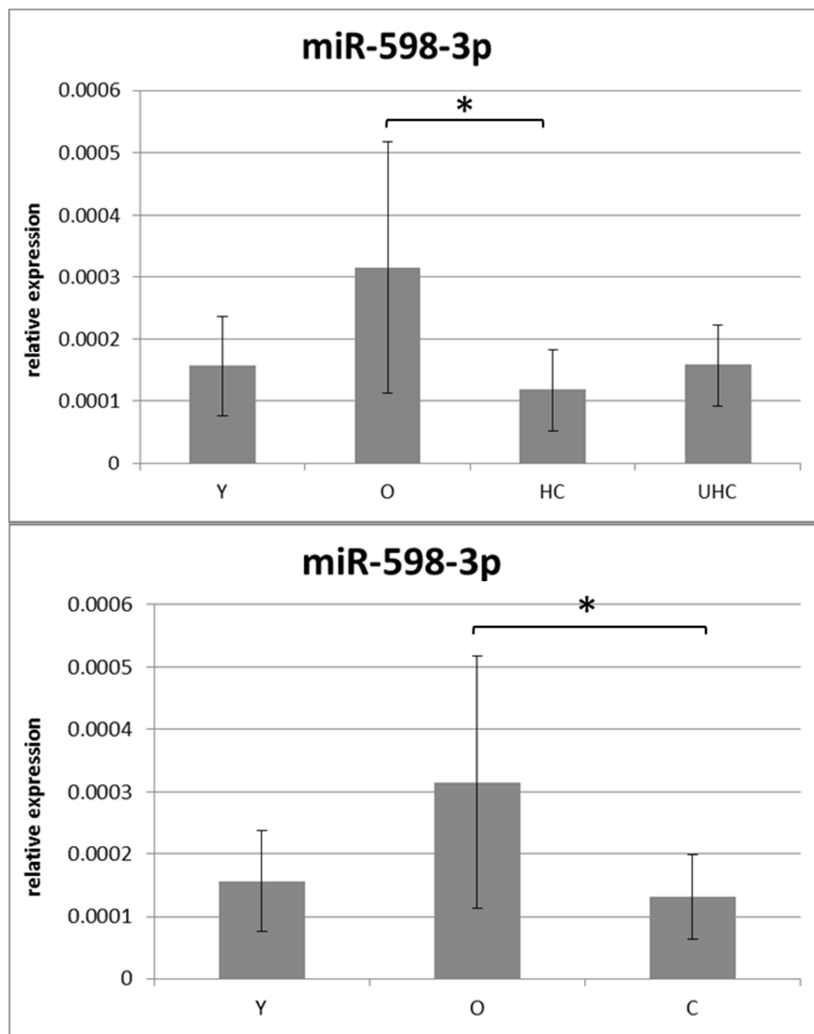

Figure 4.15. Relative expression of miR-598-3p. miR-598-3p was evaluated in young (Y), old (O), healthy (HC) and unhealthy (UHC) centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together (C). Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: * = p ≤ 0.05.

MiR-766-3p validation results are reported in figure 4.16. MiR-766-3p shows the same trend evidenced in miR-598-3p but no significance was found. Similarly the two group of centenarians were joined. Nevertheless this trend was not significant.
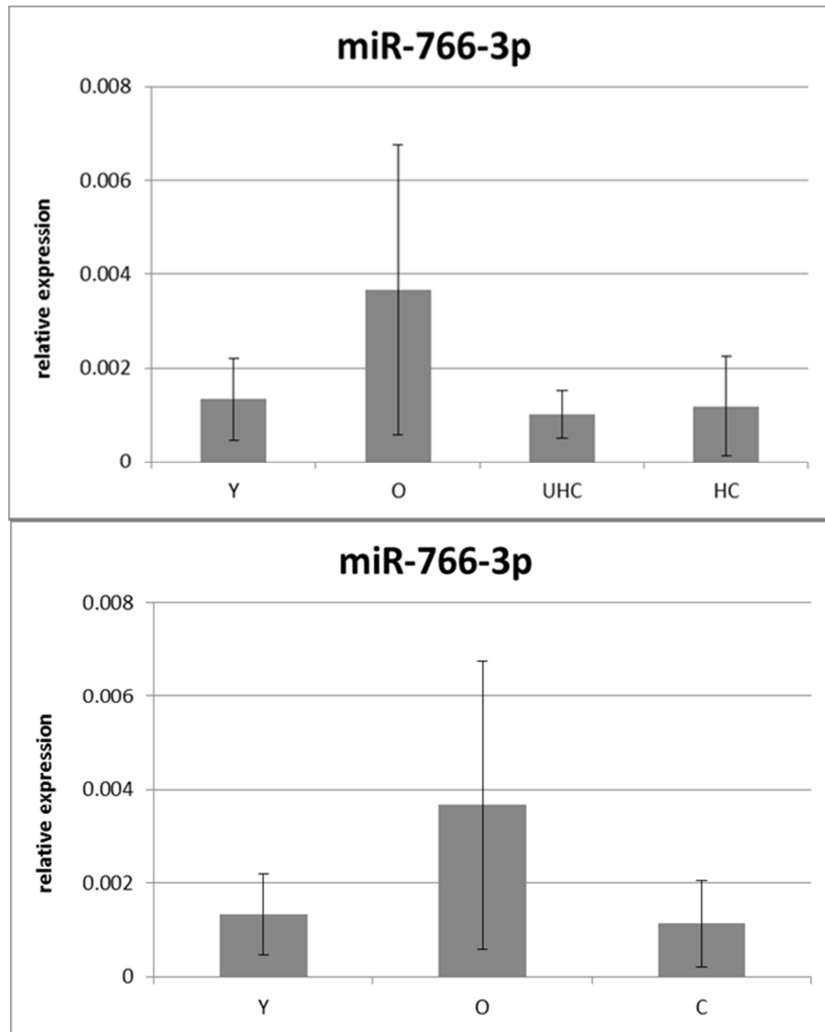

Figure 4.16. Relative expression of miR-766-3p. miR-766-3p was evaluated in young, old, healthy and unhealthy centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together. Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: $* = p \leq 0.05$.

MiR-30a-5p validation results are shown in figure 4.17. Similarly to miR-766-3p, miR-30a-5p did not show any significant trend probably due to the variability inside the different groups.
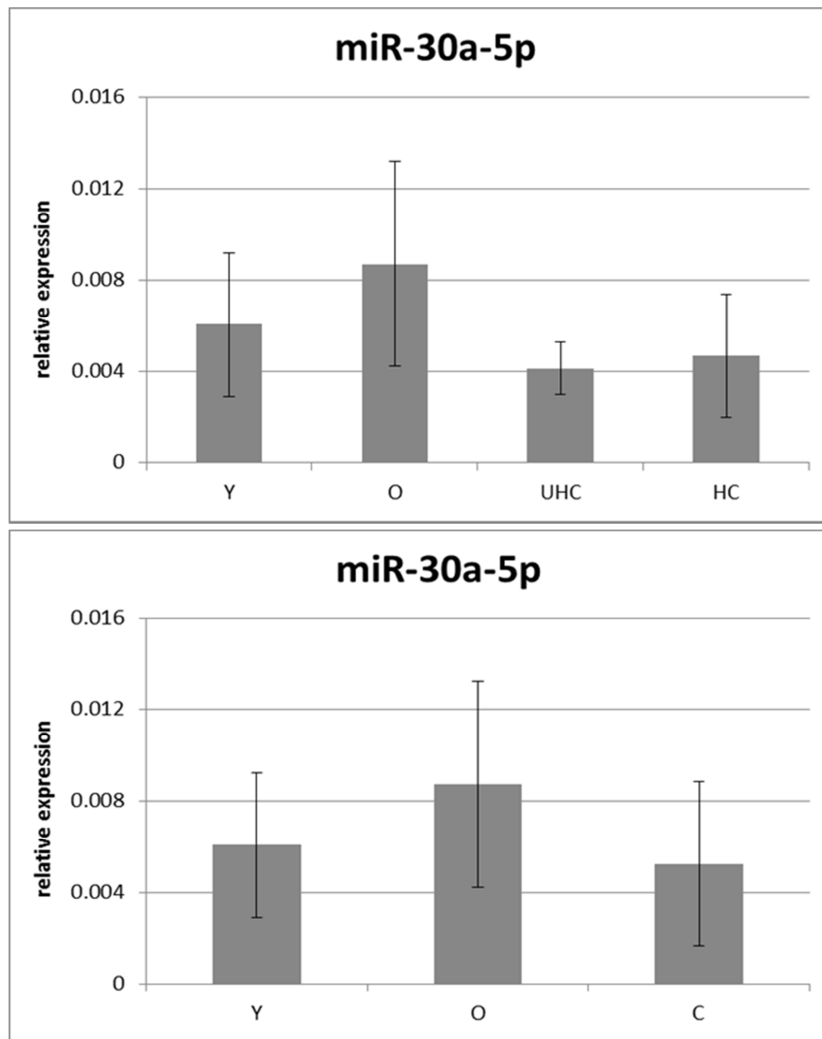
70

Figure 4.17. Relative expression of miR-30a-5p. miR-30a-5p was evaluated in young, old, healthy and unhealthy centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together. Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: * = p ≤ 0.05.

