UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculteit Geneeskunde en Levenswetenschappen
School voor Levenswetenschappen

master in de biomedische wetenschappen

Masterthesis

Study of the miR-17-92 network and potential miRNA-telomere associations in the ENVIRONAGE birth cohort

Ann-Julie Trippas

Scriptie ingediend tot het behalen van de graad van master in de biomedische wetenschappen, afstudeerrichting klinische moleculaire wetenschappen

PROMOTOR :
Prof. dr. Michelle PLUSQUIN

COPROMOTOR :
dr. Karen VRIJENS

2018
2019

De transnationale Universiteit Limburg is een uniek samenwerkingsverband van twee universiteiten in twee landen: de Universiteit Hasselt en Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculteit Geneeskunde en Levenswetenschappen

School voor Levenswetenschappen

master in de biomedische wetenschappen

Masterthesis

Study of the miR-17-92 network and potential miRNA-telomere associations in the ENVIRONAGE birth cohort

Ann-Julie Trippas

Scriptie ingediend tot het behalen van de graad van master in de biomedische wetenschappen, afstudeerrichting klinische moleculaire wetenschappen

PROMOTOR :
Prof. dr. Michelle PLUSQUIN

COPROMOTOR :
dr. Karen VRIJENS

Table of content

List of abbreviations

ENVIR*ON*AGE	The ENVIRonmental influence *ON* early AGEing
NBAS	Neonatal Behavior Assessment scale
DNA	Deoxyribonucleic acid
TL	Telomere length
PM	Particulate matter
PM_{10}	Particulate matter with an aerodynamic diameter of ≤ 10 μm
$PM_{2.5}$	Particulate matter with an aerodynamic diameter of ≤ 2.5 μm
WHO	World Health Organisation
IARC	International Agency for Research on Cancer
miRNA or miR	Micro ribonucleic acids
RNA	Ribonucleic acid
Pri-miRNA	Primary micro ribonucleic acids
Pre-miRNA	Precursor micro ribonucleic acids
RISC	RNA-induced silencing complex
mRNA	Messenger RNA
FDR	False discovery rate
GO	Gene ontology
C13orf25	Chromosome 13, open reading frame 25
RT-qPCR	real-time quantitative polymerase chain reaction
MYC	MYC proto-oncogene, BHLH transcription factor
STAT3	Signal Transducer and Activator of Transcription 3
E2F1	E2F Transcription Factor 1
TP53	Tumour Protein p53
BCL2L11 or BIM	BCL2 like 11
PTEN	Phosphatase and Tensin Homolog
CDKN1A or P21	Cyclin-Dependent Kinase Inhibitor 1A
TGF-β	Transforming Growth Factor Beta
TBFBRII	Transforming Growth Factor Beta Receptor 2
MAPK14	Mitogen-Activated Protein Kinase 14
HIF-1α	Hypoxia Inducible Factor 1 Subunit Alpha
BMI	Body mass index
NO_2	Nitrogen dioxide
BC	Black Carbon
dNTP's	Deoxyribonucleotide triphosphates
cDNA	Complementary deoxyribonucleic acid
FRET	Fluorescent resonance energy transfer
IRC	Inter-run calibrator
NTC	Non-template control
SD	Standard deviation
CI	Confidence interval
IQR	Interquartile range
SD	Standard deviation
CI	Confidence interval
ERCC1	Excision repair cross-complementation
EXOSC10	Exosome component 10
FBXO4	F-box protein 10
HISTH3	Histone cluster 3
HNRNPD	Heterogeneous nuclear ribonucleoprotein F
PPP1C	Phosphatase 1 catalytic subunit beta isozyme
SLX1A	SLX1 structure-specific endonuclease subunit beta isozyme
SMC5	Structural maintenance of chromosome 5
TINF2	TERF1-interacting nuclear factor 2
WRAP53	WD repeat containing antisense to TP53
TFIP11	Tuftelin interacting protein 11
TOX4	TOX high mobility group box family member 4
HUVEC	Human umbilical vein endothelial cells
NSCLC	Non-small cell lung cancer
PE	Preeclampsia

Acknowledgements

The last eight months when doing research and writing this thesis, I realized that my last moments as a student passed very quickly. However, my achievements could not be accomplished without the help of a number of people. First of all, I would like to thank my promotor Prof. dr. Michelle Plusquin as well as Prof. dr. Tim Nawrot for giving me the opportunity to complete my senior internship in their research group at the Centre for Environmental Sciences. It was a pleasure to do research at the CMK and to broaden my knowledge in epidemiological research.

Special thanks go out to my co-promotor and supervisor dr. Karen Vrijens. I would like to thank you for your daily guidance, the tips and tricks and all your time and devotion you spend to me. Regardless of your busy agenda, you always made time to answer my questions and explain me things even if I needed some extra information. Thanks to you, I learned to push my limits and chasing my dreams. It was a great pleasure to meet you and working with you.

Next, I would like to thank some other people at the CMK. When I had some questions, there was always someone to help me. Here, I would like to thank, especially, Verena Iven. You were always there for me when I needed you. I would like to thank you for your help and advice in the lab, but also your aid with writing this thesis. I would also like to thank Stephanie Vandionant for the help with the problems during my online solicitation.

My fellow students deserve also special thanks, in particular, Laurien Geebelen. It was a pleasure to have you around. We spend a lot of time together in the office with many fun moments, but also some less happy moments. We helped each other, if possible, to get through every situation and we always gave our honest opinion, even with the layout of our figures and tables. Thanks a lot for the joyful spirit, your enthusiasm and the optimism, which made me smile every day. It is a time to never forget.

Last but not least, I would like to thank my family and boyfriend for the interest in my internship, for listening when I had a hard time and for supporting me throughout my study period. Here, I would like to thank my parents in particular, for always supporting me, for giving me advice and for giving me the opportunity to study. But most of all for listening and consoling me, when I was freaking about deadlines or busy days. Without them, my achievement would not have been possible.

Abstract

Introduction: Already early in life, people are exposed to disturbances and environmental hazards, transported via the placenta to the foetus. *In utero* air pollution exposure can result in adverse health effects and disease risk later in life, including ageing and cancer risk. Epigenetic mechanisms, such as miRNAs, are known to have a key role in ageing, via telomere controlling, and potentially function in diseases induced by air pollution exposure. Telomere length (TL) at birth has already been associated with lifespan in zebra-finches as it determines the capacity of the cell to cope with inflammatory and oxidative processes throughout lifetime. Therefore, TL is associated with age-related diseases, such as cancer. In colorectal cancer, several miRNAs were associated with TL, via TERT regulation. Therefore, it is hypothesised that placental candidate miRNAs are differentially expressed in association with TL at birth.
Meanwhile, the oncogenic miR-17/92 cluster was shown to be related to prenatal air pollution exposure in cord blood, which indicates the relation between miRNAs and early life environmental exposure. The cluster is involved in normal development and carcinogenesis by targeting genes in important biological processes, including cell proliferation (MAPK14 and TGFBRII), cell death (BCL2LL1) and even hypoxia (HIF-1α). However, up-stream targets are also involved in critical cell functions, such as cell cycle regulation (MYC and E2F1) and cell proliferation (STAT3). It is hypothesised that the network around the oncogenic cluster is differentially expressed in association with prenatal air pollution exposure.
Here, we investigate potential alterations (1) in the expression of the network, including regulators and targets of the cluster, in association with prenatal air pollution exposure and (2) in the miRNA expression of all known human miRNAs in the placenta in association with telomere length at birth.

Methods: (1) To determine the association between the miR-17/92 cluster network and prenatal air pollution exposure, the relative mRNA expression of E2F1, C-MYC, STAT3, TP53, PTEN, BCL2L11, CDKN1A, MAPK14, HIF-1α, TGFBRII was measured in 511 cord blood samples from the ENVIR*ON*AGE birth cohort by RT-qPCR. The exposure to particulate matter with a diameter less than 2.5µm ($PM_{2.5}$), NO_2 and black carbon (BC) for the different trimesters and the entire pregnancy is estimated, using a spatiotemporal model. The associations between the relative mRNA expression and the exposure windows during pregnancy for $PM_{2.5}$, NO_2 and BC were analysed using linear regression models adjusted for potential confounders, such as maternal age, pre-gestational BMI, parity, educational status, smoking status, gestational age, date of delivery, newborn's sex, ethnicity and apparent temperature for each specific time window of pregnancy. (2) To determine the association of placental miRNA expression and telomere length at birth, the relative miRNA expression of all known human miRNAs was measured in the placenta of 60 individuals from the ENVIR*ON*AGE birth cohort, using microRNAarray analysis. Based on *in silico* work, using gene ontology terms including the word "telomere", a miRNA selection of miRarray data was made for further analysis. Sex-specific associations were analysed for the relative miRNA expression and the placental TL, using mixed linear model adjusted for potential confounders, including date of hybridisation, father's age, gestational age, parity, season of delivery, date of delivery and maternal educational status.

Results: (1) The miR-17/92 cluster network was found to be differentially expressed in cord blood in association with prenatal air pollution exposure. For each 5 µg/m^3 increase in $PM_{2.5}$ for the entire pregnancy, the relative E2F1 expression declines with 25.51% (95% CI: -37.10% to -11.79% p=0.0007) and STAT3 expression rises with 11.46% (95% CI: 0.08% to 24.13%, p=0.05). MYC was significantly associated in the first trimester of pregnancy and reduces 7.38% (95% CI: -12.00% to -2.51%, p=0.003) for each 0.5 µg/m^3 increase in BC exposure. Downstream targets BCL2L11, HIF-1α, MAPK14 and TGFBRII were inversely associated with *in utero* $PM_{2.5}$ exposure. (2) TL at birth is significantly associated with 2 miRNAs in the placenta. In newborn boys, Placental expression of hsa-miR-584-5p (β=3,38, SD=±0.95, p=0.002) and hsa-miR-4291 (β=1.51, SD=±0.68, p=0.04) were found to be positively associated with TL.

Conclusion: Prenatal air pollution exposure is associated with activators and targets of the miR-17/92 cluster in cord blood. In addition, at birth TL is associated with miR-584 and miR-4291 in the placenta of newborn boys. Both miRNAs are involved in the development of several disorders. The function of miR-4291 is not known yet. Because of its high number, not much is known about this miRNA. However, miR-584 is known to be involved in tumorigenesis, in which cell proliferation is important. During cell proliferation, TL will shorten gradually. This can result in increased disease risk, e.g. cancer, which has already been shown to be associated with shorter TL. The identified associations can provide insight how early life air pollution exposure might induce predisposition for disease development via the miR-17/92 cluster and how miRNAs could be involved in the TL variability at birth, which might be involved in predispositions of age-related diseases.

