TRANSCRIPTOMICS ANALYSIS OF CODING GENES AND LONG NON-CODING RNAS IN AGEING IN *Drosophila melanogaster*

MOHAMMAD MORTEZA HOSSEINZADEH BAJGIRAN

UNIVERSITI SAINS MALAYSIA

2022

TRANSCRIPTOMICS ANALYSIS OF CODING GENES AND LONG NON-CODING RNAS IN AGEING IN *Drosophila melanogaster*

by

MOHAMMAD MORTEZA HOSSEINZADEH BAJGIRAN

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

January 2022

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my supervisor Dr. Intan Haslina and my co-supervisor Professor Dr. Shaharum Shamsuddin for their help and support. Especially, I have to thank my co-supervisor Associate Professor Dr. Ghows Azzam for his advice, support and guidance. This thesis would not have been written without his support, patience and encouragements. No words can adequately describe my debt to him. I would like to dedicate this thesis to my devotee parents. I should thank Dr. Mardani for his support and all my lab mates those who I enjoyed so much by their friendliness. Finally, I would like to acknowledge financial support from Malaysian Ministry of Higher Education grant (FRGS: 203.PPSK.6171226).

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

°C	Degree Celsius
%	Percentage
×	Times (multiplier)
M	Molar
mg	Milligrams
mL	Milliliters
nm	Nanometer
µL	Microliters
µm	Micrometers
g	Grams

LIST OF ABBREVIATIONS

D1	1 day old
D30	30 days old
D60	60 days old
Ogg1	8-oxoguanine DNA glycosylase
EC	Absorptive enterocyte
ANT	Adenine nucleotide translocase
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
AMPK	AMP-activated kinase
BER	Base excision repair
cDNA	Complementary DNA
COX4L	Cytochrome c oxidase subunit 4-like
DR	Dietary restriction
DDR	DNA damage responses
DSB	Double-strand break repair
ESC	Embryonic stem cell
ER	Endoplasmic reticulum
EB	Enteroblast progenitor
ECM	Extracellular matrix
FDR	False discovery rate
F1	First filial
FRKM	Fragments Per Kilobase of transcript per Million
GO	Gene ontology
Gnf1	Germ line transcription factor 1
GSC	Germline stem cell
Gnu	Giant nuclei
GR	Glucocorticoid receptor
Gnmt	Glycine N-methyltransferase
GnRH	Gonadotropin releasing hormone
Grp	Grapes
GAS5	Growth arrest-specific 5

HSF1	Heat shock factor 1
Hsp	Heat-shock protein
HCC	Hepatocellular carcinoma cell
HR	Homologous recombination
HOTAIR	HOX transcript antisense RNA
HDF	Human diploid fibroblast
IIS	Insulin and IGF signalling
IGF-1	Insulin like growth factor 1
IGF-2	Insulin-like growth factor 2
IGFR	Insulin/IGF-1 transmembrane receptor
ISC	Intestinal stem cell
LPS	Lipopolysaccharide
lncRNA	Long non-coding RNA
mTOR	Mechanistic target of rapamycin
MSC	Mesenchymal stem cells
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
Met	Methionine
MBD1	Methyl-CpG–binding domain protein 1
miRNA	MicroRNA
MMR	Mismatch repair
OXPHOS	Mitochondrial oxidative phosphorylation
mtSSB	Mitochondrial single stranded DNA-binding protein
MuSC	Muscle stem cell
m6A	N6-methyladenosine
NAD	Nicotinamide adenine dinucleotide
Not	Nitric oxide synthase
NHEJ	Nonhomologous end joining
NER	Nucleotide excision repair
OD	Optical density
OreR	Oregon-R
PANDA	P21-associated noncoding RNA DNA damage activated
Png	Pan gu
IPSC	Pluripotent stem cell
Plu	Plutonium
PCR	Polymerase chain reaction

Pont	Pontin
Pro	Pre-mRNA processing
PCNA	Proliferating cell nuclear antigen
PCNA2	Proliferating cell nuclear antigen 2
qPCR	Quantitative polymerase chain reaction
ROS	Reactive oxygen species
RS	Replicative senescence
Rbf	Retinoblastoma-family protein
RMST	Rhabdomyosarcoma 2- associated transcript
RNA-seq	RNA sequencing
NeST	Salmonella pas Theiler's
EE	Secretory enteroendocrine cells
SAHF	Senescence-associated heterochromatin foci
SASP	Senescence-associated secretory phenotype
SSB	Single-strand break repair
snRNP	Small nuclear ribonucleo protein
snRNA	Small nuclear RNA
SuUR	Suppressor of Under-Replication
TERT	Telomerase reverse transcriptase
TERC	Telomerase RNA component
Tefu	Telomere fusion
GO	The gene ontology
TNFα	Tumour necrosis factor alpha
UCLH1	Ubiquitin carboxy-terminal hydrolase L1
XIST	X-inactive specific transcript XIST

LIST OF APPENDICES

ANALYSIS TRANSKRIPTOMIK GEN PENGEKOD DAN RNA PANJANG BUKANPENGEKOD DALAM PENUAAN *Drosophila melanogaster*

ABSTRAK

Penuaan dicirikan oleh kehilangan fungsi fisiologi dan selular, kehilangan struktur tisu, dan kemerosotan interaksi molekular, yang membawa kepada penyakit kronik seperti diabetes, kanser, penyakit neurodegeneratif, dan penyakit lain yang berkaitan dengan usia. Disebabkan oleh pemuliharaan genetik, gen dan laluan yang mendasari penuaan didapati mempunyai corak yang sama dalam spesies yang berbeza. Gen ini dan laluan telah dikategorikan kepada sembilan ciri selular dan molekular, iaitu, ketidakstabilan genomik, atrisi telomere, kehilangan proteostasis, disfungsi mitokondria, perubahan epigenetik, penderiaan nutrien yang tidak dikawal atur, keletihan sel stem, selular penuaan dan komunikasi antara sel yang diubah. Objektif utama penyelidikan ini, adalah untuk memahami bagaimana tahap ekspresi gen pengekodan berubah semasa penuaan dan untuk mengenal pasti lncRNA yang terlibat dalam proses penuaan. Untuk menentukan titik masa terbaik yang mewakili dewasa muda, dewasa pertengahan umur dan dewasa tua *D. melanogaster*, lengkung kelangsungan hidup ditentukan dan dengan itu tiga titik masa, iaitu hari 1, hari 30, dan hari 60, telah dipilih. Selepas itu, pemetaan transkriptom sekitar genom bagi proses penuaan dalam *D. melanogaster* telah dilakukan, di mana, analisis transkriptom telah dijalankan pada titik masa yang ditentukan sebagai muda, pertengahan umur dan lalat tua. Ekspresi pembezaan dan analisis pemerkayaan menunjukkan bahawa bilangan gen pengekodan yang tinggi terlibat dalam pengubahsuaian splisosom dan epigenetik adalah dikawalselia secara menaik akibat penuaan. Pengawalseliaan gen