4.4.2 MyomiRs analysis

Taking into account the different health status between the two group of centenarians (unhealthy centenarians were demented and bedridden), myomiRs were also investigated in this study. MiR-206 and miR-133a-3p are the two most important miRs involved in muscle function and diseases and for that reasons they were selected additionally to be assessed in this enlarge cohort (48 donors).

Mir-133a-3p was measured in the different aged groups as reported in figure 4.18. This miR increases in elderly people and decreases in healthy and unhealthy centenarians. Specifically the down-regulation of miR-133a-3p in healthy centenarians respect to old people was statistically significant, and considering all centenarians together the decrease was confirmed to be significant.
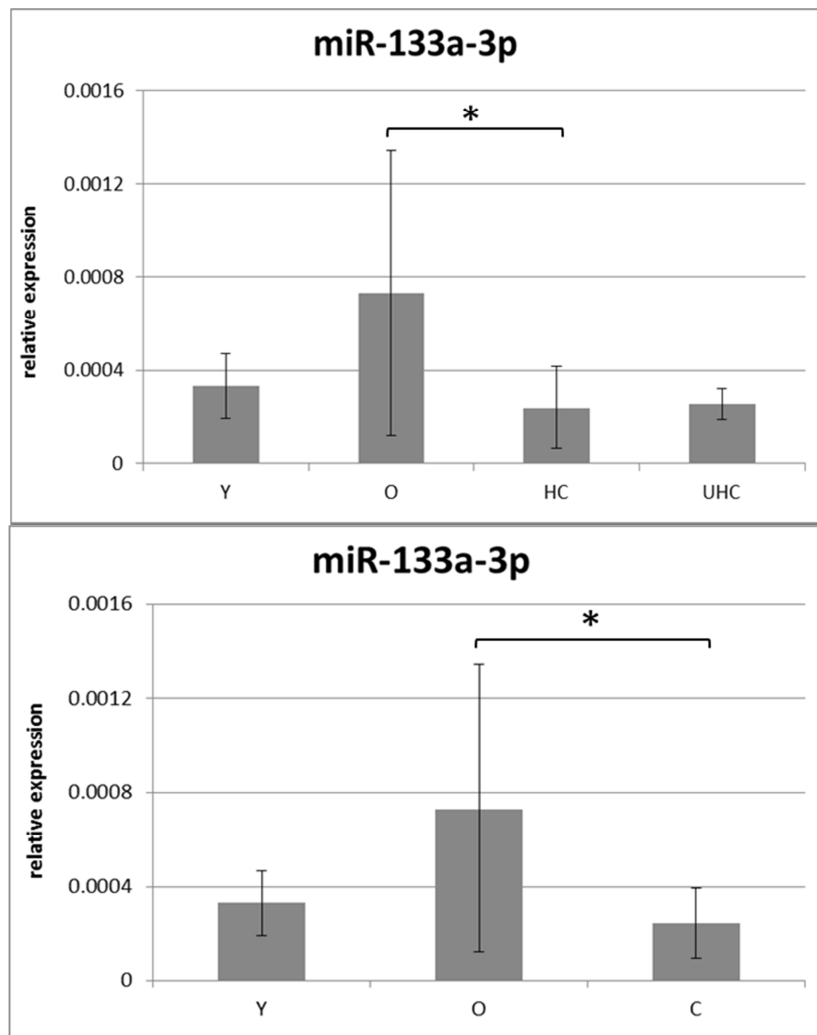

Figure 4.18. Relative expression of miR-133a-3p. miR-133a-3p was evaluated in young, old, healthy and unhealthy centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together. Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: * = p ≤ 0.05.

MiR-260, another important myomiR, was analyzed. In figure 4.19 relative expressions of miR-206 in each group are indicated. This myomiR increases along the different aged group and even more in

unhealthy centenarians, showing a very clear relationship between age and miR expression, further confirmed joining data obtained by centenarians (lower panel).
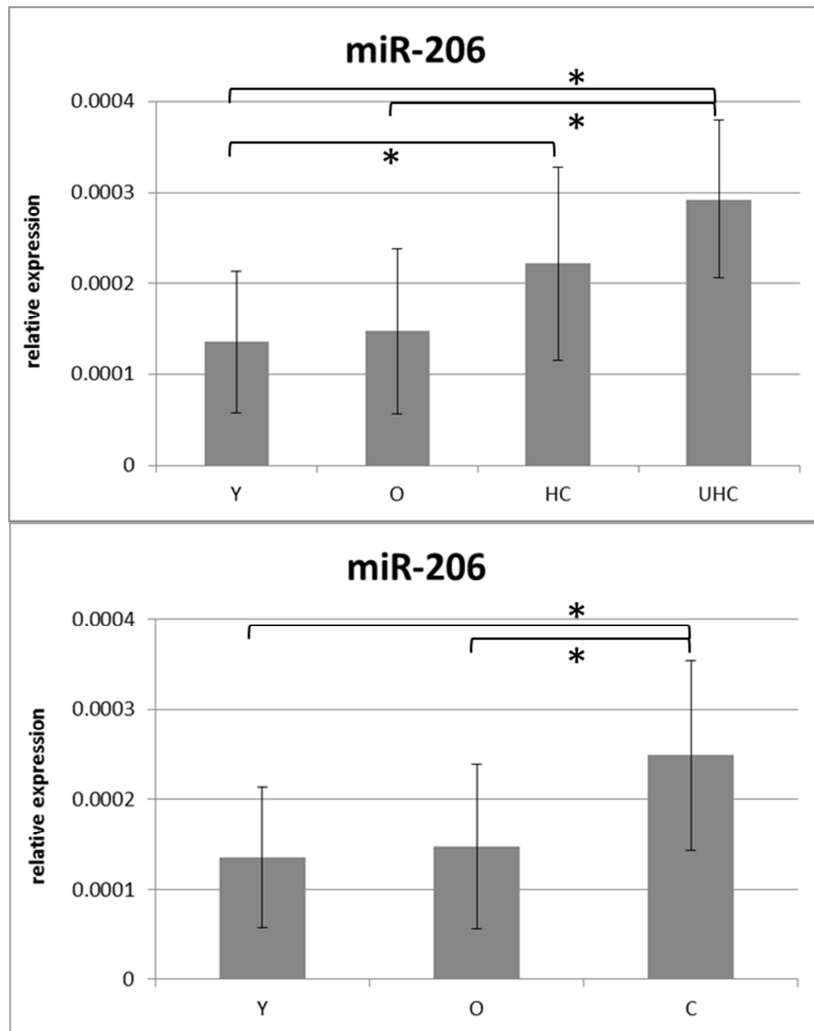

Figure 4.19. Relative expression of miR-206. In the upper panel miR-206 was evaluated in young, old, healthy and unhealthy centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together. Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: * = p ≤ 0.05.

4.4.3 MiR-16 analysis

MiR-16 is an important circulating miR which is often measured to evaluate the level of hemolysis (this miR is contained in red cells) or for sample normalizator. It was evaluated in all 48 donors and results, as reported in figure 4.20, suggest that it decreases with age. These data are consistent with a

result influenced by the aging process thus miR-16 is not suitable to be used as normalizator. Further, the level of variance is kept constant intra-groups, thus excluding a possible effect of random hemolysis in samples.
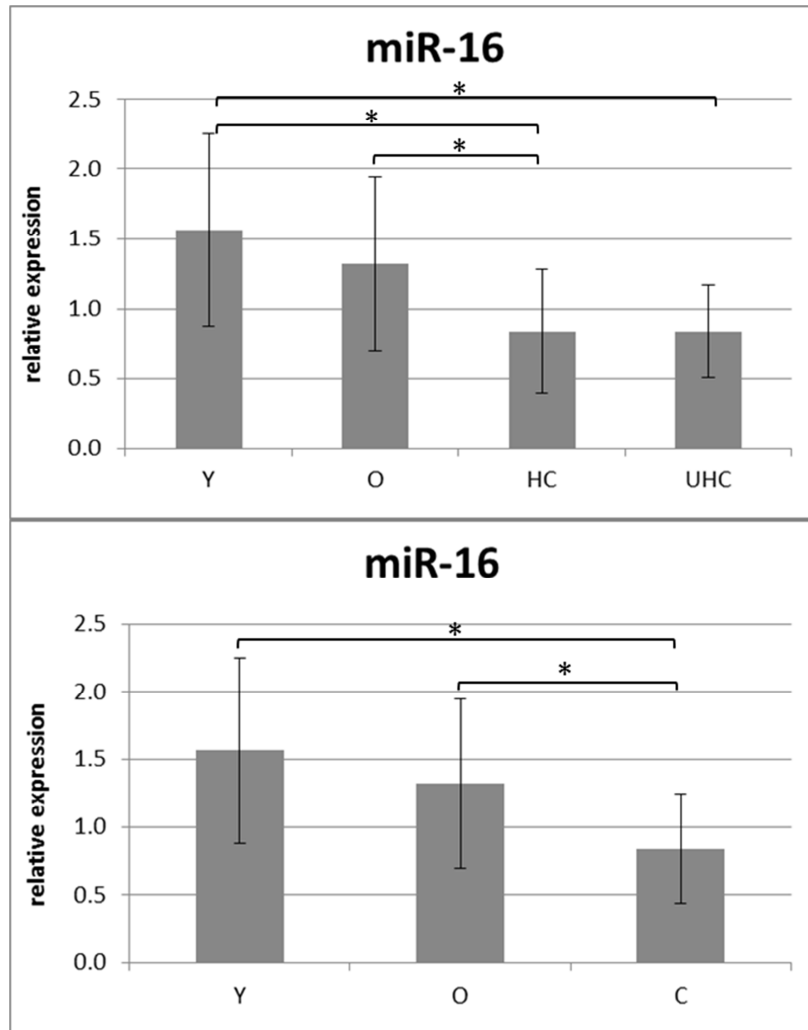