Samenvatting

Introductie: Al vroeg in het leven worden mensen blootgesteld aan omgevingsinvloeden en verstoringen die via de placenta tot bij de foetus komen. *In utero* blootstelling aan fijn stof kan negatieve gezondheidseffecten en een hoger risico op ziekten met zich meebrengen, en hierdoor bijdragen aan veroudering en kanker risico. Epigenetica en zijn mechanismen, waaronder miRNAs, zijn gekend om hun sleutelrol in veroudering, via de regulatie van telomeren alsook hun bijdragen aan ziekten, ten gevolge van luchtvervuiling. De lengte van de telomeren (TL) bij de geboorte bij zebravinken is geassocieerd met de levensverwachting, vermits ze de capaciteit van de cel bepaalt om inflammatoire en oxidatieve processen doorheen het leven te incasseren. Hierdoor is TL gelinkt met verouderingsziekten zoals kanker. In darmkanker is de TL geassocieerd, via TERT regulatie, met verschillende miRNAs. Vandaar veronderstelt men dat er kandidaat miRNAs zijn die anders tot expressie komen in de placenta in associatie met TL.
De oncogene cluster miR-17/92 daarentegen heeft in navelstrengbloed een verband getoond met prenatale blootstelling aan luchtvervuiling, wat de relatie tussen miRNAs en omgevingsblootstelling vroeg in het leven aanhaalt. De cluster is betrokken in de ontwikkeling alsook in tumorontwikkeling, dit door in te werken op genen betrokken bij belangrijke biologische processen zoals celproliferatie (MAPK14 en TGFBRII), celdood (BCL2L11) en hypoxie (HIF-1α). Ook activators van de cluster hebben een functie in kritische celfuncties zoals celcyclusregulatie (MYC en E2F1) alsook in celproliferatie (STAT3). Men vermoedt dat het oncogene netwerk van de cluster verschillend tot uiting wordt gebracht in associatie met *in utero* blootstelling aan luchtvervuiling. In deze studie onderzoeken wij de mogelijke veranderingen: (1) in de expressie van het netwerk, dat zowel regulatoren als targets bevat van de cluster, in associatie met prenatale luchtvervuilingsblootstelling en (2) in de miRNA expressie van alle gekende menselijke miRNAs in de placenta in relatie met telomeerlengte bij de geboorte.

Methoden: (1) Om de associatie na te gaan tussen het netwerk van de oncogene cluster en prenatale luchtvervuilingsblootstelling, werd de relatieve mRNA expressie van E2F1, C-MYC, STAT3, TP53, PTEN, BCL2L11, CDKN1A, MAPK14, HIF-1α en TGFBRII gemeten in 511 navelstrengbloedstalen van het ENVIRONAGE geboortecohort, door middel van RT-PCR. De blootstelling aan fijn stof kleiner dan 2.5 µm ($PM_{2.5}$), NO_2 en koolstofdeeltjes (BC) werd geschat voor de verschillende trimesters alsook de volledige periode van de zwangerschap op basis van een spatiotemporeel model. De associaties tussen de relatieve mRNA-expressie en de verschillende tijdsvensters van de zwangerschap voor $PM_{2.5}$, NO_2 en BC werden nagegaan met behulp van lineaire regressieanalyse. Deze werden gecorrigeerd voor mogelijk bijdragende variabelen, zoals leeftijd van de moeder, BMI voor de zwangerschap, aantal zwangerschappen, opleidingsniveau, rookgedrag, bevallingsdatum, geslacht van de baby, zwangerschapsduur, etniciteit en temperatuur voor elk tijdsvenster van de zwangerschap. (2) Om de associatie tussen de miRNA-expressie en de TL bij geboorte te onderzoeken, werd de relatieve expressie van alle gekende miRNAs gemeten in placentastalen van 60 individuen uit het ENVIRONAGE geboortecohort d.m.v microRNAarray-analyse. Bij *in silico* werk gebruik makend van Gene Ontology-termen met het woord "telomeren", werd er een selectie miRNAs van de miRarray dataset voor verdere analyse gemaakt. Geslacht-specifieke associaties, tussen de relatieve miRNA-expressie en de TL in de placenta, werden geanalyseerd d.m.v. mixed-lineaire modellen gecorrigeerd voor mogelijke variabelen, zoals hybridisatiedatum, leeftijd van vader en moeder, zwangerschapsduur, datum en seizoen van de bevalling en opleidingsniveau van de moeder.

Resultaten: (1) et netwerk van de miR-17/92-cluster komt anders tot expressie in associatie met prenatale blootstelling aan luchtvervuiling en dit in navelstrengbloed. Bij elke stijging in $PM_{2.5}$-blootstelling van 5 µg/m^3 tijdens de volledige zwangerschapsperiode zal de relatieve expressie van E2F1 afnemen met 25.52% (95% CI: -37.10% tot -11.79% p=0.0007). De expressie van STAT3 daarentegen stijgt met 11.46% (95% CI: 0.08% tot 24.13%, p=0.05). In het eerste trimester van de zwangerschap zal de relatieve expressie van MYC dalen met 7.38% (95% CI: -12.00% tot -2.51%, p=0.003) bij elke toename in BC-blootstelling van 0.5 µg/m^3. De targets van de cluster, BCL2L11, HIF-1α, MAPK14 en TGFBRII, zijn omgekeerd evenredig met *in utero* blootstelling aan $PM_{2.5}$. (2) Bij de geboorte is de TL significant geassocieerd met twee miRNAs in de placenta. Bij pasgeboren jongens is de miRNA-expressie van hsa-miR-584-5p (β=3,38, SD=±0.95, p=0.002) en hsa-miR-4291 (β=1.51, SD=±0.68, p=0.04) in de placenta positief gerelateerd met de TL.

Conclusie: Prenatale blootstelling aan luchtvervuiling is geassocieerd met activators alsook targets van de miR-17/92-cluster in navelstrengbloed. Bijkomend zien we een associatie tussen TL bij de geboorte en de expressie van miR-585 en miR-4291 in de placenta van pasgeboren jongens. Deze twee miRNAs zijn betrokken in de ontwikkeling van verschillende aandoeningen, zoals tumorgenese. Celproliferatie is belangrijk in tumorgenese. Tijdens dit proces neemt de lengte van de telomeren geleidelijk af. Dit kan een hoger risico voor bepaalde ziekten, bv. kanker, tot gevolg hebben. Hierbij is de associatie met kortere telomeren al aangetoond. De aangetoonde verbanden geven ons inzicht in hoe vroegtijdige blootstelling aan luchtvervuiling kan leiden tot een aanleg voor bepaalde ziekten, via het miR-17/92-cluster-netwerk en hoe miRNAs betrokken zijn in de verschillen in TL bij de geboorte hetgeen kan leiden tot verouderingsziekten later in het leven.

1 Introduction

1.1 The ENVIRonmental influence ON early AGEing birth cohort

Ageing is an unavoidable and universal process which does not occur the same for everyone. This process can be influenced by many environmental stressors, for example, radiation, tobacco smoke and air pollution (4). From early life onwards, many environmental hazards will influence this process via metabolic and physiological adaptations (4-6). D.J. Barker stated that during *in utero* life, early life perturbations, such as environmental exposures, could be transported via the placenta to the foetus. These predispositions will affect adult health and the child's risk for disease later in life (5, 6). The ENVIRonmental influence *ON* early AGEing (ENVIR*ON*AGE), a population-based longitudinal birth cohort situated in Belgium, is designed to study these interactions between environmental factors and the process of ageing (4).

Since February 2010, mother-newborn pairs have been recruited within the ENVIRONAGE birth cohort, when they arrive at the East-Limburg Hospital in Genk. The children will be followed at the age of four to six years old during a follow-up meeting. Lifestyle, medical, clinical and neurological data were collected of both mother and child. The daily air pollutant exposure levels during pregnancy will be modelled after delivery, based on the residential home address. Moreover, all medical information, such as anthropometry, foetal ultrasound records and lifestyle factors will be compiled. At time of birth, several biological specimens were taken, such as placenta, cord blood, meconium, maternal blood and urine. In addition, neurocognitive development assessment will be determined with the Neonatal Behaviour Assessment Scale (NBAS) and three days after birth, blood pressure will be measured of both mother and child. The study design is more extensively described in Janssen *et al* (4).

Because of this unique design, both subclinical and molecular parameters can be combined with early life environmental exposures. Moreover, this design will also cover some other objectives. The first objective assessed in the ENVIRONAGE birth cohort is the investigation of the association between environmental (e.g. air pollution exposures) or lifestyle factors (e.g. maternal body mass index) and molecular targets of ageing, including telomere length and mitochondrial function. A second objective will study the link between early life environmental or lifestyle factors and clinical outcomes in childhood including cardiovascular function and neurobehavioral performance. The last objective covers the interaction between molecular targets of ageing or molecular signatures in early life and clinical outcomes in childhood (4).

1.2 Telomeres

One of the "hallmarks" of ageing, is the progressive shortening of telomeres (7). Telomeres are found at the end of the chromosomes and are noncoding nucleoprotein structures consisting of tandem repeats of the sequence -TTAGGG- and function in the maintenance of genomic stability and the protection of degradation and/or end to end fusion of the chromosomes (8). With every cell division, telomeres will shorten gradually until a critical length, known as the Hayflick limit. At this point, the cell is not able to divide due to the inability of deoxyribonucleic acid (DNA) polymerase replicate the lagging strand completely, which results in senescence and apoptosis (9).

An animal-based study, performed in zebra-finches has shown that the telomere length (TL) at birth was associated with the lifespan of the animals, which is also found to be a prediction of adult TL. The variety of TL is partially due to genetic factors (10, 11). However, also environmental factors, such as air pollution exposure, can explain the variability (12). The telomere length at birth determines the capacity of the cell to cope with inflammatory or oxidative processes during lifetime, which means that longer telomeres have a better cellular function against these conditions. Therefore, TL is related to ageing and has been shown to be associated with age-related diseases (10).

1.3 Air pollution

An important environmental factor, studied within the ENVIRONAGE birth cohort, is air pollution. Air pollution is defined as each contamination with chemical, biological or physical substances that alter the native features of the atmosphere, which can threaten human health (13, 14). Air pollution can be divided into four subgroups, namely gaseous pollutants, persistent organic compounds, heavy metals and particulate matter (PM) (14). The latter is the major component of air pollution and carries the most important risk for human health. This category consists of a mixture of particles, such as sulphate, nitrates, ammonia, sodium chloride, black carbon, mineral dust, water and breathing air.

The different particles of PM vary both in composition and size. They have arisen from natural or anthropogenic activities like combustion engines, solid fuel, combustion for energy production and other industrial activities. The particles are divided based on size. There are different types of PM. Aerodynamic particles with a diameter smaller than ten or 2.5 µm (respectively PM_{10} and $PM_{2.5}$) are divided into different categories. The first class, ultrafine particles, are particles with a diameter smaller than 0.1 µm. When the diameter is smaller than 1µm, they are classified as fine particles. The last category is coarse particles and these particles have a diameter larger than 1µm. The size of the particles determines the effects on the respiratory tract. Smaller particles are more inhalable, which makes them capable of penetrating into the lungs more deeply, and they can enter the bloodstream (15, 16). In Belgium, the annual average $PM_{2.5}$ concentrations in 2017 were situated between 0 and 20 $\mu g/m^3$, which is under the European limit of 25 $\mu g/m^3$ (17).