secara meningkat yang signifikan berkait dengan pengubahsuaian splisosom dan epiginetik berpotensi menyumbang kepada perubahan dalam pengawalaturan gen dan isoform yang sepadan membawa kepada peningkatan ketidakstabilan genomik. Begitu juga, peningkatan ketara dalam tahap pengekspresan gen yang terlibat dalam tindak balas kerosakan DNA telah diperhatikan. Gen dikawalselia secara menaik yang terlibat dalam tindak balas kerosakan DNA kebanyakannya adalah antara komponen teras laluan pembaikan kerosakan DNA yang berbeza termasuk pembaikan ketidakpadanan, pembaikan eksisi nukleotida, dan pembaikan eksisi dasar. Dalam nada yang sama, data menunjukkan bahawa komponen utama bebenang ganda dua memecahkan laluan pembaikan DNA termasuk Ku70, Nse1, Spn-D, dan Spn-A dikawalselia secara meningkat dengan signifikan, walau bagaimanapun, pengawalseliaan secara menaik adalah lebih penting dalam pembaikan rekombinan homolog berbanding dengan penyambungan hujung tidak homolog. Selain itu, analisa menunjukkan bahawa gen berkaitan dengan fasa S dan pusat pemeriksaan kerosakan mitosis DNA G2 termasuk Rad9, Cdk1, tefu, Tctp, mms4 dan mre11 mempunyai tahap ekspresi yang lebih tinggi dalam lalat tua, membayangkan kadar yang lebih tinggi tekanan replikasi dan hentian kitaran sel dalam fasa G2 kitaran sel. Lebih-lebih lagi, ekspresi pengawalseliaan menurun secara am bagi kebanyakan gen yang terlibat dalam kompleks fosforilasi oksidatif mitokondria (OXPHOS), respirasi serta perubahan dalam komposisi lipid membran dalam mitokondria telah ditentukan. Penurunan pengawalseliaan menurun gen berkaitan mitokondria secara besar-besaran, berpotensi, membawa kepada disfungsi mitokondria dan kadar kerosakan oksidatif yang lebih tinggi. Selepas menjalankan analisis pembezaan ekspresi gen pengekodan, analisis lanjut dijalankan untuk mengenal pasti lncRNA yang terlibat dalam penuaan dan untuk menentukan potensi peranan regulasi dalam penuaan proses. Analisis menunjukkan bahawa

lncRNA yang diekspres secara berbeza bersebelahan dengan gen yang berkaitan dengan fosforilasi oksidatif dan pembaikan ketidakpadanan adalah diperkaya, memungkinkan bahawa lncRNA terlibat dalam proses penuaan melalui disfungsi mitokondria dan ketidakstabilan genomik. Bersama, kajian ini menunjukkan bahawa peristiwa pencantuman yang tidak dikawal, pengubahsuaian epigenetik yang diubah, tekanan replikasi dan disfungsi mitokondria memainkan peranan penting dalam meningkatkan ketidakstabilan genomik pada lalat tua yang seterusnya membawa kepada ekspresi berlebihan gen yang terlibat dalam tindak balas kerosakan DNA. Ini menunjukkan bahawa, lncRNA mempunyai penglibatan penting dalam proses penuaan dengan memainkan peranan pengawalseliaan dalam laluan yang berkaitan dengan fungsi mitokondria dan pembaikan DNA. Penemuan ini memberikan gambaran yang lebih jelas tentang gen dan laluan yang terlibat dalam penuaan dan menyumbang untuk mewujudkan rangka kerja resolusi yang lebih tinggi untuk kajian masa depan tentang penuaan dan penyakit berkaitan usia.

TRANSCRIPTOMICS ANALYSIS OF CODING GENES AND LONG NON-CODING RNAS IN AGEING IN *Drosophila melanogaster*

ABSTRACT

Ageing is characterised by the loss of physiological and cellular functions, the loss of tissue structure, and deterioration of molecular interactions, leading to chronic diseases such as diabetes, cancers, neurodegenerative diseases, and other age-related diseases. Due to the genetic conservation, the genes and pathways underlying ageing were found to have comparable pattern across different species. These genes and pathways have been categorised into nine cellular and molecular hallmarks, namely, genomic instability, telomere attrition, loss of proteostasis, mitochondrial dysfunction, epigenetic alterations, deregulated nutrient sensing, stem cell exhaustion, cellular senescence and altered intercellular communication. The main objectives of this research are to understand how the expression levels of coding genes change during ageing and to identify lncRNAs involved in the process of ageing. To determine the best time points that represent young adult, middle age adult and old adult of *D. melanogaster*, the survival curve was determined and accordingly three time points, day 1, day 30, and day 60, were chosen. Afterwards, genome wide transcriptome mapping of ageing process in *D. melanogaster* was performed, in which, transcriptomic analysis was conducted on determined time points as young, middle age and old flies. Differential expression and enrichment analysis showed that high number of coding genes involved in spliceosome and epigenetic modifications are up-regulated due to ageing. The significant up-regulation of genes related to spliceosome and epiginetic modifications potentially contribute to changes in regulation of genes and their corresponding isoforms leading to increased genomic instability. Likewise,

a significant increase in expression levels of genes invloved in DNA damage reponse was observed. The up-regulated genes involved in DNA damage response mostly are among core components of different DNA damage repair pathways including mismatch repair, nucleotide excision repair, and base excision repair. In the same vein, the data showed that major components of double strand break DNA repair pathways including *Ku70, Nse1, Spn- D*, and *Spn-A* were significantly up-regulated, however, the up-regulation was more significant in homologous recombination repair compared to non-homologous end joining. Additionally, the analysis showed that genes related to S phase and mitotic G2 DNA damage checkpoint including *Rad9, Cdk1, tefu, Tctp, mms4* and *mre11* have higher expression levels in aged flies, implying the higher rate of replication stress and cell cycle arrest in G2 phase of cell cycle. Moreover, a general down-regulation in expression of most genes involved in *mitochondrial oxidative phosphorylation (OXPHOS)* complexes, respiration and as well as changing in lipid composition of inner mitochondria membrane was determined. The massive down-regulation of mitochondria related genes, potentially, leads to mitochondria dysfunction and higher rates of oxidative damages. After conducting differential expression analysis of coding genes, further analysis was performed to identify lncRNAs involved in ageing and to determine their potential regulatory roles in ageing process. The analysis showed that differentially expressed lncRNAs neighbouring the genes related to oxidative phosphorylation and mismatch repair are enriched, implying that lncRNAs are involved in ageing process through mitochondrial dysfunction and genomic instability. Together, this study shows that deregulated splicing events, altered epigenetics modifications, replication stress and mitochondrial dysfunction play important roles in increasing the genomic instability in aged flies which in turn leads to over-expression of genes involved in DNA damage response. lncRNAs have

significant involvement in ageing process by playing regulatory roles in pathways related in mitochondrial functions and DNA repair. These findings provide a clearer picture of the genes and pathways involved in ageing and contributes to establish a higher resolution framework for future studies on ageing and age-related diseases.