Figure 4.20. Relative expression of miR-16. miR-16 was evaluated in young, old, healthy and unhealthy centenarians (upper panel). In the lower panel, healthy and unhealthy centenarians were considered together. Data are reported as mean values ± standard deviation. Data were analyzed with Kruskal-Wallis test: $* = p \leq 0.05$.

4.4.4 Aging and longevity trajectories

Taking into account those miRs resulted significant among the different aged groups, it is possible to individuate trajectories along aging and longevity. Identified miRs can distinguish not only aging process, when their level is simply increased or decreased (miR-206; -16) with age, but also longevity trajectory when other miRs, i.e. miR-598-3p and -133a-3p, have level of expression changed in old people but restored at young level in centenarians group. The different trajectories identified are reported in figure 4.21.
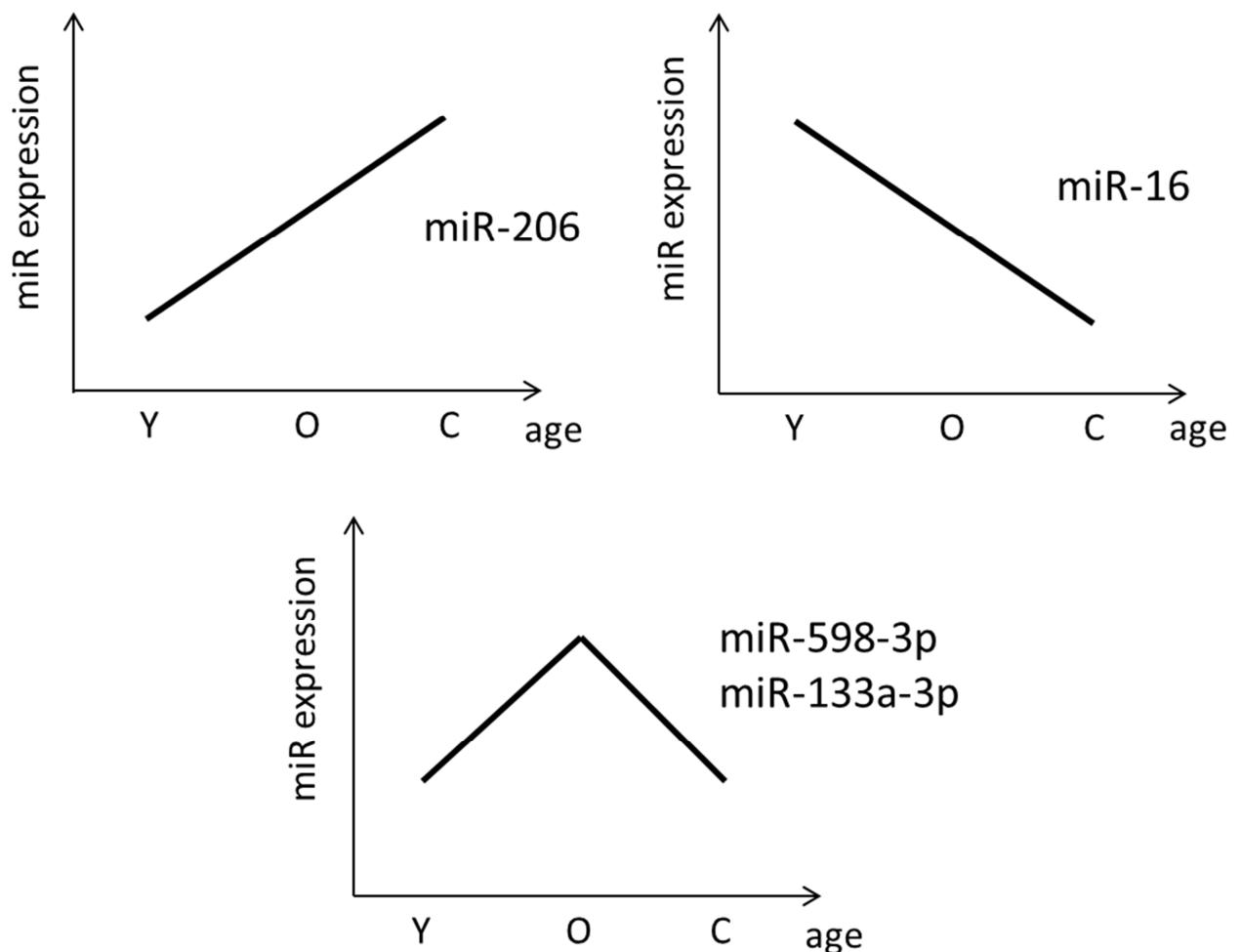

Figure 4.21. Trajectories of aging and longevity. The first two top panels represent the trajectories of aging, found with the expression of blood circulating miR-206 and miR-16. In the lower panel the longevity trajectory is described, as expression of miR-598-3p and miR-133a-3p, being centenarians restored to young group.

4.4.5 MiRs expressions and hemato-biochemical parameters

Additionally, the challenge to distinguish between healthy and unhealthy phenotype in centenarians groups was also evaluated by analyzing hemato-biochemical parameters and some of them resulted significant, or almost significant, as reported in table IV.

Table IV. Hemato-biochemical parameters significant different between healthy and unhealthy centenarians.
Factors are reported as mean value of each group ± standard deviation (sd). T test between the two groups is shown for each parameters as p value.

		wbc	mchc	rdwcv	alb	prot	pch
HC	mean	5.65	31.82	14.68	3.82	7.10	3.41
	sd	0.92	0.78	0.99	0.58	0.48	4.41
UHC	mean	7.55	30.43	16.30	3.25	6.46	11.57
	sd	2.41	1.32	2.23	0.23	0.36	10.18
	p value	0.039	0.019	0.063	0.039	0.028	0.041

WBC: white blood cell count (4.8-8.5 x1000/µl). MCHC: mean corpuscular hemoglobin concentration (33-38 gr/dl). RDWCV: red blood cell distribution (11.5-14.5 %). ALB: albumin (3.5-5.2 gr/dl). PROT: total protein (6.2-8 gr/dl). CRP: C-reactive protein (max6 mg/l).

These parameters (WBC, MCHC, RDWCV, ALB, PROT, CRP) jointed with all identified miRs are able to distinguish healthy and unhealthy centenarians (excluding one case) as shown in the figure 4.22. PLS-DA analysis of all centenarians clearly suggest a possible signature of longevity in good (green circles) or bad (blue circles) status, underlined by the cross validated predictive fraction (Q2)=0.382.
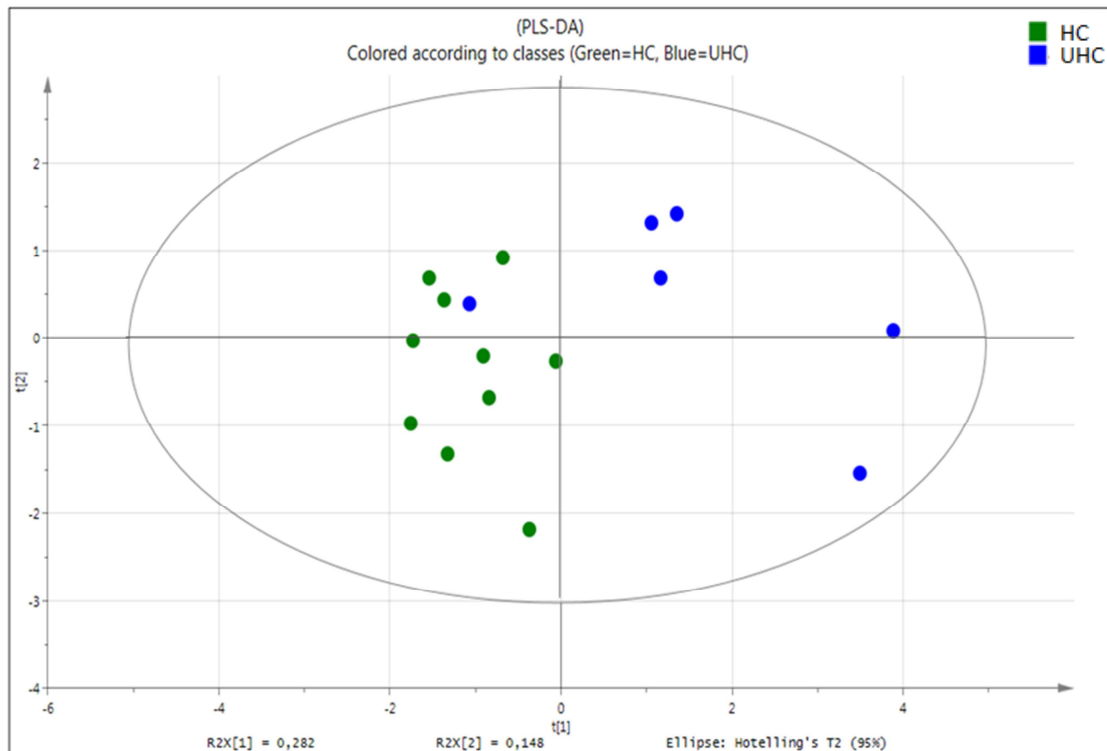
Figure 4.22. PLS_DA score plot of miRs expressions data and hemato-biochemical parameters from healthy centenarians (green circles) and unhealthy centenarians (blue circles). For PLS-DA, cumulative explained variance (R2Y) = 0.688 and cross validated predictive fraction (Q2) = 0.382 were calculated.