Fine-particulate air pollution was found to have adverse effects causing health problems in the human population (4, 18). A well-documented extreme air pollution exposure episode in history is situated in London in 1952. During this period, a thick blanket of air pollution covered the city due to the weather conditions. The toxic air caused an increase in respiratory and cardiovascular disorders. They estimate that this phenomenon occasion approximately 4000-12000 deaths (19).

Nowadays, fine-particulate air pollution is considered to be the main reason for human premature death worldwide (19). The World Health Organisation (WHO) estimated that air pollution causes seven million premature deaths each year, which were mostly due to cardiovascular diseases, stroke, chronic obstructive pulmonary disease, acute respiratory infections and cancer(15). The International Agency for Research on Cancer (IARC) classified air pollution as carcinogenic (class 1) for human beings (20). This classification is built on evidence from carcinogenicity studies in humans, animals and mechanistic in vitro studies (20). O Raaschou-Nielsen *et al*. performed a prospective analysis on data from 17 cohort studies in nine countries received from the European Study of Cohorts for Air Pollution Effects. This research showed that residential particulate air pollution was associated with the risk of lung cancer (21).

1.4 Epigenetics and micro-ribonucleic acids

Epigenetics is known to have a key role in the regulation of ageing, via telomere controlling. However, epigenetics can also contribute to disease pathogenesis, and potentially place a role in diseases induced by air pollution exposure (7, 22). Epigenetics is defined as the heritable changes in the gene expression that cannot be clarified by alterations in the DNA sequence (23-25). There are different types of epigenetic marks, such as histone modifications, DNA methylation and non-coding micro ribonucleic acids (miRNAs) (23, 26). Epigenetic modulations are involved in the regulation of the gene expression, which can explain partially the phenotypic differences of organisms with an identical DNA content (26). Epigenetic modifications are unstable and subject to environmental exposures, such as PM. Via targeting the gene expression, these marks will affect the process of ageing and the pathogenesis of diseases, for example, cancer development (7, 23, 25, 26).

Epigenetic mechanisms have an influence on the gene expression, as already explained. miRNAs will have a regulatory function at a post-transcriptional level. miRNAs are a type of non-coding, single-stranded ribonucleic acids (RNAs) with a length of approximately 20 to 24 nucleotides. (23, 25, 27, 28). The post-transcriptional regulation of miRNAs affects several cellular processes, such as proliferation, apoptosis and tissue differentiation (25). Within the human genome, one to four per cent is estimated to encode for miRNAs and these miRNAs manage the expression of 30% of all human genes (25, 29).

The regulation of gene expression via miRNAs starts with the biogenesis (Figure 1) within the nucleus. The miRNAs will be transcribed by RNA polymerase II. This transcript is called the primary transcript (pri-miRNA) and is capped and deadenylated. Once formed, the pri-miRNA will be processed, by Drosha RNase III endonuclease in shorter fragments of approximately 70 nucleotides. Moreover, these precursor miRNAs (pre-miRNA) are characterised by a stem-loop (1, 30, 31). Then, the pre-miRNA is exported to the cytoplasm, where it immediately will be processed by Dicer. Dicer, a Ribonuclease and RNase II endonuclease, will cleave the pre-miRNA into a double-stranded RNA molecule, named miRNA:miRNA*, with a length of approximately 20-22 nucleotides. One strand of this complex will be the mature miRNA and will be loaded to the RNA-induced silencing complex (RISC) (1, 30, 31). Finally, the mature miRNA can carry out its regulatory function. It will guide RISC to its target messenger RNA (mRNA) molecules and will subsequently interfere with its gene expression. This will result in the degradation or translational repression of the target, which changes the expression of the target molecules (1, 30, 31).

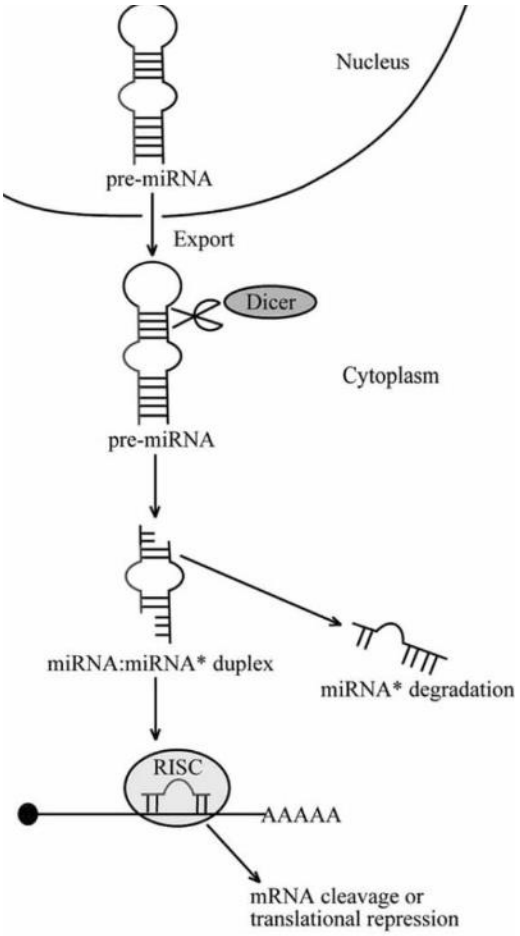

Figure 1. The biogenesis of miRNAs: within the nucleus, miRNAs will be transcribed as pri-miRNAs and directly processed in pre-miRNAs by the Drosha enzyme. Next, it will be exported to the nucleus and cleaved by Dicer in a double-stranded miRNA complex. One strand will become the mature miRNA and will be loaded into the RISC complex. Next, the miRNA will guide RISC and bind to its mRNA target. Finally, this will lead to target degradation or translational repression (1).

Because of its function in cellular processes like proliferation and apoptosis, miRNAs need strict regulation. Throughout the years, more evidence is obtained that environmental pollution can affect gene expression. These alterations can be induced by changes in the miRNA expression patterns (32). Different environmental exposures have already been shown to alter miRNA expression. Examples of these exposures are cigarette smoking and metals, like arsenic and cadmium. However, also particulate matter has been shown to be related to the expression patterns of miRNAs (32). To our knowledge, the pioneer within this topic, studied the effect of air pollution exposure and more specifically diesel exhaust. They found dramatic changes in the expression of 62.9% of the detected miRNAs in differentiated human bronchial epithelial cells (33). However, also PM has been shown to change miRNA expression patterns. Bollati *et al.* found that PM exposure during work increases the expression of miR-222 and miR-21 in steel factory workers (34). Not only adults will be affected by PM exposure, but also prenatal exposure has been found to induce altered miRNA patterns both in cord blood and placental tissue. This can have adverse effects on foetal development and health outcomes later in life. Within the ENVIRONAGE birth cohort, an association was found between prenatal PM exposure and the expression of three specific miRNAs, in particular, miR-21, miR-146a and miR-222 (29).

miRNAs play a role in almost every critical cellular process. They balance the expression of regulatory gene networks, by which they define cell fate (35, 36). When this balance is dysregulated, they can contribute to ageing by progressive telomere shortening. Like previously explained, TL is partly due to genetic factors. However, environmental factors could also explain telomere shortening. For example, long-term PM exposure has already been shown to be associated with shorter TL. Moreover, also miRNAs have already been shown to be associated with PM exposure. Therefore, it is important to determine possible mechanisms by which miRNAs could be involved in telomere attrition. Within the ENVIR*ON*AGE birth cohort, they already have been shown the association of miRNAs and TL. They found higher expression patterns of the miRNAs miR-34a, miR-146a, miR-210 and miR-222 in the placenta in association with placental telomere length at birth in new-born girls. Moreover, other studies strengthen the hypothesis that microRNAs contribute to ageing. ElSharawy *et al*. showed a differential significant expression of 80 miRNAs in blood of a population of 90 years old people relative to middle-aged people, such as miR-20a, miR-21 and miR-222 (7, 37, 38).

A possible mechanism by which miRNAs might contribute to telomere shortening is by interfering with the activity of telomerase and telomerase reverse transcriptase (TERT). TERT is an enzyme, which advances re-alignment of the template to synthesise numerous DNA repeats continuously. TERT is a component of telomerase and together with the RNA element TERC, it will attach to the 3' single-stranded telomeric end. This will contribute to the maintenance of the TL and the protection of the chromosome ends. Several studies have already shown that miRNAs are involved in telomere length via the regulation of TERT and telomerase activity. In a population-based study of individuals with colorectal cancer, Slattery *et al.* showed that 34 miRNAs were differentially expressed in association with TL. Moreover, 75 miRNAs were found to be associated with TERT in the same population. miR-615-5p, for example, has already been shown to be involved in TERT regulation. In cancer cells, increased expression of miR-615-5p, miR-296-5p or miR-512-5p is related with repression of TERT. Moreover, TERT inhibition is also associated with shorter telomeres. Another example is miR-380-5p. Increased expression of this miRNA in mouse fibroblast STO cell lines induced inhibition of the telomerase activity and this was associated with a little increase of the TL with approximately 1 Kb. These studies suggest the influence of miRNAs on telomere length via telomerase and TERT regulation (39, 40).

Based on these observations, it is hypothesised that other miRNAs might also be involved in telomere shortening. Recently, within the ENVIR*ON*AGE birth cohort study, all known human miRNAs were measured in the placenta of 60 individuals, using microRNA microarray analysis. These individuals were selected based on telomere length and were divided into two groups. Both groups included 30 subjects. However, the TL of both groups was significantly different. An association with telomere length has not been found after false discovery rate (FDR) correction, which is a method to visualize type I errors when testing the null hypothesis using multiple comparisons.

In the current research, we will perform some *in silico* work, to increase the power of the previously performed study. In the prior study, a small sample size caused a low power. Power is the ability to determine significant association. Therefore, it is hypothesised that placental candidate miRNAs will be differentially expressed in association with TL at birth. To test the hypothesis in the current research, a first objective is to select a subset of miRNAs known to be involved in telomeres based on Gene Ontology (GO). By selecting miRNAs for further analysis, the power will be increased as well as the chance to detect significant associations. Therefore, the second objective is to determine the association between the placental miRNA expression and TL at time of birth.

However, deregulation of miRNA expression can also contribute to the pathogenesis of diseases (35, 36). This association between disease phenotypes and miRNAs is already found in approximately 70 different diseases, including cancer (25, 28, 31).

Cancer is the second most lethal disease in the modern world. Every year, approximately 9.6 million people die from cancer (41). Cancer is an overall term for all diseases characterised by the uncontrolled growth of abnormal cells, which have the malignant capacity to invade in neighbouring tissues or spread to other organs in the body (41). The major cause of cancer is a change in the genes that play an important role in cellular function, such as cell growth and cell division. These differences can be genetically inherited, due to DNA damage induced by carcinogenic substances, such as the different fractions of air pollution, or due to mutations during cell division (42).

miRNAs can contribute to tumorigenesis by modifying different oncogenic or tumour suppressor genes. Thereby, they are implied in malignant conversion by modulating mechanisms such as cell cycle control, replying to DNA damage, (stem cell) differentiation, metastasis, etc. (31, 35, 36). Currently, a substantial amount of studies showed the involvement of abnormal expression of miRNAs and the development of malignancies (31, 35). Constinean and colleagues showed the development of B-cell leukaemia and high-grade lymphomas in a transgenic mice model, which overexpressed miR-155 in their B-cells. Nowadays, it is known that miR-155 is increased in several malignancies, such as breast cancer, pancreatic cancer and some types of lymphomas (35, 36, 43). Furthermore, many other miRNAs have been found to be related to human tumours, like the miR15/16 family. In human B-cell chronic lymphocytic leukaemia, a deletion might occur at chromosomal locus 13q14. A deletion at this locus is characterised by downregulation of both miR-15 and miR-16. These data suggest that miRNAs can be a major contributor to tumorigenesis. This influence can either be oncogenic or tumour suppressive (31, 35, 36).