CHAPTER 1

INTRODUCTION

Ageing can be defined as the process of getting chronologically older or progressively decline in physiological functions. Ageing also can be represented by the loss of tissue structure and function, declining in molecular interaction as well as cellular functions (López-Otín et al., 2013a; Wang et al., 2014). Cell dysfunctions may lead to chronic diseases such as diabetes, cancers and other age-related diseases.

Thanks to the advancement of healthcare, human population is getting older and life expectancy is increasing. According to the United Nations Department of Economic and Social Affairs (UNDESA), population aged 60 or more from 900 million people in 2015 will increase to 2 billion by 2050 (UN, 2015). However, the increase in life expectancy is not associated with significant improvement in the rate of disabilities, hence, the future aged population will bring a pressure on healthcare systems and governments. As a result, understanding the biological mechanism of aging and age-related disease is required in order to reduce the negative social, economic and health consequences of aging.

From genetic point of view, however, non-genetic factors such as environment and nutrition affect ageing, approximately 35% of the variance in lifespan of mammals and invertebrate caused by heritable components. Several studies over the past decades were done to identify longevity genes in human, including candidate-gene association analysis, linkage analysis, and longitudinal studies (Barger *et al.*, 2003; Nybo *et al.*, 2003). Accordingly, several coding genes implicated in ageing such as Clock (Clk), insulin/IGF-I receptor Daf-2, Apolipoprotein C3 (APOC3), and Paraoxonase 1 (PON1), and several pathways involved in ageing such as target of rapamycin (TOR) and insulin/insulin-like growth factor signaling (IIS) were identified. These genes and

pathways were classified into nine main hallmarks of ageing, namely, genomic instability, telomere attrition, ,loss of proteostasis, epigenetic alterations, deregulated nutrient sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication (López-Otín *et al*., 2013b). Despite recent progresses in identifying coding genes and pathways implicated in ageing, due to the complexity of the aging process, the underlying mechanism of ageing and the molecular connection between different hallmarks still remain unknown (López-Otín *et al*., 2013b). Additionally, recent discoveries in the field of non-coding RNAs revealed a new layer of complexity in ageing process, which had been disregarded before.

Non-coding RNAs includes, but not limited, ribosomal RNA, small nucleolar RNA, transfer RNA, micro RNAs, PIWI-interacting RNAs and long noncoding RNAs (lncRNAs) (Palazzo *et al*., 2015). lncRNAs are transcripts longer than 200 bp that are not coding any proteins and play important regulatory roles in different biological processes including epigenetic regulation, transcription and post-transcription regulation (Kapranov *et al.*, 2007,L. L. Chen & Carmichael, 2010; Clark *et al*., 2012; Wilusz *et al*., 2009). Several lncRNAs have been demostrated to be involved in ageing-associated hallmarks. For example, several lncRNAs were reported to be involved in cellular senesence such as X-inactive specific transcript (XIST) and metastasis associated lung adenocarcinoma transcript 1 (MALAT1), or lncRNAs involved in stem cell exhustion such as lncRNA AK028326 and ES1(He & Liu, 2018). In spite of their importance and association with ageing process, our knowledge on the role of lncRNAs in ageing are still scarce. Therefore, investigating the role of lncRNAs in the ageing could provide insight into the complexity of this process and facilitate the discovery of new anti-ageing targets and biomarkers.

Most of the previous studies into the molecular mechanisms of aging have focused on a particular pathway or hallmark of ageing, while individual molecular event or pathway cannot adequately decipher the aging process (McAuley *et al.*, 2017). As a result, a comprehensive research, considering a large number of genes and pathways, is needed to link different molecular mechanism involved in ageing. Fortunately, next-generation sequencing (NGS) technologies helps to comprehensively study the expressional changes of coding genes during ageing. Additionally, using multiple bioinformatic tools and pipelines help to study the role of lncRNAs alongside coding genes in order to get clearer picture of the molecular events underlying ageing.

Although NGS helps to comprehensively study the genes and pathways involved in ageing and to shed light on the complexity of ageing process, Controlling various environmental conditions such as lifestyle and diet make the human ageing research challenging (Troen, 2003). Model organisms offer much more flexibility in term of controlling variations in experiment setup for ageing studies. *D. melanogaster* has been used to study aging for more than 100 years now (Taormina *et al*., 2019). They have short lifespan, low maintenance, genetically amenable and fully sequenced genome (Partridge & Tower, 2008). Near 50% of *D. melanogaster* has similarity to humans (Myers *et al*., 2000). Besides, near 75% of human diseases and disorders can be expressed in flies (Reiter *et al.*, 2001). Therefore, *D. melanogaster* is a perfect model to study human aging and age-related diseases.

Altogether, transcriptomics analysis of ageing process in flies can help to extensively study the role of coding genes and lncRNAs in ageing. In this study, genome wide transcriptome mapping of aging process in *D. melanogaster* in different ages, namely, day 1, day 30, and day 60 was performed to investigate the expression pattern of coding genes and lncRNAs during ageing with greater resolution which is

expected to contribute towards a better understanding of ageing process as well as to provide a clearer picture of the role of non-coding genes in ageing. The present study will facilitate the future research for finding anti-ageing strategies and treatments for age-related diseases.

1.1 Objectives

This research has two main objectives:

1. To determine coding genes and pathways involved in process of ageing in *D. melanogaster*.
2. To determine the involvement of lncRNAs in ageing in *D. melanogaster* and identify their potential cis-regulated coding genes.

CHAPTER 2

LITERATURE REVIEW

2.1 Ageing and its hallmarks

Aging is a complex multifactorial process characterized by deterioration of physical, and biological condition of an organism with the passage of time. The deterioration occurs at tissue, cellular and molecular levels and contributes to several age-related chronic diseases including cardiovascular malfunction, neurodegenerative disease and cancer (Bao et al., 2014; Liu et al., 2019).

The biological causes of aging are uncertain. Advances in aging research have uncovered a wide range of molecular and cellular events involved in aging. These events were classified into nine hallmarks (López-Otín *et al*., 2013). Hallmarks of ageing can be categorised into three groups, namely, primary, antagonistic, and integrative hallmarks. Primary hallmarks are genomic instability, telomere attrition, loss of proteostasis, and epigenetic alterations which are the causes of functional damages in cells. These damages result in antagonistic hallmarks including deregulated nutrient sensing, mitochondrial dysfunction, and cellular senescence which finally lead to integrative hallmarks, i.e., stem cell exhaustion, and altered intercellular communication (Aunan *et al.*, 2016; López-Otín *et al.*, 2013) (Figure 2.1). Each of said hallmarks will be detailed in this section.