4.4.6 Pathway analysis

Bioinformatics analysis of miRs 598-3p; -133a-3p; -206; -16 with mirPath software (DIANA TOOLS, v2.0) allowed the identification of a common pathway among miRs targets on the basis of both putative and experimentally validated targets by previously published data. Specifically, miR-598-3p has not yet validated targets thus only its putative targets were considered. On the other hand, miR-133a-3p, miR-206 and miR-16 were analyzed for both putative and validated targets. On the whole, the bioinformatics analysis raised up one important pathway, i.e. PI3K-Akt as reported in figure 4.23, with high significant p value (p value = 1.13e-10). This is an intracellular signaling pathway crucial for cell cycle modulation, apoptosis, NFkB/p53 signalling and mTOR (etc). This pathway was set up by miR-16 and miR-133a-3p validated targets and by miR-206 and miR-598-3p putative targets. In particular miR-16 and miR-206 seem to have some common targets, i.e. EIF4E (eukaryotic translation initiation factor 4E), CDK6 (Cyclin-dependent kinase 6), GNB1 (Guanine Nucleotide Binding Protein 1) genes (colored in orange) in the scheme below (figure 4.23).
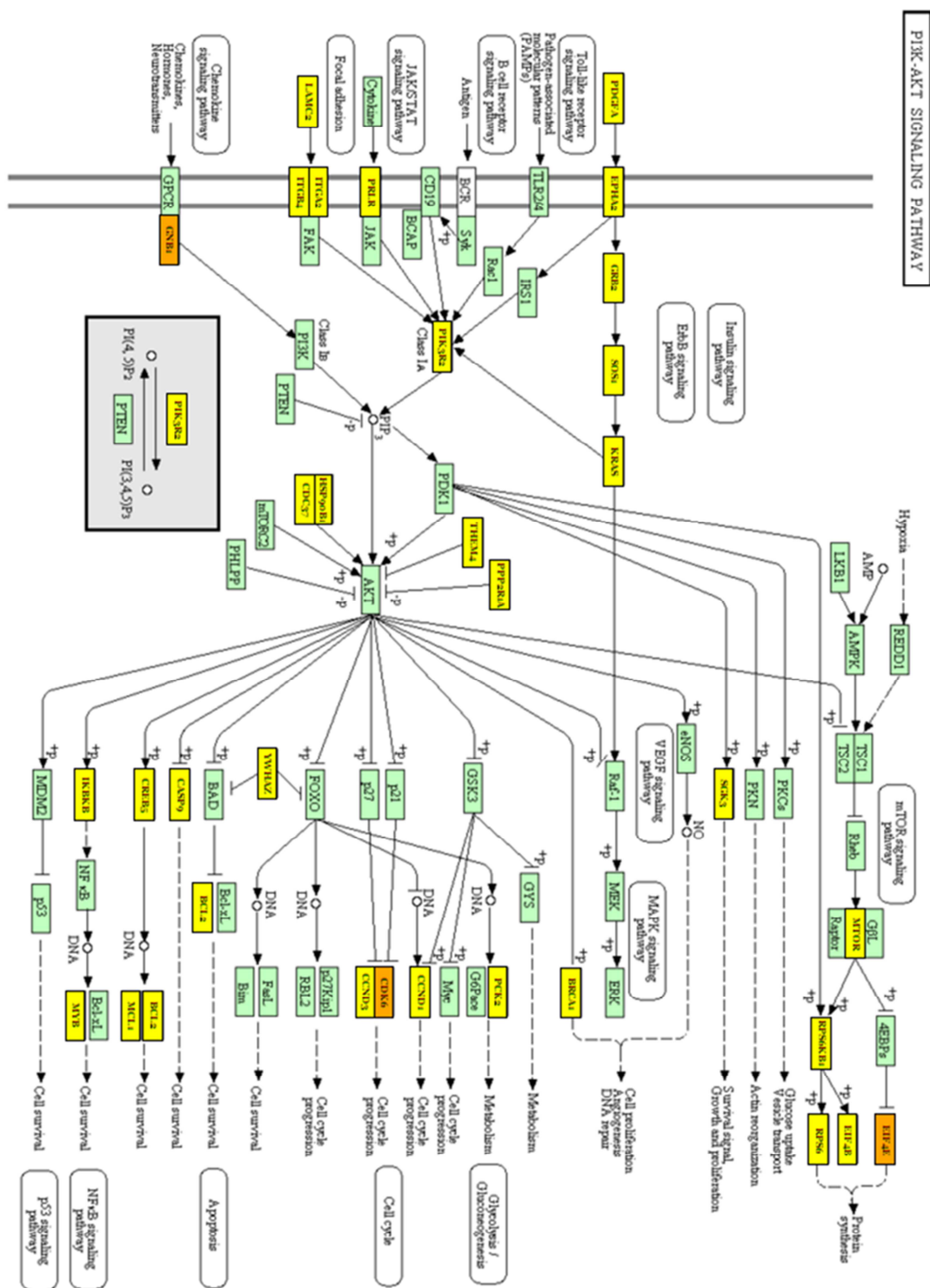
Figure 4.23. PI3K-Akt signaling, common pathway for miR-598-3p, miR-133a-3p, miR-16, miR-206. In yellow, genes which are targeted by only one miR included in the analysis; in orange, targets validated or putative in common for more than one miR.

78

5. DISCUSSION AND CONCLUSIONS

Human aging has become one of the major social, political and economic issues of the last few decades. Healthy and active aging is a crucial aspect of EU aging population accordingly with recent WHO declaration that "increased longevity without quality of life is an empty prize", thus human aging research field has to face the problem to slow down the onset of age-related pathologies and promote healthy longevity. Further, in Italy 19,096 centenarians (ISTAT at January 1st, 2015) are living and their number is projected to increase enormously, populations projections foresee the number of centenarians will further increase to 97,995 (100+ in Italy, total population 61M; World Population Prospects: the 2010 revision-USA 2011) on 2050.

Human aging is a process mainly orchestrated by gene expression deregulation (Blagoskonny et al., 2013), the accumulation of unrepaired damages, the activation of chronic low grade inflammatory process sustained by a lifelong exposure to internal and external antigenic load (Franceschi et al., 2000; Wild, 2012) which can affect health status and favor the onset of age-related disease. Population aging is not uniform, many factors can interact and determine a huge variability as counteracting process of defense and remodeling (Cevenini et al., 2008). This variability is due to the genetic complex trait, epigenetics, environment, life style and stochasticity that differently affect the rate of aging at the levels of cells, tissues or body systems within the same organism and among members from the same species. This complexity makes more difficult the identification of a unique comprehensive mechanism of aging and age-associated biomarkers able to distinguish not only the rate of the process, but different trajectories likely conducting towards healthy aging and longevity or eventually the onset of age-related pathologies.

During last decade many teams have been tried to identify new crucial biomarker of aging capable to identify biological versus chronological age (Capri et al., 2015). Among them, those related to the epigenetic clock (353CpG sites shaping an aging clock in terms of chromatin states and tissue variance), recently raised up from studies on differently aged human tissues, human obesity and Down Syndrome (Horvath, 2013; Horvath et al., 2014 and Horvath et al., 2015), reveal powerful tissue-specific biomarkers mirroring the systemic condition. These findings open new

questions on the future role of the biomarkers network which supports information of healthy status and likely could also "transport communication" among different cells, organs or tissues.