1.5 The miR-17/92 cluster

One-third of the miRNA-coding genes are clustered in one single pri-miRNA transcript. The mature miRNAs coded by this cluster, can function independently (44). The first discovered cluster with oncogenic potential, also named oncomir-1, is the miR-17/92 cluster. The cluster encodes for six mature miRNAs; miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1 and miR-92a-1, which can be divided in three families (miR-17, miR-19 and miR-92 family) (3, 44, 45). The cluster counts 800 base-pairs and is located at chromosome 13 within, open reading frame 25 (C13orf25) (3, 29, 44, 45).

The cluster is highly conserved in vertebrates, which indicates its importance in the regulation and functioning of cells. The miR-17/92 cluster is involved in "normal" development, as it functions in different critical cellular processes, such as cell cycle regulation, differentiation, proliferation, angiogenesis and apoptosis (3, 44, 45). The cluster's function in the regulation of cellular life and death elucidates that potential dysregulation might contribute to disease pathogenesis. The other name of the cluster, oncomir-1, suggests its contribution in tumorigenesis and this is confirmed by both the oncogenic or tumour suppressive function of the members within this cluster (44, 45). Overexpression of the cluster is found in many hematopoietic malignancies and solid tumours. For instance, He *et al.* showed that the overexpression of the cluster, within an *in vivo* mouse B-cell lymphoma model, was associated with an accelerated onset and progression of B-lymphomagenesis (3, 44, 45).

However, in some disease phenotypes, a loss of function of the miR-17/92 cluster is seen. For example, during ageing, several members of the cluster (miR-17, miR-19b and miR-20a) have already been shown to be downregulated (2, 3, 44, 45). Moreover, this can be observed in age-related diseases, such as Alzheimer's disease (2, 3, 29). However, also for cancer cells, a loss of function of these miRNAs can be advantageous. Hossain *et al*., for example, observed a reduction in the proliferation of cancer cells after introduction of miR-17 in a breast cancer cell line. This implicates a deletion of the miR-17/92 cluster in this type of cancer, which can also be seen in other tumour types such as acute myeloid leukaemia, ovarian cancer, melanoma and retinoblastoma (2, 3, 44, 45).

Because of the miR-17/92 cluster's association with different disease phenotypes, strict regulation of its expression is necessary. However, as already explained above, miRNA expression can be influenced by environmental exposures. This is also be observed for the miR-17/92 cluster. Recently, within the ENVIR*ON*AGE birth cohort, the alterations in the expression of the cluster members were observed in cord blood in association with the environmental pollutant particulate matter. Within a population of 408 mother-newborn pairs, the relative expression of miR-17, miR-20a and miR-92a were measured in cord blood, using quantitative Real-Time polymerase chain reaction (qRT-PCR). Multiple linear regression was performed to associate the miRNA expression with $PM_{2.5}$ exposure during the different time windows of pregnancy. They observed an inverse relation in the expression of the miR-17/92 cluster and prenatal particulate matter exposure (data not published).

Further research is necessary to determine the long-lasting consequences of these epigenetic alterations. Because of the cluster's association with tumorigenesis, it is important to determine how the network (Figure 2) around the cluster is influenced (2, 3, 29, 44, 45). The question "Is there an alteration in the mRNA expression of both oncogenic or tumour suppressor members of the pathway of miR-17/92 associated with perinatal air pollution exposure?" arose and therefore the current research has the goal to investigate a network of both up- and downstream targets of the miR-17/92 cluster to determine potential alterations in the expression of these targets in association with prenatal air pollution exposure. Because of the knowledge that the miR-17/92 cluster is downregulated in some cancer types and the preliminary results of its association with perinatal air pollution exposure, it is hypothesised that the mRNA expression of both up- and downstream targets, known to be involved in cancer initiation, of three specific miR-17/92 cluster members (miR-17, miR-20a and miR-92a) are associated with prenatal exposure to air pollution.

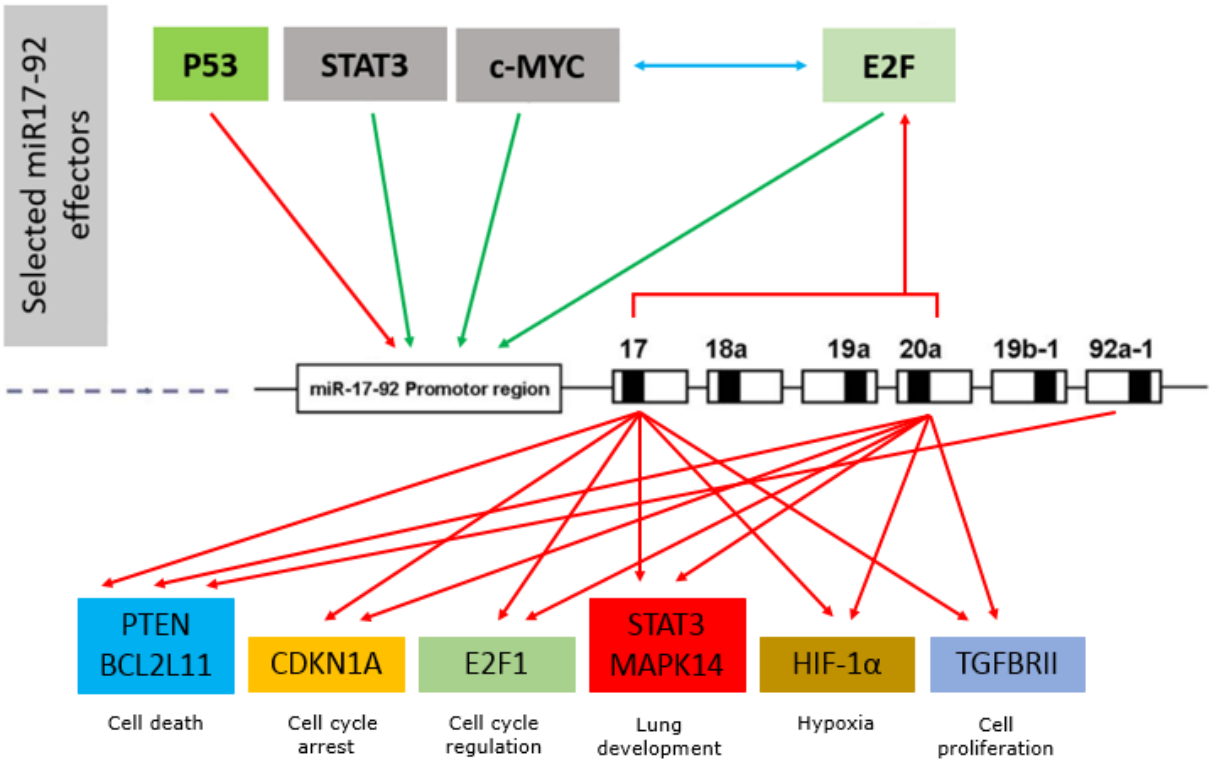

Figure 2. Overview of the regulation of transcription and target mRNAs of the miR-17/92 cluster. The red arrows indicate inhibition, the green arrows indicate upregulation and the blue arrow indicate a feedback loop. Adapted from (2, 3)

In the current research, Both up- and downstream targets will be investigated for alterations in the expression in association with air pollution exposure during pregnancy. The miR-17/92 cluster is regulated (Figure 2) by different transcription factors, such as MYC proto-oncogene, BHLH transcription factor (MYC), Signal Transducer and Activator of Transcription 3 (STAT3), E2F Transcription Factor 1 (E2F1) and Tumour Protein p53 (TP53). Effector molecules, MYC, STAT3 and E2F1 activate the cluster and induce transcription of the individual members by binding to the promotor region. On the other hand, TP53 will inhibit the cluster and its transcription (1-3, 30). Within this regulatory circuit, transcriptional activation of E2F1 will also be induced by MYC and the other way around. Moreover, the transcription factor E2F1 will be inhibited by two members of the cluster, namely miR-17-5p and miR-20a. In this way, an auto-regulatory feedback-loop controls the activation of E2F1 via MYC. These regulatory targets function in cellular processes, such as cell proliferation, cell growth, apoptosis and cell differentiation and are already shown to be involved in tumorigenesis (1-3, 30, 46, 47). In Table 1, a detailed description of the function of the different regulators of the cluster is given.

Table 1. Overview of regulating factors of the miR-17-92 cluster. These regulators either activate or repress the transcription of the miR-17/92 cluster (1-3, 30, 47).

Name	Type of gene	General function	Function in the cluster
MYC Proto-oncogene, BHLH Transcription Factor (MYC)	Proto-oncogene	- Cell cycle control - Apoptosis - Cellular transformation	- Activation of the cluster - Induction of transcription
Signal Transducer and Activator of Transcription 3 (STAT3)	Transcription factor	- Cell proliferation - Cell growth - DNA replication	- Activation of the cluster - Induction of transcription
E2F Transcription Factor 1 (E2F1)	Transcription factor	- Cell cycle control - DNA replication	- Activation of the cluster - Induction of transcription
Tumour-protein P53 (TP53)	Tumour suppressor gene	- Cell arrest control - Cell senescence control - Apoptosis - DNA repair mechanisms - Metabolic control	- Inhibition of the cluster - Inhibition of transcription

The cluster is known to regulate the expression of minimum 30 mRNA targets (2). The targets, chosen for the current research, regulated by three members of the cluster, specifically miR-17, miR-20a and miR-92a, are described in more detail in Table 2. These targets are involved in critical cellular processes, such as cell death, cell cycle arrest, cell cycle regulation, lung development, hypoxia and cell proliferation. These processes are known to be dysregulated in cancer (3, 44, 45).

The first process regulated by the cluster is cell death, which will be regulated by two targets of the miR-17/92 cluster. BCL2 like 11 (BCL2L11or BIM) is a tumour suppressor gene. It acts as a pro-apoptotic facilitator. The expression of this gene is inhibited by miR-92, miR-17 and miR-20a. Moreover, Phosphatase and Tensin Homolog (PTEN), a tumour suppressor gene, also functions within this process and is inhibited by miR-17, miR- 20a and miR-19b-1. Inhibition of these targets results in less cell death, which increases the survival of tumour cells (3, 30, 47).