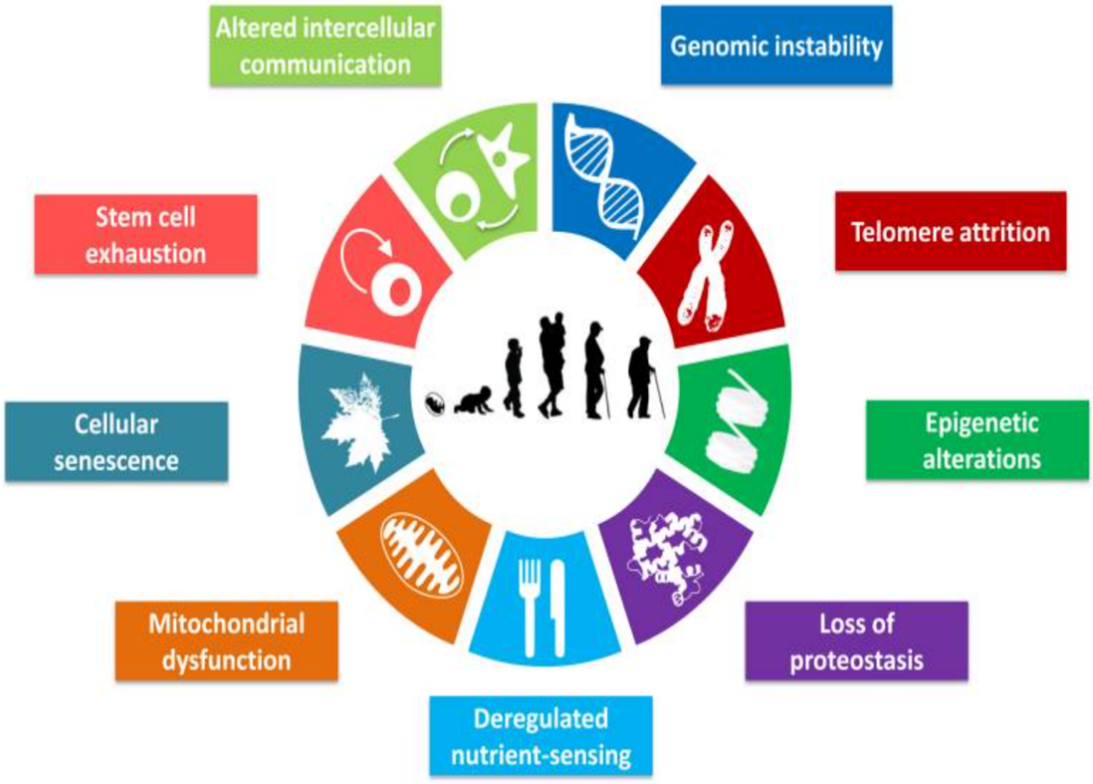

Figure 2-1. Nine main hallmarks of ageing. Image was taken from López-Otín *et al*., 2013

2.1.1 Genomic instability

Genomic stability is an important characteristic of organisms to ensure the preservation of genome integrity through the generations. This stability depends on an accurate replication system and a DNA repair system to correct the replication mistakes. In contrast genome instability refers to the appearance of a wide range of alternations in genome including point mutations, insertion and deletion, and chromosomal rearrangements (Kovalchuk, 2016).

In the 1950s Failla and Szilard separately suggested that random mutations can lead to gene inactivation and cause aging (Failla, 1958; Szilard, 1959). Later, it was evidenced that many age-related diseases such as cancer, cardiovascular diseases, autoimmune diseases, and neurodegenerative diseases are resulted from accumulation of mutations and DNA damages (Hanisch & Kettenmann, 2007; McMurray & Gottschling, 2003; Nowell, 1976; Vijg & Suh, 2013). Therefore, genomic instability has been classified as one of the primary hallmarks of aging which disturbs the cellular processes and causes aging (López-Otín *et al*., 2013b). The age associated genomic instability is the consequence of long-term exposure to a broad range of endogenous and exogenous DNA damaging factors. The most important endogenous DNA damaging factors are oxidative stress, hydrolysis, and error-prone DNA replication and repair (Vijg & Suh, 2013).

2.1.1(a) Oxidative DNA damage

Oxidative stress contributes significantly in aging as it was proposed in the "free radical theory of aging". This theory is in accordance with previous findings that metabolic activity inversely correlates with longevity via producing reactive oxygen species (ROS) and increasing oxidative damage (Harman, 1956). Oxidative

phosphorylation in the inner mitochondrial membrane generates different types of ROS including superoxide, hydrogen peroxide, hydroxyl radicals, and singlet oxygen. Thereafter, these ROS oxidise the DNA resulting in DNA damage (Vijg & Suh, 2013). Moreover, other evidences including an accumulation of oxidative DNA damage and an increase in mitochondrial ROS production with age support the contribution of oxidative stress in the process of aging (Lee & Wei, 2013; Moskalev *et al*., 2013).

2.1.1(b) Hydrolysis

Another endogenous cause of DNA damage is hydrolysis. As a result of continuous exposure of DNA to water, the glycosylic bonds are hydrolysed. This hydrolysis happens up to 10,000 sites per cell per day that can lead to depurination and depyrimidination producing an abasic site (Lindahl & Nyberg, 1972). Besides, deamination sometimes causes point mutations. Especially, deamination of 5-methylcytosine forms thymine which can result in transition mutation. The reason that this conversion can result in mutation is that the thymine created from deamination of 5-methylcytosine is a normal nucleotide, hence, it is difficult for DNA repair system to detect it as a damage (Cooper & Youssoufian, 1988). Taking into account the significant role of CpG methylation in epigenetic regulation, deamination of 5-methylcytosine not only can cause point mutation, but also it can have epigenetic regulatory effects (Riggs, 1975).

2.1.1(c) Replication stress

In addition to ROS and hydrolysis, DNA is at risk of replication error and replication stress. However, the cell has several mechanisms to preserve the accuracy of DNA replication, there is still about 1 error in 100,000 bases which 99% of them

are repaired by proofreading process. Additionally, DNA template and synthesizing DNA have to align accurately during replication, otherwise, it can lead to some deletion or insertion *(*Kunkel, 2009; Streisinger *et al.*, 1966). Replication stress also can result in DNA damage. Replication stress refers to any event or condition that causes slowing or stopping of DNA replication. For instance, reduced nucleotide pool, unrepaired DNA damage, impaired nucleosome assembly, and mutation in genes involved in replication can slow down or stall the replication (Burhans & Weinberger, 2007). Long-term replication stalling can lead to replication fork collapse which in turn induces single-strand and double-strand DNA breaks. Repairing of double-strand DNA breaks usually is via homologous recombination (HR), however, in the absence of HR nonhomologous end joining (NHEJ) would repair the damage (Aguilera & García-Muse, 2013). Nonetheless, deterioration of DNA repair pathways in aging has been demonstrated, causing the accumulation of DNA damages resulted from replication errors (Alejandro Lagunas-Rangel & María Bermúdez-Cruz, 2019).

2.1.1(d) Errornous DNA repair

As stated above, genome-maintenance mechanisms and DNA repair pathways deteriorate in aging (Alejandro Lagunas-Rangel & María Bermúdez-Cruz, 2019). Depending on the type of damage, cells employ different DNA damage responses (DDR) in order to detect DNA lesions and trigger subsequent repair pathways (Jackson & Bartek, 2009a). Thus, changes in regulation of genes involved in DDR can lead to accumulation of DNA damages and genomic instability. It has been shown also that cells with defects in genes related to DDR accelerate aging process (Gorbunova *et al.*, 2007).