In this context epigenetic markers appear to be the most promising for the best identification of aging trajectories and modeling applications. For instance, the DNA methylation levels of peripheral blood mononuclear cells (PBMCs) could be a strong biomarker of biological age. In fact, semi-supercentenarians' offspring have a lower epigenetic age than age-matched controls (age difference = 5.1 years) and centenarians are younger (8.6 years) than expected based on their chronological age (Horvath al., 2015).

A broaden literature suggests a strong role of circulating blood (c-)miRs as potential biomarkers for age-related pathologies (Jung et al., 2014). In this context the work of the current thesis is presented together with the basic idea of distinguishing healthy and unhealthy longevity in terms of molecular pathway deregulation and the most significant gene products. The challenge is to find treatments, likely in a precise medicine vision, able to slow down, at early stage, the onset of age-related pathologies and promote healthier aging and longevity.

MiRs are important regulator of gene expression at post-transcriptional level and they are clearly involved in many physiologically process (Kloosterman et al., 2006). Indeed many studies have shown that miRs levels change during the lifespan of several species and can be associated with age-related diseases (Harries, 2014). From the first evidence that miRs are released in the body fluid, such as blood, urine, breast milk, saliva and so on, c-miRs were considered not only epigenetic gene regulators but also non-invasive markers, in different physiologically and pathologically conditions, including aging process.

The main purpose of the current thesis was to investigate the role of blood c-miRs and their expression profile during human aging and longevity by the means of the most advanced technology, i.e. smallRNA sequencing with Next Generation Sequencing (NGS) platform. After a period spent for optimization of the sequencing protocol, in collaboration with the BOKU University in Vienna, miRs profiling was conducted in two different phases, i.e. discovery and validation. The former was performed on a few samples but at highest resolution using smallRNA-seq. miRs profile was performed on samples from 12 donors (9 healthy subjects: 3 young, 3 old, 3

centenarians and 3 unhealthy centenarians). Donors were recruited in Bologna and centenarians were deeply characterized to distinguish healthy from unhealthy status. Plasma samples were used for the analysis of c-miRs. The sequencing was conducted on Illumina platform from Exiqon company and, as first analysis, differentially expressed miRs between healthy and unhealthy (demented and bedridden) centenarians were considered. Only 13 miRs showed a significantly different expression between the two group and among these miR-598-3p, miR-766-3p and miR-30a-5p revealed an interesting trajectory along aging process (young and old groups), considering the average read counts in every groups.

The latter phase was spent to validate data on the selected miRs using RT-qPCR on a larger cohorts of subjects (48 total subjects), i.e. 16 healthy young donors, 16 healthy old donors and 16 centenarians, divided in 10 healthy and 6 unhealthy centenarians. Expression of miR-598-3p, miR-766-3p and miR-30a-5p showed the same type of trend, i.e. an increase expression in elderly people and a decrease in centenarians. However these changes resulted statistically significant only in miR-598-3p, specifically the down-regulation of miR-598-3p between old and centenarians groups was strongly significant, likely due to healthy centenarians as confirmed by statistical analysis. MiR-30a-5p and miR-766-3p did not show any significant differences among different age groups.

Taking into account most relevant results coming from literature and from different models of aging investigated by the team at the University of Bologna, further miRs were considered, e.i. miR-133a-3p; -206 and miR-16. MiR-133a and -206 are two important myomiRs which are recognized to be involved in muscle development and phenotype (Kirby and McCarthy, 2013) while miR-16 is an important miR often used as internal reference for normalization or hemolytic marker (Kirschner et al., 2013).

miR-133a-3p showed the same trend observed in miRs-598-3p; -766-3p; -30a-5p , in particular the decrease between old and centenarians' groups was found significant, due to healthy centenarians as showed by statistical analysis. miR-598-3p, for the first time identified in the aging process, and miR-133a-3p reflect the longevity trajectory since miR expression increases in old donors but decrease in centenarians, thus restoring the level of miRs found in young donors.

miR-206 levels showed a linear age-dependent increase, which was strongly significant in unhealthy centenarians suggesting the possible role of this miR to identify unhealthy trajectories. To this regards, it is known that miR-206 is an important marker for cardiopathy (Oliveira-Carvalho et al.,

2013) likely suggesting that alteration of its targets can lead to unfavorable aging process mostly affecting cardiovascular system (Bonn et al., 2013).

Interestingly, miR-16 levels showed a progressive and significant decrease with age, thus this specific trend strongly suggest that miR-16 cannot be used as internal normalizator. MiR-16 has already been described differentially expressed in elderly people, it was found upregulated in aged unstimulated B cells and positively correlated with serum and B cell–intrinsic TNF-α level (Frasca et al., 2015). MiR-16 is also clearly involved in apoptosis and senescence process (Kitadate et al., 2015). In this current work miR-16 showed an opposite trend respect to miR-206, describing a human aging trajectory, with a strong decrease in healthy centenarians.

A limitation of the work here presented is that smallRNA sequencing did not produce data which were validated in RT-qPCR. Various assumptions could be proposed about this apparently discordant results, in principle technical problems and/or not sufficient power in the discovery phase cannot be excluded. Basically, high-throughput NGS has different limitations, beyond relatively high costs, the substantial computational infrastructure needed for data analysis and interpretation, as well as sequence-specific biases related to enzymatic steps in cDNA library preparation methods that favor capture of some miRs over others. Further, miR quantification is expressed as a value relative to the total number of sequence reads for a given sample; thus comparisons between samples with high variance in miR distribution may not be appropriate, and absolute quantification is not possible.

As far as RT-qPCR validation method is concerned, this is considered the most sensitive and reliable technique. However some technical considerations should highlighted (Pritchard et al., 2012), such as the difficulties to quantify the small amount of RNA, to select the ad hoc normalizator and different types of spike-in (Mitchell et al., 2008). However an effective normalization strategy for biological variability is currently not well-developed. At present, any specific miR or set of miRs cannot be approved as suitable endogenous controls in blood analysis, even if specific c-miR are used as normalizer only in specific cohort of study when their stable expression is detected (Sanders et al., 2012; Kok et al., 2015; Schlosser et al., 2015).
The increasing interest in c-miRs as biomarkers has raised the issue regarding the influence of hemolysis on the measurements of miRs levels in serum/plasma and the reliability of these data

(Pritcher et al., 2011). Different evaluation methods of hemolysis level were proposed in literature (Kirschner et al., 2011; Blondal et al., 2013; Fortunato et al., 2014) but the lack of consensus regarding a general method leads to a difficult evaluation of sample quality, without a reproducible and clear evaluation. However, the result reported in this work suggest that miR-16 level, the most cited for hemolytic level control, had stable variance among the different age groups (no hemolysis was directly observed on the samples), thus permitting to exclude random hemolysis effect (Yamada et al., 2014).

The current work answered two main questions:

1. May c-miRs profile distinguish between healthy and unhealthy longevity?

Looking at miRs validated profiles exclusively, no clear evidence of a signature between healthy and unhealthy centenarians was found. On the other hand, miRs levels jointed with hemato-biochemical parameters (i.e. white blood cell count, mean corpuscular hemoglobin concentration, red blood cell distribution, albumin, total protein, C-reactive protein), produced a good "signature" to distinguish the two groups of centenarians. In fact, a specific analysis, i.e. PLS-DA, can discriminate healthy and unhealthy phenotype. One important parameter among those used is the C-reactive protein (CRP), normally found in the blood, whose levels rise in response to inflammation. CRP was already described as correlated to inflamma-miR, i.e. in circulating miR-21, (Olivieri et al., 2012) and importantly is confirmed to be a good marker for phenotypical risk profiles, as previously suggested (Abbatecola et al., 2004; Morrisette-Thomas al., 2014).
The parameters used to discriminate the two different health status also suggest the role of immune system cells and in particular the number of leucocytes which increase in unhealthy centenarians, thus further confirming the role of immune system at oldest ages (Bucci et al., 2014).

2. May c-miRs profile specifically characterize aging and longevity trajectories?

Levels of miRs 598-3p -133a-3p; -206; -16 clearly evidenced different aging trajectories and in particular miR-598-3p and -133a-3p appear to follow a longevity trajectory. Bioinformatics analysis of their putative and validated targets (see appendix A) identified a common pathway, i.e. PI3K-Akt signaling. This pathway regulates multiple biological processes including cell survival,

proliferation, growth, glycogen metabolism, apoptosis etc. It was already described the role of this signaling, in synergy with others crucial pathways, i.e. insulin/IGF-1 pathway and mTOR pathway, in human aging process (Capri et al., 2014; Franceschi et al., 2007). Furthermore long-term calorie restriction (CR), that is one of the most robust intervention for extending lifespan in animal models and preventing/delaying age-related disease (Longo et al., 2015), showed a significantly down-regulation effect on PI3K and AKT transcripts (Mercken et al., 2013). Modulation of these miRs and their specific targets, could be proposed to mimic CR and likely delay the aging process, as recently proposed for specific drugs which can modify these signalings according with CR effects (Madeo et a., 2014).