Also, cell cycle arrest and regulation will be managed by the cluster. In cell cycle regulation, E2F1 is involved at cell cycle checkpoint G1/S, which is previously explained. An arrest of the cell cycle is negatively regulated at checkpoint G1/S by Cyclin-Dependent Kinase Inhibitor 1A (CDKN1A or P21). This can initiate tumour growth. CDKN1A is directly downregulated by miR-17, miR-18a and miR-20a, which results in uncontrolled cell cycle progression (1, 3, 30, 47).

Transforming Growth Factor Beta (TGF-β) signalling pathway is involved in the regulation of cell proliferation. Transforming Growth Factor Beta Receptor 2 (TBFBRII) is important within this pathway and regulates multiple processes, such as wound healing, immunosuppression, cell proliferation and cell cycle arrest. Dysregulation can contribute to tumorigenesis. TGFBRII is directly targeted by two members of the cluster, namely miR-17 and miR-20a (3, 30, 47).

STAT3 initiates the expression of the cluster as previously explained. However, it also targeted by the cluster and functions in the development of the lung, like Mitogen-Activated Protein Kinase 14 (MAPK14). Moreover, MAPK14 is also involved in cell proliferation, differentiation, transcriptional regulation and developmental processes. These two genes will be regulated by miR-17/92, more specifically by miR-17 and miR-20a (1, 3, 47).

Table 2. Overview of different targets of the miR-17-92 cluster. These targets will be inhibited by the cluster (1-3, 30, 47).

Name	Inhibition via	Function
Phosphatase and Tensin Homolog (PTEN)	- miR-17 - miR- 20a - miR-19b-1	- Cell spreading - Cell death - Cell cycle progression - Cell survival
Transforming Growth Factor Beta Receptor II (TGFBRII)	- miR-17 - miR-20a	- Wound healing - Immunosuppression - Tumorigenesis - Cell proliferation - Cell cycle arrest
BCL2 like 11 (BCL2L11)	- miR-92 - miR-17 - miR-20a	- Apoptosis (pro-apoptotic gene)
Cyclin Dependent Kinase Inhibitor 1A (CDKN1A)	- miR-17 - miR-18a - miR-20a	- DNA damage - DNA replication - Cell cycle progression
Mitogen-activated Protein Kinase 14 (MAPK14)	- miR-17 - miR-20a	- Transcription regulation - Lung development - Cell proliferation - Cell differentiation
Hypoxia Inducible Factor 1 Subunit Alpha (HIF-1α)	- miR-17 - miR-20a	- Angiogenesis - Apoptosis - Regulating hypoxia - Energy metabolism

Hypoxia, a condition characterised by a bad oxygen supply in a certain tissue or region of the body, is regulated by the cluster via Hypoxia Inducible Factor 1 Subunit Alpha (HIF-1α). HIF-1α is involved in both the systemic and cellular homeostasis of oxygen supply and will respond to hypoxia by activating genes involved in energy metabolism, angiogenesis and apoptosis, which is critical in many cancer types. Inhibition by miR-17 and miR-20a negatively influence its expression and can contribute via this mechanism to the progression of cancer (3, 47).

Not much is known about the effects of *in utero* air pollution exposure on the alteration in the expression of both up- and downstream targets of the miR-17/92 cluster (in particular miR-17, miR-20a and miR-92a) in the intra-uterine environment. Therefore, the objectives for the current research are: (1) build up a network around the cluster including both up- and downstream targets (MYC, E2F1, STAT3, TP53, MAPK14, TGFBR2, HIF-1α, CDKN1A, PTEN and BCL2L11), which are known to be involved in tumorigenesis and (2) investigate potential alterations in the expression of these regulatory molecules and targets in cord blood in association with *in utero* air pollution exposure by using RT-qPCR. This is a method to detect and quantify the products of a certain gene of interest generated during each PCR cycle (48).

This research will provide more biological insight in the underlying epigenetics mechanisms, by microRNA expression, on the regulation of telomere length and the oncogenic mechanisms of the miR-17/92 cluster in early life, in response to prenatal air pollution exposure.

2 Material and methods

2.1 Study design and study population

The ongoing, longitudinal, population-based ENVIR*ON*AGE birth cohort enrols mother-child pairs when arriving for delivery at the East-Limburg Hospital (ZOL, Genk, Belgium). The catchment area of the study population is approximately 2 422 km^2, with a population density of 82 to 743 inhabitants/km^2 and compromises rural, suburban and urban regions.

After delivery, written informed consent was obtained from all eligible participants. The only inclusion criterium was the ability to fill out a questionnaire in Dutch. At the post-delivery ward, detailed information was obtained from study questionnaires on maternal age, parity, education level, occupation, smoking behaviour, alcohol consumption, residential address, use of medication, newborn's ethnicity and paternal age. Perinatal parameters were also collected, such as birth date, gestational age, newborn's gender, birth weight and length, duration of the labour, Apgar score, pH of the arterial cord blood and ultrasonographic data.

Within this research, two different hypotheses were covered. In Figure 3, the study design of this research was visually presented. The study consisted of two panels. The hypothesis "mRNA expression of both up- and downstream targets, known to be involved in cancer initiation, of three specific miR-17/92 cluster members (miR-17, miR-20a and miR-92a) were associated with prenatal exposure to air pollution", was situated in Figure 3, panel A. First of all, the expression of the miR-17/92 cluster was measured in cord blood with RT-PCR on 432 individuals (Figure 3, panel A, part (1)) and in the placenta using miRNA microarray analysis in 60 individuals (Figure 3, panel A, part (3)). These data sets were available in the research group. In this research, the association was determined between the miR-17/92 cluster members and prenatal air pollution exposure. Moreover, the expression of the paralogue clusters, miR-106a-363 and miR-106b-25, was measured with microarray analysis (Figure 3, panel A, part (3), microarray data were already available) and its association with prenatal air pollution exposure was checked within the present study. Within the current research, the mRNA expression of up- and downstream targets of the miR-17/92 cluster was measured in 432 individuals with RT-PCR within the current research (Figure 3, panel A, part (2)) and checked for associations with air pollution exposure during pregnancy.

In the first part of this research (Figure 3, panel A), the expression of the network around the miRNA cluster miR-17/92 is measured in cord blood (Figure 3, panel A, part (2)), using RT-PCR. Therefore, 511 mother-newborn pairs, recruited between February 2015 and February 2017, were selected for the present study (only singletons). Individuals were selected based on the population, of which data of the miRNA expression in cord blood were available within the research group, form part (1) in panel A of Figure 3.

The second part of this research (Figure 3, panel B) covered the hypothesis: "The placental candidate miRNAs will be differentially expressed in association with TL at birth". In the placenta, mRNA gene expression was measured in 200 individuals, using mRNA microarray analysis (Figure 3, Panel B, part (1)). This population was independent of the individuals in which miRNA expression of miR-17/92 and mRNA expression of its network was measured in cord blood (Figure 3, Panel A, part (1) and (2)). From these 200 subjects, 60 individuals were selected based on TL. Here, 30 individuals had a short TL, including 15 boys and 15 girls, the other 30 individuals had a long TL and also with an equal amount of boys and girls. In this population of 60 individuals, all known human miRNAs were measured with miRNA microarray analysis (Figure 3, Panel B, part (2)). Both the mRNA and miRNA microarray data were already available within the research group. Using GO terms and the mRNA microarray dataset, a subset of miRNAs of the whole miRNA microarray dataset was selected to determine their association with TL at birth.

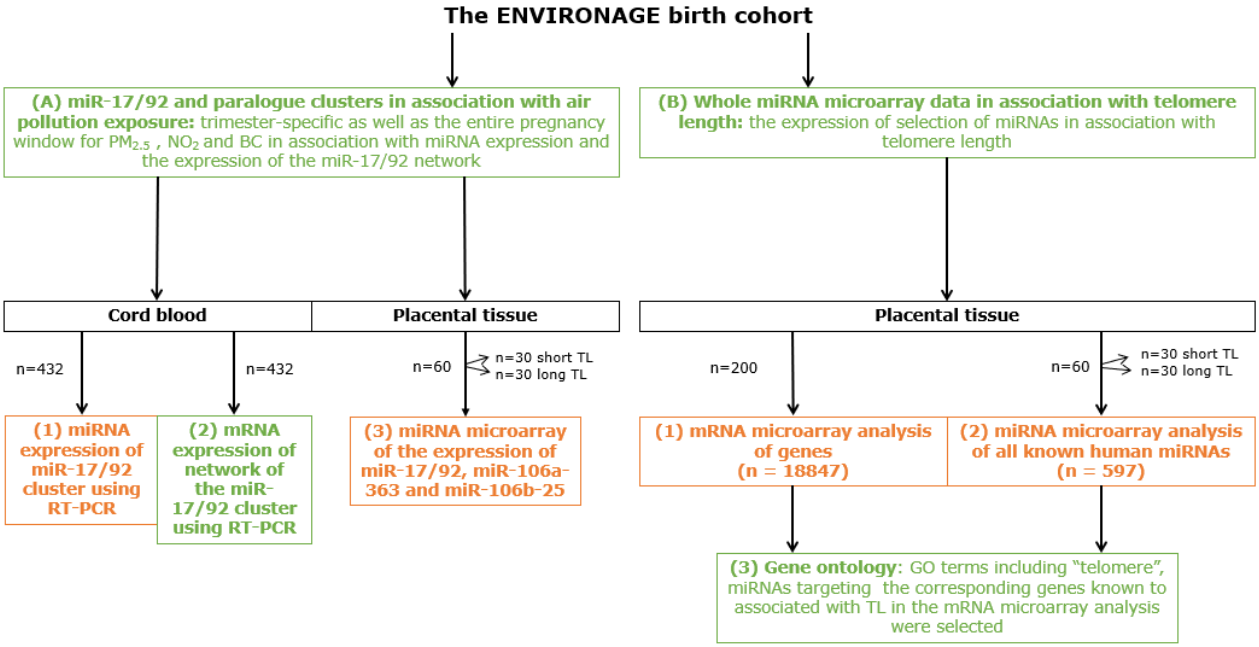

Figure 3. Flowchart of the study design. The current research consists of two independent panels, within the setting of the ENVIR*ON*AGE birth cohort. **(A)** The association between prenatal air pollution exposure and the miRNA expression of miR-17/92, miR-106a-363 and miR-106b-25 as well as the association between the mRNA expression of the up-and downstream targets of the miR-17/92 cluster and prenatal air pollution exposure was determined. This panel includes data from three subparts **(1)** The expression of the miR-17/92 cluster was determined in cord blood of 432 new-born children using RT-PCR. **(2)** In the placenta, the expression was measured of the miR-17/92 cluster as well as the paralogue clusters miR-106a-363 and miR-106b-25 using miRNA microarray analysis. The study population consisted of 60 individuals, divided into 2 groups of 30 individuals. One group included individuals with short telomere length (TL), the other group included individuals with a long TL. Both groups had an equal amount of boys and girls. **(3)** The mRNA expression of both up- and downstream targets were measured in cord blood of 432 individuals, using RT-PCR and checked for associations with prenatal air pollution exposure. **(B)** The association was determined between the miRNA expression in the placenta and TL at birth. **(1)** In the placenta of 200 individuals, the expression of 18847 genes was measured using mRNA microarrays and associated with TL. **(2)** The expression of all known human miRNAs was measured in the placenta, using microRNA microarray analysis. This population included 60 individuals selected from the population in (B1). The population is divided into two groups of 30 individuals, one group included individuals with short TL and the other group long TL, and both groups had an equal amount of boys and girls. **(3)** Using Gene ontology terms, including the word "telomere", and both the mRNA microarray and miRNA microarray dataset, a selection of miRNAs was made to test for associations with telomere length at birth. Data, which were already available in the research group were indicated in orange and were used in the current research. In green, the panels were given, which were performed in the current research. RT-PCR = real-time polymerase chain reaction, TL = telomere length, $PM_{2.5}$ = particulate matter <2.5µm, BC = black carbon.