There are different type of DNA repair mechanisms including mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), and double-strand break repair (DSB) (Chatterjee & Walker, 2017). MMR is responsible for repairing DNA lesions resulted from erroneous DNA replication or repair. MMR defects cause point mutations or microsatellite instability. Previous studies have shown a decline in MMR ability due to aging (Annett et al., 2005; Krichevsky *et al*., 2004; Neri *et al.*, 2005). BER pathway repair DNA damages caused by oxidation, deamination, and single strand breaks (SSB) (Krokan & Bjørås, 2013a). It has been shown that efficiency of BER pathway decline with age and it considerably affects the longevity (Debrabant *et al.*, 2014). Besides, defective BER genes increase the risk of cancer in mice and decrease the longevity in yeast (Cabelof *et al.*, 2006; Maclean *et al.*, 2003). NER corrects the oxidative damages caused by UV. The ability of skin fibroblasts to repair UV-induced damages decrease with age, suggesting a decrease in efficiency of NER pathway in aging (Genome Stability, 2016). Moreover, defects in genes involved in NER leads to premature aging syndromes in humans (Goukassian, 2000). DSBs are generated as a consequence of a wide range of damaging agents including oxidation, radiation, and replication errors which can result in loss of chromosomal segments. There are two pathways responsible for DSB repair: homologous recombination (HR) and nonhomologous end joining (NHEJ). While HR repair the damage accurately, NHEJ is more subject to errors resulting in deletion or addition of nucleotides (Chatterjee & Walker, 2017). It is worth noting that HR mostly is active in S and G2 phases of cell cycle, however, NHEJ is responsible for repairing in G1 phase and play more important role in senescent cells (Siebert & Puchta, 2002). Interestingly, HR repair decline with age which would result in utilizing NHEJ pathway by cell in order to repair DSBs (Frasca et *al.*, 1999). At the same time, cellular

senescence in aging can increase the activity of NHEJ which in turn can cause more damages (Feng *et al*., 2007).

2.1.2 Telomere attrition

Telomeres are the specific regions with repetitive nucleotide sequences at both end of chromosomes (Van Steensel *et al.*, 1998). Since DNA polymerase is unable to replicate terminal ends of DNA, telomeres shorten at each DNA replication. Thus, after several replication round, the telomere will reach a critical length which is not functional anymore and causes chromosomal degradation and cell death (Jackson & Bartek, 2009b; Shin *et al*., 2006). In addition, DNA damage to telomeres cannot be detected by DNA repair machinery because of a nucleoprotein complex known as shelterin. As a result, not only telomere become shorter with time but also DNA damages can accumulate in telomere regions resulting in telomere deficiency which can lead to cellular senescence and aging (Palm & De Lange, 2008). Previous studies showed that shortening the telomeres of mice decreases their lifespan and lengthening them causes an increase in lifespan (Armanios *et al.*, 2009; Rudolph *et al.*, 1999; Tomás-Loba *et al*., 2008).

2.1.3 Loss of proteostasis

Another molecular feature of ageing is deterioration of protein quality control systems leading to expression of unfolded, misfolded or aggregated proteins (Balch, 2008; Proctor & Lorimer, 2011). It has been evidenced that loss of proteostasis results in some age-related diseases including Alzheimer's and Parkinson's (Powers *et al.*, 2009). For instance, studies have shown that chaperone-mediated protein folding as

one of the core mechanisms that regulates protein homeostasis deteriorates with age (Calderwood *et al.*, 2009). Besides, impaired autophagy and decreased proteasomal activity have been observed in ageing, leading to accumulation of damaged proteins and organelles in old cells (Rubinsztein *et al*., 2011; Tomaru *et al*., 2012).

2.1.4 Epigenetic alterations

Epigenetic alterations include changes in DNA methylation pattern, histone modification, and chromatin remodelling (Berger *et al.*, 2009). Previous studies have shown that association of epigenetic alterations with genetic changes can result in wide range of age-related diseases and affect the lifespan (Gonzalo, 2010). In particular, alteration in pattern of 5-methylcytosine which is the most common form of DNA methylation has been demonstrated during ageing (Greenberg & Bourc'his, 2019; Horvath, 2013; E. Li & Zhang, 2014). Additionally, recent studies have indicated a loss of constitutive heterochromatin and heterochromatin redistribution during aging, implying the alteration of histones modifications with age (Celona *et al*., 2011; Jambhekar *et al.*, 2019; Sen *et al.*, 2016).

Aside from DNA and histones, RNAs can be modified epigenetically (Barbieri *et al.*, 2017; Berulava *et al.*, 2015; Squires *et al*., 2012). Particularly, RNA methylation to form N6-methyladenosine (m6A) in mRNA is responsible for the most common epigenetic RNA modification which has a wide range of roles in regulation of gene expression (Zaccara *et al.*, 2019). Abnormality of m6A mRNA methylation in different age-related diseases including cancer and Alzheimer were observed (Han *et al*., 2020; Lan *et al*., 2019). The role of m6A methylation in ageing is still unclear, however, m6A RNA profiling of human peripheral blood mononuclear cells identified a decrease in RNA methylation levels with age (Min *et al.*, 2018).

2.1.5 Cellular senescence

One of the strategies that somatic cells developed in order to respond to the stress is cellular senescence which means irreversible cell cycle arrest (Kuilman *et al.*, 2010). Although, cellular senescence prevents the transmission of mutations and damages, the accumulation of senescent cells can lead to a wide range of age-related disease (Grimes & Chandra, 2009; Kuilman *et al*., 2010; Salama *et al*., 2014). According to nature of the stimuli, cellular senescence can be classified as telomere dependent senescence and non-telomeric senescence (Bodnar *et al*., 1998; Collado *et al*., 2007a; Hayflick & Moorhead, 1961). In telomere dependent senescence, it has been found that cells after a specific number of cell divisions will stop replicating because of progressive shortening of telomeres. (Hayflick & Moorhead, 1961; Salama *et al*., 2014). Additionally, a wide spectrum of non-telomeric factors stimulates cellular senescence. For instance, damaged and derepressed INK4/ARF locus are possible to cause cellular senescence and, interestingly, both of them occur in aged cells (Collado *et al*., 2007b).

2.1.6 Deregulated nutrient sensing

Nutrient sensing is an essential cellular process to regulate and provide adequate and proper levels of nutrients for ideal metabolic activity (Efeyan *et al*., 2015). Cellular damages, over time, lead to deregulated nutrient sensing resulting in generation of impaired and undesirable by-products causing different types of stresses (López-Otín *et al*., 2013a). It has been shown that deregulated nutrient sensing process often gives rise to age-related diseases (Patti & Kahn, 2004). There are several pathways involved in nutrient sensing that their accurate regulation is vital for cellular

homeostasis. Among these pathways, target of rapamycin (TOR) and insulin/insulin-like growth factor signaling (IIS) pathways play important roles in ageing (Bhaumick *et al*., 1981; Kapahi *et al.*, 2004a; Theurey *et al.*, 2016; Vellai *et al.*, 2003). For instance, attenuation of IIS pathway was shown to increase the lifespan in several model organisms including *D. melanogaster* (Clancy *et al.*, 2001; Tatar *et al.*, 2001). Besides, depletion of TOR in *S. cerevisiae* was shown to increase the life span (Powers *et al*., 2006). Likewise, in *D. melanogaster*, overexpression of TOR suppressors extends the longevity (Kapahi *et al*., 2004b).