To this regard, nutrition intervention is one of the most proposed as anti-aging approach, as well as Mediterranean diet, to favor anti-inflammaging effects in elderly (Ostan et al., 2015). Further, pharmacological interventions, physical exercise together with modification of the lifestyle were suggested to prolong life and/or sustain health late in life (de Cabo et al., 2014).

In conclusion, the topic of the current thesis is inserted in the complex scenario of human aging markers in order to identify new anti-aging intervention. Data obtained strongly suggest that the use of many parameters, not only miRs, permit to distinguish different phenotypes, such as healthy and unhealthy centenarians, and further data could be used in a more complex analysis to create algorithms for modeling in the context of a personalized and precise medicine.

Future analysis will involve a larger panel of miRs in differently aged groups and the combination of them with others factors will try to disentangle the complexity of human aging process. To this regard, system medicine combine data from different levels of analysis, such as transcriptomics, proteomics, metabolomics using ad hoc experimental design (cross sectional and longitudinal cohorts), thus providing new multidimensional biomarkers capable of distinguishing among different pathological conditions and aging (Castellani et al., 2015).

6. BIBLIOGRAPHY

Abbatecola AM, Ferrucci L, Grella R, Bandinelli S, Bonafè M, Barbieri M, Corsi AM, Lauretani F, Franceschi C, Paolisso G. Diverse effect of inflammatory markers on insulin resistance and insulin-resistance syndrome in the elderly. J Am Geriatr Soc 2004; 52(3):399-404.

Adachi T, Nakanishi M, Otsuka Y, Nishimura K, Hirokawa G, Goto Y, Nonogi H, Iwai N. Plasma microRNA 499 as a biomarker of acute myocardial infarction. Clin Chem 2010; 56:1183–1185.

Ajit SK. Circulating microRNAs as Biomarkers, Therapeutic Targets, and Signaling Molecules. Sensors 2012; 12(3):3359-3369.

Altilia S, Santoro A, Malagoli D, Lanzarini C, Ballesteros Alvarez JA, Galazzo G, Porter DC, Crocco P, Rose G, Passarino G, Roninson IB, Franceschi C, Salvioli S. TP53 codon 72 polymorphism affects accumulation of mtDNA damage in human cells. Aging 2012; 4:28–39.

Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc Natl Acad Sci 2011; 108:5003–08.

Baraibar MA, Friguet B. Oxidative proteome modifications target specific cellular pathways during oxidative stress, cellular senescence and aging. Exp Gerontol 2013; 48(7):620-5.

Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004; 116(2):281–297.

Blagosklonny MV. Aging is not programmed: genetic pseudo-program is a shadow of developmental growth. Cell Cycle 2013; 12(24):3736-42.

Blondal T, Jensby Nielsen S, Baker A, Andreasen D, Mouritzen P, Wrang Teilum M, Dahlsveen IK. Assessing sample and miRNA profile quality in serum and plasma or other biofluids. Methods 2013; 59(1):S1-6.

Boon RA, Iekushi K, Lechner S, Seeger T, Fischer A, Heydt S, Kaluza D, Tréguer K, Carmona G, Bonauer A, Horrevoets AJ, Didier N, Girmatsion Z, Biliczki P, Ehrlich JR, Katus HA, Müller OJ, Potente M, Zeiher AM, Hermeking H, Dimmeler S. MicroRNA-34a regulates cardiac ageing and function. Nature 2013; 495(7439):107-10.

Borer RA, Lehner CF, Eppenberger HM, Nigg EA. Major nucleolar proteins shuttle between nucleus and cytoplasm. Cell 1989; 56:379–390.

Bucci L, Ostan R, Giampieri E, Cevenini E, Pini E, Scurti M, Vescovini R, Sansoni P, Caruso C, Mari D, Ronchetti F, Borghi MO, Ogliari G, Grossi C, Capri M, Salvioli S, Castellani G, Franceschi C, Monti D. Immune parameters identify Italian centenarians with a longer five-year survival independent of their health and functional status. Exp Gerontol 2014; 54:14-20.

Capri M, Santoro A, Garagnani P, Bacalini MG, Pirazzini C, Olivieri F, Procopio A, Salvioli S, Franceschi C. Genes of human longevity: an endless quest? Curr Vasc Pharmacol. 2014; 12(5):707-17.

Capri M, Moreno-Villanueva M, Cevenini E, Pini E, Scurti M, Borelli V, Palmas MG, Zoli M, Schön C, Siepelmeyer A, Bernhardt J, Fiegl S, Zondag G, de Craen AJ, Hervonen A, Hurme M, Sikora E, Gonos ES, Voutetakis K, Toussaint O, Debacq-Chainiaux F, Grubeck-Loebenstein B, Bürkle A, Franceschi C. MARK-AGE population: From the human model to new insights. Mech Ageing Dev 2015; 151:13-7.

Cevenini E, Invidia L, Lescai F, Salvioli S, Tieri P, Castellani G, Franceschi C. Human models of aging and longevity. Expert Opin Biol Ther 2008; 8(9):1393-405.

Cevenini E, Monti D, Franceschi C. Inflamm-ageing. Curr Opin Clin Nutr Metab Care 2013; 16:14-20.

Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, Wang J, Jiang X, Xiang Y, Xu C, Zheng P, Zhang J, Li R, Zhang H, Shang X, Gong T, Ning G, Zen K, Zhang CY. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 2008; 18:997–1006.

Cheng Y, Tan N, Yang J, Liu X, Cao X, He P, Dong X, Qin S, Zhang C. A translational study of circulating cell-free microRNA-1 in acute myocardial infarction. Clin Sci 2010; 119(2):87-95.

Cordes KR, Sheehy NT, White MP, Berry EC, Morton SU, Muth AN, Lee TH, Miano JM, Ivey KN, Srivastava D. MiR-145 and miR-143 regulate smooth muscle cell fate and plasticity. Nature 2009; 460:705–710.

De Cabo R, Carmona-Gutierrez D, Bernier M, Hall MN, Madeo F. The search for antiaging interventions: from elixirs to fasting regimens. Cell 2014; 157(7):1515-26.

Deelen J, Beekman M, Capri M, Franceschi C, Slagboom PE. Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges. Bioessays 2013; 35(4):386-96.

ElSharawy A, Keller A, Flachsbart F, Wendschlag A, Jacobs G, Kefer N, Brefort T, Leidinger P, Backes C, Meese E, Schreiber S, Rosenstiel P, Franke A, Nebel A. Genome-wide miRNA signatures of human longevity. Aging Cell 2012; 11(4):607-16.

Etheridge A, Lee I, Hood L, Galas D, Wang K. Extracellular microRNA: a new source of biomarkers. Mutation research 2011; 717(1-2):85-90.

Fortunato O, Boeri M, Verri C, Conte D, Mensah M, Suatoni P, Pastorino U, Sozzi G. Assessment of circulating microRNAs in plasma of lung cancer patients. Molecules 2014; 19(3):3038-54.

Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, De Benedictis G. Inflamm-aging. An evolutionary perspective on immunosenescence. Ann N Y Acad Sci 2000; 908:244-54.

Franceschi C, Bonafè M. Centenarians as a model for healthy aging. Biochem Soc Trans 2003; 31:457–461.

Franceschi C, Capri M, Monti D, Giunta S, Olivieri F, Sevini F, Panourgia MP, Invidia L, Celani L, Scurti M, Cevenini E, Castellani GC, Salvioli S. Inflammaging and antiinflammaging: a systemic perspective on aging and longevity emerged from studies in humans. Mech Ageing Dev 2007; 128:92– 105.

Frasca D, Diaz A, Romero M, Ferracci F, Blomberg BB. MicroRNAs miR-155 and miR-16 Decrease AID and E47 in B Cells from Elderly Individuals. J Immunol 2015; 195(5):2134-40.

Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AM, Gentilini D, Vitale G, Collino S, Rezzi S, Castellani G, Capri M, Salvioli S, Franceschi C. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell 2012; 11(6):1132-1134.

Garm C, Moreno-Villanueva M, Bürkle A, Larsen LA, Bohr VA, Christensen K, Stevnsner T. Genetic and environmental influence on DNA strand break repair: a twin study. Environ Mol Mutagen 2013; 54:414-20.

Geekiyanage H, Jicha GA, Nelson PT, Chan C. Blood serum miRNA: non-invasive biomarkers for Alzheimer's disease. Exp Neurol 2012; 235(2):491-6.

Gombar S, Jung HJ, Dong F, Calder B, Atzmon G, Barzilai N, Tian XL, Pothof J, Hoeijmakers JH, Campisi J, Vijg J, Suh Y. Comprehensive microRNA profiling in B-cells of human centenarians by massively parallel sequencing. BMC Genomics 2012; 13:353.

Grasedieck S, Sorrentino A, Langer C, Buske C, Döhner H, Mertens D, Kuchenbauer F. Circulating microRNAs in hematological diseases: principles, challenges, and perspectives. Blood 2013; 121(25):4977-4984.