For the second part of this research (Figure 3, Panel B), 200 individuals were selected from the ongoing ENVIRONAGE cohort (Figure 3, Panel B, part (1)). From this population, 60 individuals were selected, based on TL as previously explained, in which miRNA expression was measured in placental tissue (Figure 3, Panel B, part (2)). For Panel A, part (3), in which the miRNA expression of the miR-17/92, miR-106a-363 and miR-106b-25 cluster was measured, the data of the miRNA microarray data was used (Figure 3, Panel B, part (2)), to select the three clusters for further analysis.

The study protocol is approved by the ethical committee of both the University of Hasselt and the East-Limburg Hospital. Moreover, it is according to the declaration of Helsinki for research on human subjects. This subpopulation showed no significant differences in parameters, such as maternal age, pre-gestational body mass index (BMI), parity, and ethnicity of the new-born in comparison with the birth register of Flanders, Belgium (49).

2.2 miR-17/92 expression in cord blood and placental tissue and miR-17/92 network expression in cord blood in association with prenatal air pollution exposure

2.2.1 Exposure to particulate matter air pollution measurement

The regional background concentrations of $PM_{2.5}$, nitrogen dioxide (NO_2) and black carbon (BC) in $\mu g/m^3$, were estimated for each maternal home address, using a spatial-temporal interpolation method. This method provided daily information about the different pollutant concentrations by creating a high-resolution receptor grids with a surface of 25x25 m^2. This method used the information of land covering data and $PM_{2.5}$ pollution data, which were obtained respectively of satellite images and the official fixed-site monitoring network. The interpolation method is combined with a dispersion model (kriging). This method inserted telemetric air quality network data together with point and line sources. Meteorological data from the Royal Meteorological Institute, like the speed and direction of wind and temperature were also implemented in this method. This tool explains more than 80% of the variability in the Flemish region of Belgium (50-53).

The exposure to $PM_{2.5}$ during the whole pregnancy period was calculated. Critical exposure windows were determined by calculating the average daily exposure levels for the three trimesters of pregnancy, with trimester 1 (1-13 weeks), trimester 2 (14-26 weeks) and trimester 3 (27 weeks to delivery). Based on the first day of the last menstrual cycle and the first ultrasound examination, the date of conception was determined. Information of the residential home address during pregnancy was obtained by questionnaires and checked in the medical records. If mothers did move during pregnancy, the trimester-exposures were calculated in a way that allows changes in residence.

2.2.2 Sample collection and RNA extraction from cord blood and placental tissue

After delivery, umbilical cord blood was stored in PAXgene Blood RNA tubes (Preanalytix GmbH, Feldbachstrasse, Switzerland) and the samples were stabilized for 24 hours at room temperature. Next, the blood tubes were stored at -20°C for longer periods prior to analysis.

After stabilizing the whole blood samples, total RNA and miRNA were extracted with the PAXgene Blood miRNA isolation kit (Preanalytix GmbH, Feldbachstrasse, Switzerland). The purification proceeded according to the manufacturer's protocol and used a spin-column, silica-based RNA purification technique. The isolated RNA was quantified, using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and the quality was checked, using Agilent 2100 Bioanalyzer (Agilent Technologies, Amstelveen, the Netherlands). Next, the samples were stored until further use at -80°C.

Within the first hour after delivery, fresh placental tissue was also collected. At the foetal side, four biopsies were taken at four centimetres distance from the umbilical cord, which is described in detail by Janssen *et al* (4). Next, the biopsies were stored in RNA later (ThermoFisher Scientific, Massachusetts, USA) at minimum 12 hours and maximum 24 hours at a temperature of 4°C. Subsequently, the biopsies were stored at -20°C, until further use.

From one single biopsy, the total RNA and miRNA were extracted using the miReasy mini kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's protocol. Subsequently, an RNase-free DNase treatment (Qiagen, Venlo, The Netherlands) was carried out, according to the standardised protocol. The quantity and purity of the samples were determined by spectrophotometry (nanodrop 100, Isogen, Life Science, Belgium) as well as the RNA integrity, assessed by Agilent 2100 Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands).

2.2.3 cDNA synthesis from cord blood

The synthesised total RNA, isolated from the cord blood samples, was reverse transcribed, using the GoScriptTM reverse transcription System kit (Promega Corporation, Madison, USA) according to the manufacturer's protocol. An RNA input concentration of 1.5 µg is used and combined with both Oligo dT primer and random primer. Subsequently, this mix was denaturated at 70°C for five minutes, chilled on ice and centrifuged.

Next, a reverse transcription mixture was added, containing nuclease-free water, a GoScript reaction buffer, magnesium chloride, a PCR nucleotide mix including deoxyribonucleotide triphosphates (dNTP's), a reverse transcriptase, and a recombinant RNasin® Ribonuclease Inhibitor. Finally, complementary deoxyribonucleic acid (cDNA) was synthesized. First, an annealing step was performed at 25°C for five minutes. Next, the primers were extended at 42°C for 60 minutes. Finally, an additional step was carried out to inactivate the reverse transcriptase, namely incubation of the sample mixture at 70°C for 15 minutes. Afterwards, the cDNA is stored at -20°C until further use.

2.2.4 Primer efficiency testing

The efficiency of the primers, according to the different up- and downstream targets were determined. The targets were chosen, based on literature studies. The chosen targets were PTEN, MAPK14, HIF-1α, E2F1, CDKN1A, TGFBRII, BCL2L11, TP53, MYC and STAT3. The primer assay used and the corresponding RefSeq are given in Table 3.

Table 3. Overview of different up- and downstream targets of the miR-17/92 cluster with the corresponding primer assay and RefSeq used for RT-PCR.

Gene of interest	*Primer assay*	*RefSeq*
Up-stream targets		
E2F1	Hs.PT.58.45513742	NM_005225
C-MYC	Hs.PT.58.39142481	NM_002467
STAT3	Hs.PT.58.20367494	NM_003150
TP53	Hs.PT.58.39489752.g	NM_000546
Downstream targets		
BCL2L11	Hs.PT.58.18885697	NM_001204109
CDKN1A	Hs.PT.58.40874346.g	NM_000389
HIF-1α	Hs.PT.58.534274	NM_001243084
MAPK14	Hs.PT.58.40355791	NM_001315
PTEN	Hs.PT.58.4416071	NM_00314
TGFBRII	Hs.PT.58.2243027	NM_0010224847

To test the efficiency of the primers, the TaqMan Fast advanced protocol was performed using the 7900HT Fast Real-Time PCR system (Thermo Scientific, Wilmington, DE, USA), with the Taqman detector. First, a mixture of different cord blood samples was made with a concentration of 75 ng/µl. Next, the standard curve was made by serially diluting the mixture sample by twofold to become seven different concentrations. Mastermix was added to each dilution (2µl) and was composed of TaqMan Fast advanced Master Mix (5µl), 20x TaqMan gene expression assay (0.5µl), RNase free water (2.5µl). PCR amplification was performed according to the standard temperature protocol. This protocol started with one cycle at 50°C for two minutes. Next, one cycle of ten minutes was performed at 95°C. Finally, 40 cycles of 20 seconds were performed at 95°C, alternated by one minute at 60°C.

The different dilutions were measured in triplicate. The efficiency was calculated based on the standard curve. Therefore, the log transformation of each dilution was plotted in function of the corresponding average Ct value. Triplicates were only included if ΔCq was smaller than 0.3. Next, the efficiencies were determined, according to the formula: [efficiency (%) = $(10^{(-1/slope)} -1) *100$].

2.2.5 mRNA expression analysis

Ten specific target mRNAs were measured in cord blood, using a Taqman Fast advanced protocol. Therefore, a 7900HT Fast Real-Time PCR system (Thermo Scientific, Wilmington, DE, USA) was used, with the TaqMan detector. Within the analysis, three reference genes are included for normalization namely: CYC1, TBP and IPO8. In Table 4, the primer sequence, primer assay and RefSeq are given for each reference gene used in this research.

RT-PCR is a method used to quantify the mRNA levels in a sample. TaqMan Gene expression Assays are used to detect PCR products of interest, accumulated during PCR, and is based on a 5'endonuclease chemistry using a fluorescent probe. This special probe, named TaqMan, is labelled with a fluorescent FAM™ dye on the 5' end and a non-fluorescent quencher molecule at the 3' end. When the probe is intact, the quencher molecule captures the energy of the fluorescent dye, known as fluorescent resonance energy transfer (FRET). If FRET occurs, no fluorescent signal can be observed. The RT-qPCR process starts with a denaturation step. During this step, the temperature will be increased, and the double-stranded cDNA molecule denaturates. Next, the temperature will be lowered. During this step, both the unlabeled primers and the TaqMan probe can anneal to the target sequence. Subsequently, Taq-DNA polymerase will extend the primers and synthesises a new strand. If Taq-DNA polymerase arrives at the bound TaqMan probe, it will cleave the probe, using its endogenous 5'endonuclease activity, which separates the quencher and fluorescent dye. Each PCR cycle, more fluorescent dyes will be released, which increases the fluorescent signal. This increase in signal is proportional with the amount amplicon synthesized of the target sequence.

According to the protocol, the MasterMix (see "2.2.4 Primer efficiency testing") was added to the cord blood samples with an input concentration of 10 ng/µl cDNA. Next, PCR was performed, according to the standard temperature protocol (see "2.2.4 Primer efficiency testing").

To minimize possible technical variation between multiple runs of the same gene, inter-run calibrators (IRCs) were included on each plate. Non-template controls (NTCs), negative controls were used for each mRNA assay, in triplicate. Expression of the target genes was determined and Cq values were obtained with SDS 2.3 software (Applied Biosystems, Foster City, CA). qBase plus software (Biogazelle, Zwijnaarde, Belgium) was used to calculate the relative gene expression according to the $2^{-\Delta\Delta Cq}$ method. Sample analysis was performed in triplicate and only included if ΔCq was smaller than 0.3.

Table 4. mRNA primers of the reference genes for RT-qPCR. In the table, the primer sequence, primer assay and the RefSeq are given for the three reference genes, including CYC1, TBP and IPO8.