2.1.7 Mitochondrial dysfunction

Ageing is associated with a decrease in efficiency of electron transport chain which elevates the production of ROS and reduces the generation of ATP in mitochondria (Green et al., 2011). It has been evidenced that the rate of oxidative damage to lipids, proteins and DNA increases with age, supporting the role of elevated ROS in ageing process (Balaban et al., 2005). If ROS production affects the ageing process, then, it is reasonable to hypothesise that decreasing the ROS production would prolong the lifespan. However, studies shown unexpectedly opposite results (Van Remmen *et al*., 2004; Y. Zhang *et al.*, 2009). New studies suggested that ROSs have crucial roles in stress response aiming to reduce the age-related deterioration of cellular functions (Sena & Chandel, 2012). Hence, there are still conflicting evidences regarding the role of ROS in ageing process and further studies are needed to shed light on it. Additionally, several other mechanisms causing mitochondrial dysfunction including mutations and damages in mtDNA, oxidation of mitochondrial proteins, destabilization electron transport chain, and alteration of the lipid composition of

mitochondrial inner membranes can be associated with ageing, however, their role in ageing process is poorly understood (López-Otín *et al.*, 2013b).

2.1.8 Stem cell exhaustion

The stem cell exhaustion and deterioration of regenerative potential of cells and tissues is one of the most important hallmarks of aging (López-Otín *et al.*, 2013b). Stem cells have distinctive mechanisms to reduce the risk of damages over time including suppression of metabolic activity, cellular quiescence, and asymmetric cell division (Sameri *et al*., 2020). Among these mechanisms, asymmetric cell division is of interest in ageing research.

Asymmetric cell division is a conserved mechanism in which one stem cell divides into a stem cell and a differentiating cell (Morrison & Kimble, 2006). Through this mechanism, damaged proteins and mitochondria are passed on to the differentiating daughter cell, protecting the stem cell from damages and aged proteome (Charville & Rando, 2011). However, during ageing, this process declines, deteriorating the regenerative ability of stem cells (Sameri *et al*., 2020). One of the most important signalling pathways underlying asymmetric cell division is p38-MAPK signalling pathway (Morrison & Kimble, 2006). It has been shown that aged muscle stem cells (MuSC) exhibit higher levels of p38-MAPK signalling, reducing their renewal and regenerative ability (Segalés *et al*., 2016).

Apart from asymmetric cell division, there are other studies on different features of stem cells and ageing. For example, it has been found out that notch signalling, which mediates the coordination between stem cells and their niche, is up-regulated in female germline stem cells (GSCs) with age in *D. melanogaster* (Tseng

et al., 2014). In contrast, notch signalling is reduced in aged MuSCs and ectopic expression of notch can rescue the regenerative ability of MuSCs (Mu *et al*., 2015).

2.1.9 Altered intercellular communication

Cells signal each other either directly or indirectly through exchanging released substances (Mittelbrunn & Sánchez-Madrid, 2012). There are several studies highlighting the alteration of intercellular communication due to ageing. Among different types of intercellular communication, inflammation drew lots of attention in ageing research (López-Otín *et al*., 2013b). Previously, it has been indicated that inflammation is involved in a wide range of age-related disease including type 2 diabetes, cancer, and Alzheimer (Barzilai et al., 2012; Coussens & Werb, 2002; Kinney *et al.*, 2018). Besides, age-associated inflammation deteriorates the renewal ability of epidermal stem cells (Doles *et al.*, 2012). Additionally, altered intercellular communication causes a significant decline in immune system increasing the risk of cancer (Davoli & De Lange, 2011; Deeks, 2011; Senovilla *et al.*, 2012). In addition to inflammation, age-associated alteration in one tissue can induce the same changes in other tissues or organs (López-Otín *et al.*, 2013b). For instance, senescent cells induce senescence in adjacent cells through gap junctions between cells (Nelson *et al.*, 2012).

2.2 Long noncoding RNA

lncRNAs are transcripts longer than 200 bp that are not coding any proteins. Although lncRNAs are not capable of coding for functional proteins, they have many similarities with messenger RNAs (mRNAs) (Kapranov *et al.*, 2007). Same as mRNAs, lncRNAs are transcribed by RNA polymerase II. They also undergo post-transcriptional processing including capping, polyadenylation and alternative splicing (Derrien *et al.,* 2012). On the other hand, lncRNAs have some important differences from protein coding genes. For instance, lncRNAs have lower GC content and lower sequence conservation (Cabili *et al.*, 2011). In addition, lncRNAs have very specific temporal and spatial expression pattern (Ulitsky *et al.*, 2011). In spite of not being conserved, lncRNAs play vital regulatory roles in different biological processes including epigenetic regulation, transcription and post-transcription regulation (Chen & Carmichael, 2010; Clark *et al*., 2012; Wilusz *et al*., 2009).

2.2.1 Role of lncRNAs in epigenetic regulation

lncRNAs has been determined to be involved in chromatin regulation and epigenetic modifications. In fact, by recruiting chromatin-modifying factors to specific site in the genome, lncRNAs act as a scaffold to mediate the epigenetic regulation (Bhat *et al.*, 2016). Metastasis associated lung adenocarcinoma transcript 1 (MALAT1) and HOX transcript antisense RNA (HOTAIR) are two examples of transcripts that recruit chromatin-modifiers (Rinn *et al*., 2007; Zhang *et al.*, 2012). Another important lncRNA involved in epigenetic regulation is X-inactive specific transcript (XIST) that is one of the key players of X chromosome inactivation process. XIST is a giant noncoding transcript spreads along the X chromosome and mediates

the inactivation and silencing of the chromosome by recruiting chromatin-remodelling complex (Brockdorff, 2017; Brown *et al*., 1992).

2.2.2 Role of lncRNAs in transcription regulation

In addition to epigenetic regulation, lncRNAs have important roles in transcription regulation. They can modulate the activity of transcription factors by inhibiting binding of transcription factors and decoying them away from intended region in the genome. Besides, they can inhibit the activity of transcription factors by directly blocking their active site or by allosterically inhibiting them (Long *et al.*, 2017). Growth arrest-specific 5 (GAS5) and P21-associated noncoding RNA DNA damage activated (PANDA) are two examples of lncRNAs that act as molecular decoy of transcription factors (Hung *et al*., 2011; Kino *et al.*, 2010). lncRNAs also can recruit transcription factors and direct them to a certain region in the genome in order to activate gene expression (Long *et al.*, 2017). For instance, lncRNAs rhabdomyosarcoma 2- associated transcript (RMST) and EVF2 interact with Sox2 and DIX2, respectively, in order to stimulate gene expression (Cajigas *et al.*, 2015; Ng *et al.*, 2013).