Güller I, Russell AP. MicroRNAs in skeletal muscle: their role and regulation in development, disease and function. J Physiol 2010; 588:4075-87.

Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 2005; 309(5740):1577-81.

Jung HJ, Suh Y. Circulating miRNAs in ageing and ageing-related diseases. J Genet Genomics 2014; 41(9):465-72.

Hamrick MW, Herberg S, Arounleut P, He HZ, Shiver A, Qi RQ, Zhou L, Isales CM, Mi QS. The adipokine leptin increases skeletal muscle mass and significantly alters skeletal muscle miRNA expression profile in aged mice. Biochem Biophys Res Commun 2010; 400(3):379–83.

Hanke M, Hoefig K, Merz H, Feller AC, Kausch I, Jocham D, Warnecke JM, Sczakiel G. A robust methodology to study urine microRNA as tumor marker: microRNA-126 and microRNA-182 are related to urinary bladder cancer. Urol Oncol 2010; 28:655–661.

Harries LW. MicroRNAs as Mediators of the Ageing Process. Genes 2014; 5(3):656-70.

Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013; 14:R11.

Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, Heits N, Bell JT, Tsai PC, Spector TD, Deloukas P, Siebert R, Sipos B, Becker T, Röcken C, Schafmayer C, Hampe J. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci 2014; 111(43):15538-43.

Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio AM, Giuliani C, Tung S, Vinters HV, Franceschi C. Accelerated epigenetic aging in Down syndrome. Aging Cell. 2015; 14:491-5.

Karolina DS, Tavintharan S, Armugam A, Sepramaniam S, Pek SL, Wong MT, Lim SC, Sum CF, Jeyaseelan K. Circulating miRNA profiles in patients with metabolic syndrome. J Clin Endocrinol Metab 2012; 97(12):E2271-6.

Kim TN, Choi KM. Sarcopenia: definition, epidemiology, and pathophysiology. J Bone Metab 2013; 20:1-10.

Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nature Rev Mol Cell Biol 2009; 10:126–139.

Kirby TJ, McCarthy JJ. MicroRNAs in skeletal muscle biology and exercise adaptation. Free Radic Biol Med 2013; 64:95-105.

Kirschner MB, Kao SC, Edelman JJ, Armstrong NJ, Vallely MP, van Zandwijk N, Reid G. Haemolysis during Sample Preparation Alters microRNA Content of Plasma. PLoS ONE 2011; 6(9):e24145.

Kirschner MB, Edelman JJ, Kao SC, Vallely MP, van Zandwijk N, Reid G. The Impact of Hemolysis on Cell-Free microRNA Biomarkers. Front Genet 2013; 4:94.

Kitadate A, Ikeda S, Teshima K, Ito M, Toyota I, Hasunuma N, Takahashi N, Miyagaki T, Sugaya M, Tagawa H. MicroRNA-16 mediates the regulation of a senescence-apoptosis switch in cutaneous T-cell and other non-Hodgkin lymphomas. Oncogene 2015; doi: 10.1038/onc.2015.435.

Kloosterman WP, Plasterk RHA. The diverse functions of miRNAs in animal development and disease. Dev Cell 2006; 11:441-450.

Kok MG, Halliani A, Moerland PD, Meijers JC, Creemers EE, Pinto-Sietsma SJ. Normalization panels for the reliable quantification of circulating microRNAs by RT-qPCR. FASEB J 2015; 29(9):3853-62.

Kong L, Zhu J, Han W, Jiang X, Xu M, Zhao Y, Dong Q, Pang Z, Guan Q, Gao L, Zhao J, Zhao L. Significance of serum microRNAs in pre-diabetes and newly diagnosed type 2 diabetes: a clinical study. Acta Diabetol 2011; 48(1):61-9.

Kosaka N, Iguchi H, Yoshioka Y, Takeshita F, Matsuki Y, Ochiya T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. J Biol Chem 2010; 285:17442–52

Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, Dhanasekaran SM, Chinnaiyan AM, Athey BD. New class of microRNA targets containing simultaneous 59-UTR and 39-UTR interaction sites. Genome Res 2009; 19:1175-1183.

Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochrinic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993; 75:843-854.

Li E, Zhang J, Yuan T, Ma B. miR-145 inhibits osteosarcoma cells proliferation and invasion by targeting ROCK1. Tumour Biology 2014; 35(8):7645-7650.

Lin CC, Liu LZ, Addison JB, Wonderlin WF, Ivanov AV, Ruppert JM. A KLF4-miRNA-206 autoregulatory feedback loop can promote or inhibit protein translation depending upon cell context. Mol Cell Biol 2011; 31:2513–2527.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ. Argonaute2 is the catalytic engine of mammalian RNAi. Science 2004; 305:1437-1441.

Liu X, Fortin K, Mourelatos Z. MicroRNAs: Biogenesis and Molecular Functions. Brain Pathology 2008; 18:113-121.

Longo VD, Antebi A, Bartke A, Barzilai N, Brown-Borg HM, Caruso C, Curiel TJ, de Cabo R, Franceschi C, Gems D, Ingram DK, Johnson TE, Kennedy BK, Kenyon C, Klein S, Kopchick JJ, Lepperdinger G, Madeo F, Mirisola MG, Mitchell JR, Passarino G, Rudolph KL, Sedivy JM, Shadel GS, Sinclair DA, Spindler SR, Suh Y, Vijg J, Vinciguerra M, Fontana L. Interventions to Slow Aging in Humans: Are We Ready? Aging Cell. 2015; 14(4):497-510.

López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell 2013; 153(6):1194–1217.

Madeo F, Pietrocola F, Eisenberg T, Kroemer G. Caloric restriction mimetics: towards a molecular definition. Nat Rev Drug Discov 2014; 13(10):727-40.

Malafarina V, Uriz-Otano F, Iniesta R, Gil-Guerrero L. Sarcopenia in the elderly: diagnosis, physiopathology and treatment. Maturitas 2012; 71(2):109-14.

Maniataki E, Mourelatos Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes Dev 2005; 19:2979–2990.

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci 2008; 105:10513–10518.

Monteys AM, Spengler RM, Wan J, Tecedor L, Lennox KA, Xing Y, Davidson BL. Structure and activity of putative intronic miRNA promoters. RNA 2010; 16(3):495–505.

Morrisette-Thomas V, Cohen AA, Fülöp T, Riesco É, Legault V, Li Q, Milot E, Dusseault-Bélanger F, Ferrucci L. Inflamm-aging does not simply reflect increases in pro-inflammatory markers. Mech Ageing Dev 2014; 139:49-57.

Nawa Y, Kawahara K, Tancharoen S, Meng X, Sameshima H, Ito T, Masuda Y, Imaizumi H, Hashiguchi T, Maruyama I. Nucleophosmin may act as an alarmin: implications for severe sepsis. J Leukoc Biol 2009; 86:645–653.

Noren Hooten N, Fitzpatrick M, Wood WH 3rd, De S, Ejiogu N, Zhang Y, Mattison JA, Becker KG, Zonderman AB, Evans MK. Age-related changes in microRNA levels in serum. Aging 2013; 5(10):725-40.

Oliveira-Carvalho V, da Silva MM, Guimarães GV, Bacal F, Bocchi EA. MicroRNAs: new players in heart failure. Mol Biol Rep. 2013; 40(3):2663-70.

Olivieri F, Spazzafumo L, Santini G, Lazzarini R, Albertini MC, Rippo MR, Galeazzi R, Abbatecola AM, Marcheselli F, Monti D, Ostan R, Cevenini E, Antonicelli R, Franceschi C, Procopio AD. Age-related differences in the expression of circulating microRNAs: miR-21 as a new circulating marker of inflammaging. Mech Ageing Dev 2012; 133(11-12):675-85.

Ørom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 2008; 30(4):460–71.

Ostan R, Lanzarini C, Pini E, Scurti M, Vianello D, Bertarelli C, Fabbri C, Izzi M, Palmas G, Biondi F, Martucci M, Bellavista E, Salvioli S, Capri M, Franceschi C, Santoro A. Inflammaging and cancer: a challenge for the Mediterranean diet. Nutrients 2015; 7(4):2589-621.

Park NJ, Zhou H, Elashoff D, Henson BS, Kastratovic DA, Abemayor E, Wong DT. Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. Clin Cancer Res 2009; 15:5473–5477.

Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Müller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 2000; 408:86-89.

Pegtel DM, Cosmopoulos K, Thorley-Lawson DA, van Eijndhoven MA, Hopmans ES, Lindenberg JL, de Gruijl TD, Würdinger T, Middeldorp JM. Functional delivery of viral miRNAs via exosomes. Proc Natl Acad Sci USA 2010; 107:6328–33.

Picot CR, Moreau M, Juan M, Noblesse E, Nizard C, Petropoulos I, Friguet B. Impairment of methionine sulfoxide reductase during UV irradiation and photoaging. Exp Gerontol 2007; 42(9):859-63.