Full name	Gene	Forward primer	Primer assay	RefSeq
Cytochrome C1	CYC1	F: 5'-ACACATCCTTGGCTATCTGG-3' R: 5'-CTCTACTTCAACCCCTACTTTCC-3	Hs.PT.58.20696349.gs	NM_001916
TATA-box binding protein	TBP	F: 5'-TCGTGGCTCTCTTATCCTCAT-3' R: 5'-CAGTGAATCTTGGTTGTAAACTTGA-3'	Hs.PT.56a.20792004	NM_003194
Importin 8	IPO8	F: 5'-CCACTTCTTACACTTCCACCAT-3' R: 5'-GAGATCTTCCGAACTATTATCGACA-3'	Hs.PT.56a.40532361	NM_001190995

2.2.6 miRNA expression profiling

The placental expression of all known human miRNAs, including the miR-17/92 cluster and the paralogue clusters miR-106a-363 and miR-106b-25, was measured in 60 individuals (Figure 3, Panel A, part (3) and panel B part (2)). The participants were selected based on the TL, 30 individuals with short TL and 30 individuals with long TL and each group included an equal amount of boys and girls. After hybridisation of the total RNA, the microRNA expression was measured, using Sureprint G3 Human V19 miRNA Agilent 8 x60K microarrays. After extraction of the data, using Agilent Feature Extraction Software, the quality assessment, pre-processing and quantile normalization were accomplished by R package AgiMicroRna. miRNAs were excluded if less than 70% of the data were present over all the samples. The missing values were attributed by K-nearest neighbour imputation (K=15). The expression data of 597 miRNAs, out of 2558 miRNAs, can be used for further analysis.

2.2.7 Statistical analysis

Statistical analysis (Figure 3, panel A, part (2)) was performed, using SAS software (Version 9.4 SAS Institute, Cary, NC, USA). The relative gene expression of the up- and downstream targets of the miR-17/92 cluster in cord blood was log-transformed (Log10), because of non-normal distribution. Continuous data were presented as mean (± standard deviation (SD)) and categorical data as numbers or frequencies (%). Multiple linear regression was performed to assess the association between prenatal air pollution exposure and the relative gene expression of both the up- and downstream targets of the miR-17/92 cluster. This model was corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m^2), newborn's sex, educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), parity (1,2 or ≥3), gestational age (weeks), ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy. The estimates, effect of air pollution on gene expression, was shown as the percentage of change (with a 95% confidence interval (CI)) in gene expression for each 5 $\mu g/m^3$ increment of $PM_{2.5}$ or NO_2 or and 0.5 $\mu g/m^3$ increment of BC exposure at each specific time window of pregnancy. The percentage of change, for example for $PM_{2.5,}$ was calculated using the formula: [change (%) = $(10^{(\beta*5)} -1) *100$]. Spearman correlations for each miRNA (miR-17, miR-20a and miR-92a) and target gene were determined as well as the correlations among the different miRNAs.

Subsequently, the association between air pollution (Figure 3, panel A, part (3)) exposure during pregnancy and the placental expression of the miRNAs belonging to the miR-17/92, miR-106a/363 and miR-106b/25 clusters were determined using multiple linear regression analysis. First, a selection of the 597 miRNAs was made, including all the measured miRNAs belonging to one of these clusters. Statistical analysis was performed using SAS software (version 9.4, SAS Institute, Cary, NC, USA). Log-transformation (LOG2) of the miRNA expression was performed, because of non-normal distribution. Multiple linear regression models were performed to determine the association between the miRNA expression of the members of both the miR-17/92 cluster and its paralogue clusters in the placenta and prenatal air pollution exposure like previously explained in above.

2.3 Whole microRNA microarray data in the placenta in association with telomere length

2.3.1 Relative telomere length determination

The relative telomere length in the placenta was already available for each individual of the present study population. After DNA extraction from the placental biopsy, the relative telomere length was determined, using a monochrome multiplex quantitative PCR method. This method is more detailed described in Martens *et al* (54, 55).

2.3.2 Transcriptomic profiling of the placenta by microarray

The relative expression of 18847 different genes was measured in the placenta, using Agilent whole human genome 8x60K microarrays (Agilent Technologies, Amstelveen, The Netherlands). These mRNA microarray data were available, within the research group, for further analysis (Figure 3, Panel B, part (1)).

2.3.3 miRNA expression profiling

In 60 subjects of the initially selected population, selected based on TL (30 individuals with long TL and 30 individuals with short TL), the placental expression of all known human microRNAs was determined using microarrays (Figure 3, Panel B, part (2)), which is described in more detail in "2.2.6 miRNA expression profiling".

2.3.4 Statistical analysis

To determine the association between telomere length and miRNA expression in the placenta, we selected a subset of microRNAs from the entire miRNA microarray dataset (Figure 3, Panel B, part (2)). A miRNA selection was made based on Gene Ontology (GO) terms including the word "telomere" (Appendix 1, Table 12). Therefore, mRNA microarray data were used. The sex-specific association between the different genes and the relative placental TL was determined using a mixed linear model (Rstudio, R version 3.3.2). The models were adjusted for batch effects, age of the father (years), maternal age (years), maternal pre-gestational BMI (kg/m^2), parity (1,2 or ≥3), maternal smoking status (never-smoker, past-smoker or current-smoker), maternal educational status (low, middle or high) newborn's ethnicity (European or non-European) and gestational age (weeks). After statistical analysis, two lists were obtained with significantly associated genes with TL for boys and girls separately. The top 50 of significant genes are given in appendix 2 for both boys (Table 13) and girls (Table 14).

Based on these data a set of miRNAs was selected. First, the mRNA microarray data were used to select a subset of genes significantly associated with TL. Using the previously mentioned GO terms, including the term "telomere", we selected significant genes when they belong to one of these GO terms. Subsequently, the miRNAs targeting these genes were predicted for boys and girls separately, using the miRDB database based on a prediction target prediction described by Liu *et al* (56). These miRNAs were used as input for statistical analysis.

A list of 552 miRNAs (for the boys, Appendix 3 Table 15) and 98 (for the girls, Appendix 3 Table 16) were used as input. The overlap with the miRNA microarray data was used for further analysis. The sex-specific association between the telomere length and the miRNA expression in the placenta was determined using a mixed linear model (Rstudio, R version 3.3.2). The models were adjusted for the date of hybridisation, the father's age (years), maternal education status (low, middle or high), parity (1,2 or ≥3), season of delivery, date of delivery and gestational age (weeks). Using the Benjamini-Hochberg false discovery rate (FDR), the p-value was adjusted for multiple testing, considering a significance level of 5%.

3 Results – miR-17/92 network expression in cord blood in association with prenatal air pollution exposure

3.1 Characteristics of the study population

Detailed information about demographic and lifestyle characteristics, of both mother and child, are shown in Table 6 and exposure characteristics are shown in Table 5. The population, described here, is situated in the overview Figure 3 in panel A, part (1) and (2). We started the experiments with a population of 511 individuals. However, due to technical errors, low expression levels, inefficient expression or the lack of available miRNA expression data of the miR-17/92 cluster, 79 individuals were excluded from the study population. To summarise, this population counts 432 individuals of which the relative expression of the miR-17/92 cluster and its network were available.

Table 5. Characteristics of air pollution exposure data (µg/m³) for the study population (n=432). The miRNA population counts 432 individuals of which the relative expression of the miRNA cluster miR-17/92 was measured in cord blood as well as the network of up- and downstream targets. The expression of the different exposures is given for each time-window of pregnancy (µg/m³) as the mean (SD). SD= Standard deviation, IQR=Interquartile Range

	miRNA (n=432)	
Exposure data	***Mean (SD)***	***IQR***
Particulate Matter <2.5µm ($PM_{2.5}$)(µg/m³)		
Trimester 1	12.0 (3.4)	4.6
Trimester 2	11.7 (3.8)	5.4
Trimester 3	11.6 (3.6)	4.9
Entire pregnancy	11.7 (2.1)	3.2
Black carbon (BC) (µg/m³)		
Trimester 1	1.2 (0.4)	0.5
Trimester 2	1.1 (0.3)	0.5
Trimester 3	1.1 (0.3)	0.4
Entire pregnancy	1.1 (0.3)	0.3
Nitrogen dioxide (NO_2) (µg/m³)		
Trimester 1	16.5 (4.9)	6.2
Trimester 2	16.0 (4.9)	7.0
Trimester 3	15.8 (4.6)	5.9
Entire pregnancy	16.1 (3.8)	5.5

Table 5 represents the residential prenatal exposure data of different pollutants ($PM_{2.5}$, NO_2 and BC) in µg/m³ for the current population, as explained above. The $PM_{2.5}$ residential prenatal exposure was, for the current research, is the lowest in the third trimester, 11.6 (±3.6) µg/m³, and the highest in the first trimester, 12.0 (±3.4) µg/m³, with corresponding interquartile range (IQR) of respectively 4.9 and 4.6. The residential exposures during pregnancy for BC showed the same pattern, with the highest exposure in the first trimester, which is 1.2 (±0.4) µg/m³ with an IQR commonly of 0.5. The exposure of the second, third as well as the entire pregnancy period amounts 1.1 (±0.3) µg/m³, with an IQR of respectively 0.5, 0.4 and 0.3. The residential NO_2 exposure is also determined and ranges between 15.8 (±4.6, IQR= 5.9) µg/m³ in the last trimester of pregnancy and 16.5 (±4.9, IQR=6.2) µg/m³ in the first trimester of pregnancy.

Lifestyle and demographic characteristics (Table 6) included a mean maternal age (±SD) of 29.7 (±4.4) years and the mother's pre-gestational body mass index (BMI) averaged (±SD) 24.6 kg/m^2 (±4.8), in the current population. 51.7% of the mothers were highly educated and 65.1% had never smoked. 86.5% of the newborns are of European descent and the newborn population is 54.1% male. The children's gestational age is on average 39.2 weeks (±1.6) and ranged between 38 and 40 weeks. At birth, newborns had an average weight (±SD) of 3412.7 (±486.3) g. The children were born equally divided over the different seasons and 51.5 % of them was their mother's first child.

Table 6. Demographic characteristics of the study population (n=432). The miRNA population counts 432 individuals of which the relative expression of the miRNA cluster miR-17/92 was measured in cord blood as well as the network of up- and downstream targets. Categorical data were presented as numbers or percentages (%) and continuous data were presented as mean (±SD). Maternal education level was characterised as 'low' (no diploma or primary school), 'middle' (high school) or 'high' (college or university degree). Ethnicity was classified as 'European-Caucasian' (when two or more grandparents of the new-born were European) or 'non-European'.

	miRNA (n=432)		
Characteristic	***Mean (SD)***	***No.***	***%***
Mother			
Age (years)	29.7 (4.4)		
Pre-gestational BMI (kg/m^2)	24.6 (4.8)		
Educational level			
Low		50	11.6
Middle		176	36.7
High		223	51.7
Parity			
1		222	51.5
2		151	35.0
≥3		58	13.5
Smoking status			
Never		278	65.1
Before pregnancy		103	24.1
During pregnancy		46	10.8
New-born			
Sex, Male		233	54.1
Ethnicity			
European-Caucasian		373	86.5
Non-European		58	13.5
Gestational age (weeks)	39.2 (1.6)		
Birth weight (g)	3412.7 (486.3)		
Season of birth			
Winter		93	21.6
Spring		118	27.4
Summer		95	22.0
Autumn		125	29.0

Characteristics of the population used to measure the expression the miR-17/92, miR-106a-363 and miR-106b-25 clusters in the placenta of 60 individuals, using miRNA microarray analysis (Figure 3, Panel A, part (3)) are given in "4.1 characteristics of the study population".