2.2.3 Role of lncRNAs in post-transcription regulation

lncRNAs are also involved in post-transcriptional regulation. For instance, MALAT1 modifies the alternative splicing events through interacting with splicing factors (Tripathi *et al., 2010)*. Moreover, it has been evidenced that some lncRNAs such as Tug1 and Meg3 can inhibit the microRNAs (miRNA) activity by decoying them from their targets. Since miRNA silences its target mRNA vis base pairing of its

seed sequence, decoying them from their target mRNAs causes de-repressing of the mRNAs (Balas & Johnson, 2018; Lozano-Vidal *et al.,* 2019*)* (Figure 2.2).

.

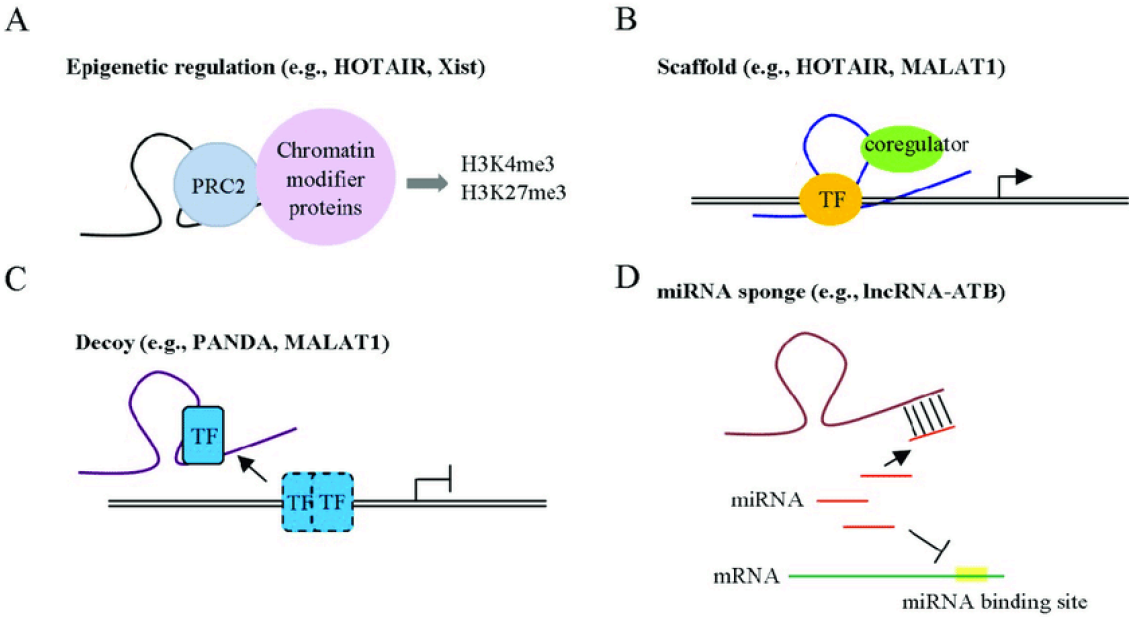

Figure 2-2. Mechanism of actions of lncRNAs. (A) lncRNAs can recruite chromatin-modifier factors such as Polycomb repressive complex 2 (PRC2) to catalyses epigenetic modification of histons and DNA. (B) lncRNAs can act as a scaffold to recruit transcription factors and direct them to a certain region in the genome in order to activate gene expression. (C) lncRNAs can inhibit binding of transcription factors and decoying them away from intended region in the genome. (D) lncRNAs can inhibit the miRNA activity by decoying them from their targets. The image was taken from Lin *et al*., 2018.

2.3 Role of lncRNAs in ageing

By regulating the gene expression in different levels, lncRNAs are increasingly recognized as the vital regulators of different biological processes in cell including proliferation, apoptosis, and differentiation (Chen et al., 2020; Rossi & Antonangeli, 2014). However, lncRNAs employ multiple mechanisms in the cells, the roles they play in ageing are mostly unknown. Here, some lncRNAs that their roles in different hallmarks of ageing have been identified will be reviewed (Figure 2.3, Table 2.1).

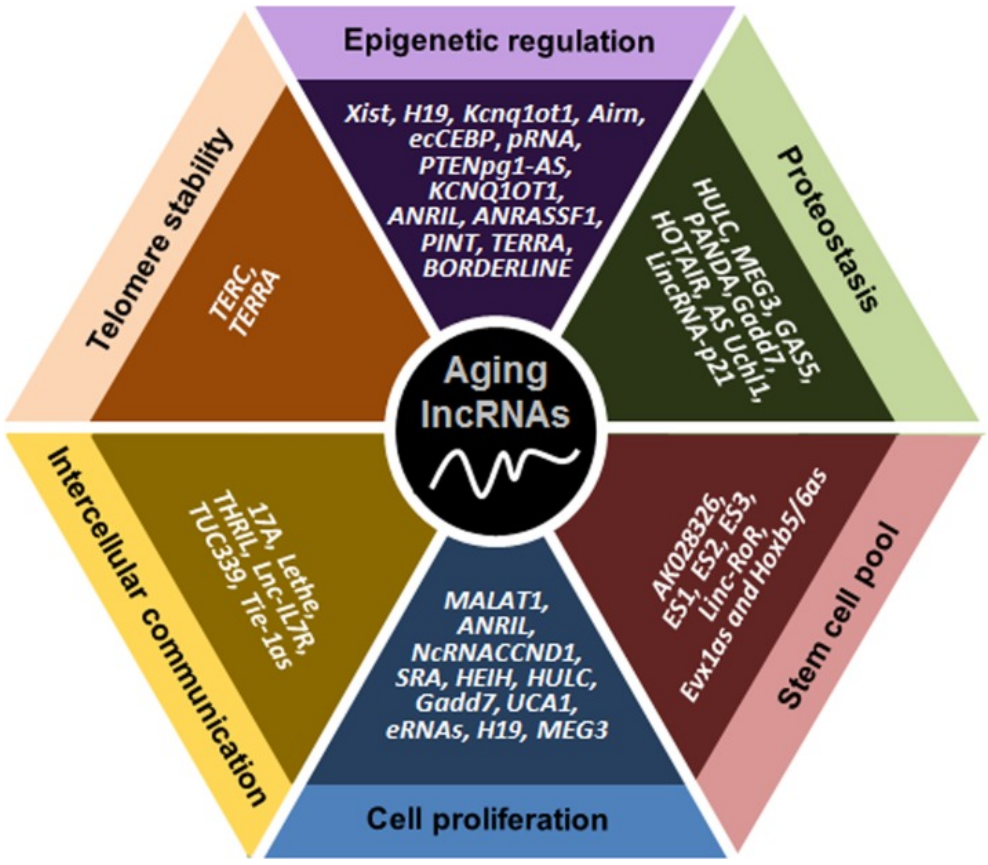

Figure 2-3. LncRNAs involved in hallmarks of ageing. Although lncRNAs linked to different hallmarks of ageing including telomere stability, epigenetic regulation, proteostasis, stem cell pool, cell prolifertion, and intercellular communication have been identified, the roles they play in ageing process are mostly unknown. The image was taken from Grammatikakis et al., 2014.