Pigati L, Yaddanapudi SC, Iyengar R, Kim DJ, Hearn SA, Danforth D, Hastings ML, Duelli DM. Selective release of microRNA species from normal and malignant mammary epithelial cells. PLoS One 2010; 5:e13515

Piva R, Spandidos DA, Gambari R. From microRNA functions to microRNA therapeutics: Novel targets and novel drugs in breast cancer research and treatment. International Journal of Oncology 2013; 43:985-994.

Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, Tait JF, Tewari M. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer Prev Res 2011; 5(3):492-7

Pritchard CC, Cheng HH, Tewari M. MicroRNA profiling: approaches and considerations. Nat Rev Genet 2012; 13(5):358-69.

Ratajczak J, Wysoczynski M, Hayek F, Janowska-Wieczorek A, Ratajczak MZ. Membrane-derived microvesicles: important and underappreciated mediators of cell-to-cell communication. Leukemia 2006; 20:1487–1495.

Raule N, Sevini F, Li S, Barbieri A, Tallaro F, Lomartire L, Vianello D, Montesanto A, Moilanen JS, Bezrukov V, Blanché H, Hervonen A, Christensen K, Deiana L, Gonos ES, Kirkwood TB, Kristensen P, Leon A, Pelicci PG, Poulain M, Rea IM, Remacle J, Robine JM, Schreiber S, Sikora E, Eline Slagboom P, Spazzafumo L, Antonietta Stazi M, Toussaint O, Vaupel JW, Rose G, Majamaa K, Perola M, Johnson TE, Bolund L, Yang H, Passarino G, Franceschi C. The co-occurrence of mtDNA mutations on different oxidative phosphorylation subunits, not detected by haplogroup analysis, affects human longevity and is population specific. Aging Cell 2014; 13(3):401-7.

Rechavi O, Erlich Y, Amram H, Flomenblit L, Karginov FV, Goldstein I, Hannon GJ, Kloog Y. Cell contact-dependent acquisition of cellular and viral nonautonomously encoded small RNAs. Genes Dev 2009; 23:1971–9.

Reinhart BJ, Slack FJ, Basson M, Bettinger JC, Pasquinelli AE, Rougvie AE, Horvitz HR and Ruvkun G. The 21 nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 2000; 403:901-906.

Salvioli S, Capri M, Valensin S, Tieri P, Monti D, Ottaviani E, Franceschi C. Inflamm-aging, cytokines and aging: state of the art, new hypotheses on the role of mitochondria and new perspectives from systems biology. Curr Pharm Des 2006; 12(24):3161-71.

Sanders I, Holdenrieder S, Walgenbach-Brünagel G, von Ruecker A, Kristiansen G, Müller SC, Ellinger J. Evaluation of reference genes for the analysis of serum miRNA in patients with prostate cancer, bladder cancer and renal cell carcinoma. Int J Urol 2012; 19(11):1017-25.

Schlosser K, McIntyre LA, White RJ, Stewart DJ. Customized Internal Reference Controls for Improved Assessment of Circulating MicroRNAs in Disease. PLoS One. 2015; 10(5):e0127443.

Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N and Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell 2003; 115:199–208.

Simm A, Müller B, Nass N, Hofmann B, Bushnaq H, Silber RE, Bartling B. Protein glycation - Between tissue aging and protection. Exp Gerontol 2015; 68:71-5.

Tan L, Yu JT, Liu QY, Tan MS, Zhang W, Hu N, Wang YL, Sun L, Jiang T, Tan L. Circulating miR-125b as a biomarker of Alzheimer's disease. J Neurol Sci 2014; 336(1-2):52-6.

Tsai N, Lin Y, Wei L. MicroRNA mir-346 targets the 5′-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression. Biochemical Journal 2009; 424:411–418.

Turchinovich A, Weiz L, Burwinkel B. Extracellular miRNAs: the mystery of their origin and function. Trends Biochem Sci. 2012; 37(11):460-465.

Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ, Lotvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol 2007; 9:654–59.

Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M. Mechanisms of miRNA-mediated gene regulation from common downregulation to mRNA-specific upregulation. Int J Genomics 2014; 2014:970607.

Vanhooren V, Dewaele S, Libert C, Engelborghs S, De Deyn PP, Toussaint O, Debacq-Chainiaux F, Poulain M, Glupczynski Y, Franceschi C, Jaspers K, van der Pluijm I, Hoeijmakers J, Chen CC. Serum N-glycan profile shift during human ageing. Exp Gerontol 2010; 45(10):738-43.

Vasudevan S, Steitz JA. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. Cell 2007; 128:1105–1118.

Vickers KC, Palmisano BT, Shoucri BM, Shamburek RD, Remaley AT. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. Nat Cell Biol 2011; 13:423–433.

Wang GK, Zhu JQ, Zhang JT, Li Q, Li Y, He J, Qin YW, Jing Q. Circulating microRNA: a novel potential biomarker for early diagnosis of acute myocardial infarction in humans. Eur Heart J 2010; 31(6):659-66.

Wang K, Zhang S, Marzolf B, Troisch P, Brightman A, Hu Z, Hood LE, Galas DJ. Circulating microRNAs, potential biomarkers for drug-induced liver injury. Proc Natl Acad Sci 2009; 106:4402–4407.

Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, Galas DJ, Wang K. The microRNA spectrum in 12 body fluids. Clin Chem 2010; 56(11):1733-41

Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012; 41(1):24-32.

Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 1993; 75(5):855-62.

Yamada A, Cox MA, Gaffney KA, Moreland A, Boland CR, Goel A. Technical factors involved in the measurement of circulating microRNA biomarkers for the detection of colorectal neoplasia. PLoS One 2014; 9(11):e112481.

Yuan A, Farber EL, Rapoport AL, Tejada D, Deniskin R, Akhmedov NB, Farber DB. Transfer of microRNAs by embryonic stem cell microvesicles. PLoS ONE 2009; 4:e4722.

Zacharewicz E, Lamon S, Russell AP. MicroRNAs in skeletal muscle and their regulation with exercise, ageing, and disease. Front Physiol 2013; 4:266.

Zampetaki A, Kiechl S, Drozdov I, Willeit P, Mayr U, Prokopi M, Mayr A, Weger S, Oberhollenzer F, Bonora E, Shah A, Willeit J, Mayr M. Plasma microRNA profiling reveals loss of endothelial miR-126 and other microRNAs in type 2 diabetes. Circ Res 2010;1 07(6):810-7.

Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Bürkle A, Caiafa P. Reconfiguration of DNA methylation in aging. Mech Ageing Dev 2015; 151:60-70.

Zhang H, Yang H, Zhang C, Jing Y, Wang C, Liu C, Zhang R, Wang J, Zhang J, Zen K, Zhang C, Li D. Investigation of microRNA expression in human serum during the aging process. Gerontol A Biol Sci Med Sci 2015; 70(1):102-9.

Zhang Y, Fan M, Zhang X, Huang F, Wu K, Zhang J, Liu J, Huang Z, Luo H, Tao L, Zhang H. Cellular microRNAs up-regulate transcription via interaction with promoter TATA-box motifs. RNA 2014; 20(12):1878-89.

APPENDIX A

List of target genes identified by pathway analysis of miRs-598-3p; -133a-3p;-206;-16

BCL2   B-cell lymphoma 2

BRCA1 BReast CAncer gene one

CASP9 Caspase 9

CCND1 Cyclin D1

CCND3 Cyclin D3

CDC37 Cell division cycle 37

CDK6   Cyclin-dependent kinase 6

CREB5 cAMP responsive element binding

EIF4E   Eukaryotic translation initiation factor 4E

EPHA2 Ephrin type-A receptor 2

GNB1   Guanine Nucleotide Binding Protein, Beta Polypeptide 1

GRB2   Growth factor receptor-bound protein 2

HSP90B1 Heat shock protein 90kDa beta member 1

IKBKB Inhibitor of nuclear factor kappa-B kinase subunit beta

ITGA2 Integrin subunit alpha 2

ITGB4 Integrin subunit beta 4

KRAS   Kirsten rat sarcoma viral oncogene homolog

LAMC2 Laminin subunit gamma 2

MCL1   Myeloid cell leukemia 1

MTOR Mammalian target of rapamycin

MYB   Myeloblastosis proto-oncogene protein

PCK2   Phosphoenolpyruvate carboxykinase 2

PDGFA Platelet-derived growth factor A

PIK3R2 Phosphoinositide-3-Kinase, Regulatory Subunit 2 (Beta)

PPP2R1A Protein Phosphatase 2, Regulatory Subunit A

PRLR   Prolactin receptor

RPS6   Ribosomal protein s6

RPS6KB1 Ribosomal protein S6 kinase, polypeptide 1

SGK3   Serum/Glucocorticoid Regulated Kinase Family, Member 3

SOS1   Son Of Sevenless Homolog 1

THEM4 Thioesterase Superfamily Member 4

YWHAZ 14-3-3 protein zeta/delta