Table 2-1. Summary of age-associated lncRNAs. The table lists lncRNAs involved in ageing process (column 1), and their molecular function in ageing (column 2).

LncRNAs	Function in ageing	Reference
LncRNAs that modulate telomere length		
TERC	promotion of telomere elongation	Collins, 2008; Porro *et al.*, 2014
TERRA	Inhibition of telomere elongation	Collins, 2008; Porro *et al.*, 2014
LncRNAs associated with epigenetic alterations		
XIST	DNA methylation	Sado & Brockdorff, 2013
H19	DNA methylation	Zhou *et al*., 2015
Kcnq1ot1	DNA methylation, involved in age-related diseases like type 2 diabetes and cancer	Grammatikakis *et al*., 2014
Airn	DNA methylation, promoting cellular senescence	Grammatikakis *et al*., 2014
ecCEBP	DNA methylation, cell cycle progression	Grammatikakis *et al*., 2014
NeST	Histone modifications, recruitment of H3K4 methyltransferase, involved in inflammation.	Gomez *et al*., 2013; Grammatikakis *et al*., 2014
ANRIL	Histone modifications, stimulation of cellular senescence	Grammatikakis *et al*., 2014
LncRNAs affecting proteostasis		
HULC	Promotion of autophagy	Ying *et al*., 2013
MEG3	Suppression of autophagy	Zhang *et al*., 2013
7SL	Suppression of autophagy, protein traffiking	Zhao *et al*., 2014
GAS5	Protein traffiking	Lee *et al*., 2012; Mourtada-Maarabouni *et al.*, 2009
PANDA	Protein traffiking, DNA damage-induced senescence	Agostino *et al*., 2006

AS Uchl1	Protein synthesis, involved in cellular senescence and neurodegeneration	Carrieri *et al*., 2012; Maraganore *et al*., 2004; Ummanni *et al*., 2011
LncRNA-p21	Translation suppression, involved in carcinogenesis	Konishi *et al*., 2008; Marchand *et al.*, 2011
LncRNAs modulation stem cell function		
AK028326	Regulation of stem cell transcription factors	Grammatikakis *et al*., 2014
ES1, ES2, and ES3	Regulation of stem cell transcription factors, promotion of neuronal differentiation	Ballas *et al*., 2005; Ng *et al.*, 2012
linc-RoR	Regulation of transcription factors in pluripotent stem cell and embryonic stem cells	Loewer *et al*., 2010
LncRNAs affecting intercellular communication		
Lethe	Modulate inflammation	Rapicavoli *et al.*, 2013
THRIL	Rehulation of TNFα	TNFα
Lnc-IL7R	Regulation of inflammatory response	Grammatikakis *et al*., 2014
TUC339	extracellular signal, regulation of tcell cycle secondary tissues	Braconi *et al*., 2011; Kogure *et al*., 2013

2.3.1 lncRNAs that modulate telomere length

As it was mentioned before, DNA replication shortens the length of telomere, hence, telomerase reverse transcriptase (TERT) is essential to extend the telomere and preserve its length (López-Otín *et al*., 2013b*).* There are two important lncRNAs associated with TERT, namely, telomerase RNA component (TERC) and TERRA (Collins, 2008; Porro *et al.*, 2014). These two lncRNAs together with TERT regulate the length of telomeres. A study on lncRNA TERC showed that TERC-deficient mice

exhibit short telomere and chromosomal instability, indicating that TERC prevents the cells from premature ageing. TERC has been reported to act as a scaffold that recruits protein subunits of telomerase together with other accessory proteins facilitating the synthesis of telomere repeats (Samper *et al.*, 2001). In contrast to TERC that promotes the elongation of telomeres, lncRNA TERRA suppresses telomere lengthening. Overexpression of TERRA is associated with premature ageing. TERRA suppresses the telomere elongation because it contains numerous copies of the UUAGGG repeat causing TERRA to be a suitable ligand for TERT (Azzalin *et al*., 2007).

2.3.2 lncRNAs associated with epigenetic alterations

There are several lncRNAs contribute to different epigenetic changes including DNA methylation, histone modification, and heterochromatin formation.

Various lncRNAs that are involved in regulation of DNA methylation during ageing have been identified such as XIST which its levels decline with age (Sado & Brockdorff, 2013). Another example of a lncRNA involved in DNA methylation during ageing is H19. H19 plays a role in development and growth (Zhou *et al*., 2015). By interacting with methyl-CpG–binding domain protein 1 (MBD1) and recruiting histone lysine methyltransferases, H19 regulates the imprinting of its own gene and also insulin-like growth factor 2 (IGF2) (Monnier *et al.*, 2013). It has been evidenced that loss of imprinting of the IGF2-H19 locus occurs with aging in human prostate tissues resulting in upregulation of IGF2 and H19 (Christofori *et al*., 1994; Fu *et al*., 2008). There are several other lncRNAs that mediate the methylation and silencing of genes and have roles in ageing process such as Kcnq1ot1 (relevant to age-associated diseases like type 2 diabetes and cancer), Airn (involved in senescence through

suppressing IGF2R), and ecCEBP (influences cell cycle progression by regulating CEBP) (Grammatikakis *et al*., 2014b).

In addition to DNA methylation, lncRNAs are involved in histone modifications. For example, the lncRNA nettoie Salmonella pas Theiler's (NeST) recruits a H3K4 methyltransferase to the IFNγ locus. It has been reported that lncRNA NeST involved in inflammation during infection and that IFNγ methylation increases with age, hence, NeST could play a role in the inflammatory response and infection in aged organisms (Gomez *et al*., 2013; Grammatikakis *et al*., 2014b). There are several other lncRNAs implicated in histone modifications that are involved in ageing process such as ANRIL which imprints cellular senescence by recruiting CBX7 enhancing H3K27 methylation resulting in repression of INK4a transcription (Grammatikakis *et al*., 2014b).

lncRNAs also can contribute to ageing by changing the pattern of heterochromatin formation. As an example, lncRNA TERRA modulates the telomeric heterochromatin by binding to a wide range of telomeric proteins including HP1 and H3K9me3. HP1 was observed to be correlated to longevity (Deng *et al.*, 2009).

2.3.3 lncRNAs and proteostasis

Loss of proteostasis is one of the major hallmarks of ageing that can lead to age-related diseases such as Alzheimer, Parkinson, and Huntington. Deteriorated proteolytic systems such as autophagy and the ubiquitin-proteasome are responsible for loss of proteostasis during ageing (López-Otín *et al*., 2013). There is a wide range of lncRNAs involved in each of these proteolytic systems that will be reviewed here.

Pathways underlying autophagy are usually decline with age (López-Otín *et al*., 2013b). Undoubtably, several lncRNAs participate in these pathways. Out of all