3.2 Primer efficiency

The efficiencies of each target primer pair are presented in percentages (%), in Table 7. The primers were found to be efficient for PCR amplification if the efficiency ranges between 90 and 110%. For the upstream targets, the lowest efficiency is seen for TP53 and has a value of 90.94%. The best primer efficiency of the upstream targets is found for C-MYC and amounts 103.94%. For the downstream targets, BCL2L11 has the least reprehensible efficiency, which is 93.15%. Meanwhile, MAPK14 is found to have the highest efficiency of the downstream targets, which amounts 100.32%. All the chosen primers range between the standard norm of 90-110% as explained above.

Table 7. The efficiency (%) of primers for both up- and downstream target primer pairs.

Gene of interest	*Primer efficiency (%)*
Up-stream targets	
E2F1	95.77
C-MYC	103.94
STAT3	99.49
TP53	90.94
Downstream targets	
BCL2L11	93.15
CDKN1A	95.66
HIF-1α	100.21
MAPK14	100.32
PTEN	96.95
TGFBRII	95.91

3.3 Cord blood expression of miR-17, miR-20a and miR-92a in association with air pollution exposure and the correlations with both their up- and downstream targets

Both the adjusted model (Figure 4), corrected for date of delivery and adjuster for some covariates: maternal age (years), pre-gestational BMI (kg/m^2), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy and the unadjusted model (Appendix 4, Table 17) shows a significant negative association between $PM_{2.5}$ and both miR-17 and miR-20a expression during the 1st and 3rd trimester of pregnancy, which were respectively -31.4% (95% CI: -45.57% to -13.43%, p= 0.002) and -29.4% (95% CI: -44.79% to -9.65%, p= 0.006) for miR-17 and -25.6% (95% CI: -40.87% to -6.36%, p= 0.01) and -35.57% (95% CI: -49.33% to -18.09%, p= 0.0004) for miR-20a. Moreover, the entire pregnancy for miR-17 is significantly associated with a decrease of 34.02% (95% CI: -54.48% to -4.37%, p=0.03) for each increment of 5 μg/m^3 of $PM_{2.5}$ exposure and a trend to significance for miR-20a. Whereas, in trimester 2 of pregnancy, an estimated change of in miR-20a is seen for a 5 μg/m^3 increase of $PM_{2.5}$.

For a 5 μg/m^3 increment of NO_2, a significant association is only found for the 1st trimester of pregnancy. Trimester-specific estimates are -16.71% (95% CI: -28.6% to -2.85%) for miR-20a and -19.00% (95%CI: -30.71% to -5.32%) for miR-17.

The relative expression of the three miRNAs of interest (miR-17, miR-20a and miR-92a) was also analysed in association with BC. A negative association is seen between the relative expression of miR-17 and BC in the 1st trimester of pregnancy, with a decrease of 27.1% (95% CI: -39.98% to -11.45%, p=0.002) for an increase of 0.5 μg/m^3 in BC exposure. However, in the 3rd trimester, a positive association is seen, with an increment of 27.05% (95% CI: 1.52% to -59.01%, p=0.04) for each increase of 0.5 μg/m^3 in BC exposure. Trimester-specific estimates of miR-20a exposure were only significant for the first trimester of BC exposure. The relative miR-20a expression decreases with 21.18% (95% CI: -34.98% to -4.46%) for an increment of 0.5 μg/m^3 exposure to BC.

The relative expression of miR-92a was only significantly associated with BC exposure in the second and third trimester of pregnancy. The estimated changes were respectively -19.66% (95%CI: -33.33% to -3.18%) and 20.74% (95% CI: 1.07% to 44.24%) for an increment in BC exposure of 0.5 μg/m^3.

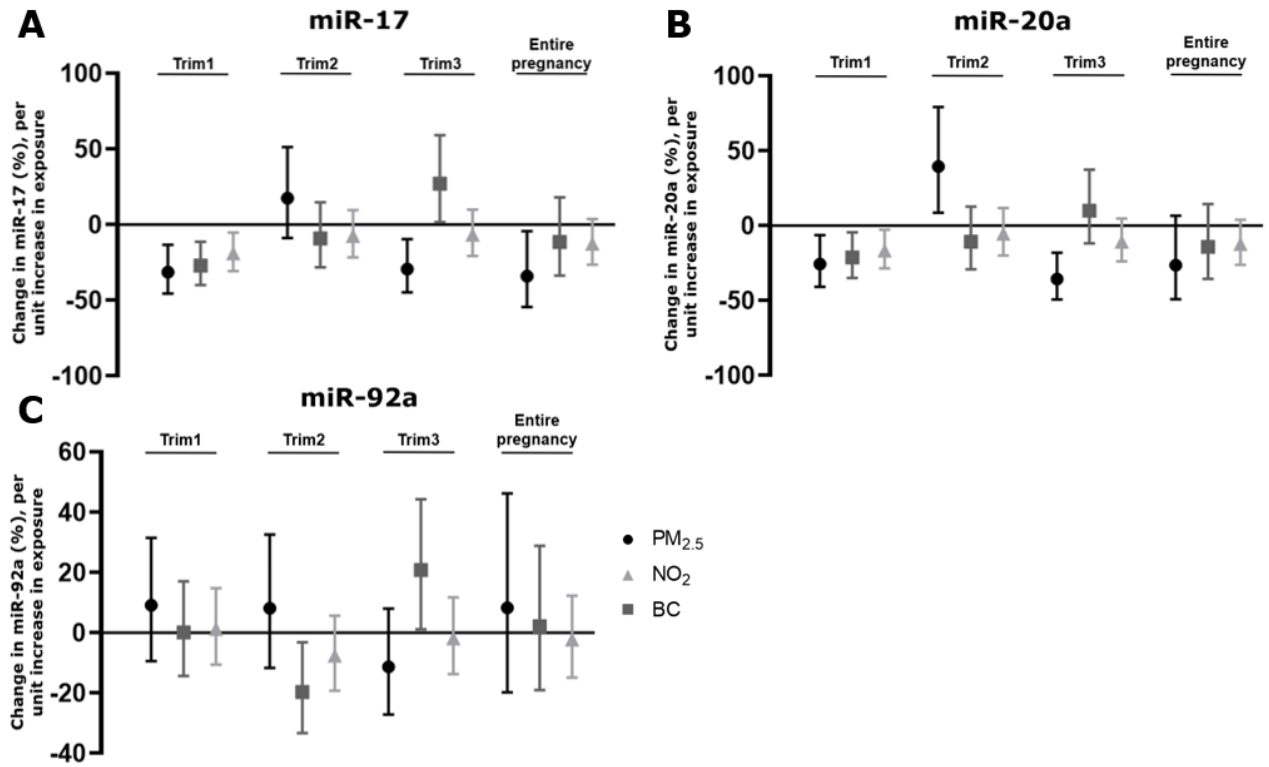

Figure 4. Estimated change (%) in relative miRNA expression in cord blood for each increase of 5 µg/m³ of $PM_{2.5}$ and NO_2 or 0.5 µg/m³ increment in BC for the different time windows of pregnancy of (A) miR-17, (B) miR-20a and (C) miR-92a. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

The Spearman correlations between the relative miRNA expression (miR-17, miR-20a and miR-92a) and the relative expression of the up- and downstream targets are shown in Appendix 4, Table 18. miR-17 expression is only significantly correlated with the gene expression of MYC. The Spearman correlation coefficient amounts -0.16. For miR-20a, a significant correlation is seen with the expression of MYC and TP53, with Spearman correlation coefficients of respectively -0.16 and -0.11. Moreover, HIF-1α and PTEN gene expression are significantly associated with miR-92a expression. The Spearman correlation coefficients are respectively -0.10 and -0.14. Correlation among the different miRNAs was determined, using Spearman correlations (Appendix 4, Table 19). Significant positive correlations are found among all three miRNAs (miR-17, miR-20a and miR-92a) with $r>0.56$ ($p<0.001$).

3.4 Placental expression of the miR-17/92 cluster and the paralogue clusters mir-106a-363 and miR-106b-25 in association with air pollution exposure

Both before (Appendix 5, Table 20 and Table 21) and after adjustment (Figure 5) for maternal age (years), pre-gestational BMI (kg/m^2), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's sex, ethnicity (European and non-European), date of delivery and the apparent temperature for each specific time window of pregnancy, a positive association is found between the different air pollutants ($PM_{2.5}$, NO_2 or BC) and five specific miRNAs of the miR-17/92 cluster (Figure 5), such as miR-17-3p, miR-17-5p, miR-18a-5p, miR-19a-3p and miR-20a-5p. Moreover, miRNAs belonging to the paralogue clusters miR-106a-363 (Appendix 5, Figure 10) and miR-106b-25 (Appendix 5, Figure 11) are also found to be positively correlated with air pollution exposure during pregnancy.

Significant positive associations are shown between prenatal $PM_{2.5}$ exposure and three specific miRNAs of the miR-17/92 cluster. For each increase in $PM_{2.5}$ of 5 $\mu g/m^3$ in the second trimester of pregnancy, the expression of miR-17-3p (Figure 5A) increases with 30.56% (95% CI: 4.52% to 63.08%, p=0.02) and the expression of miR-18a-5p (Figure 5E) increases with 49.31% (95% CI: 9.53% to 103.54%, p=0.01). However, miR-19a-3p (Figure 5F) shows only a trend to significance in the second trimester of pregnancy in association with prenatal $PM_{2.5}$ but is significantly proportionally associated with $PM_{2.5}$ exposure for the entire pregnancy period, with a p-value of 0.02. For each increase of 5 $\mu g/m^3$ exposure, the expression of miR-19a-3p will increase with 92.47% (95% CI: 12.14% to 230.33%). The whole pregnancy period is also significantly associated with $PM_{2.5}$ exposure for miR-17-3p, which shows an increase in expression of 44.27% (95% CI: 5.53% to 97.23%, p=0.02) for each increment of 5 $\mu g/m^3$ $PM_{2.5}$ exposure. miR-18a-5P only shows a borderline significant association for the whole pregnancy period. Other cluster members, such as miR-17-5p (Figure 5B), miR-20a-5p (Figure 5C), miR-19b-3p (Figure 5G) and miR-92a-3p (Figure 5D) are not found to be significantly associated with $PM_{2.5}$, nor NO_2 exposure.

Some members of the miR-17/92 cluster are also associated with NO_2 exposure during pregnancy in a positive direction. Both for miR-17-3p (Figure 5A) and miR-18a-5p (Figure 5E), significant correlations were found in the first trimester of pregnancy as well as the whole pregnancy period. In trimester 1, both miRNAs increase significantly (p=0.03) respectively with 12.73% (95% CI: 1.28% to 25.46%) and 18.48% (95% CI: 1.57% to 38.20%) per 5 $\mu g/m^3$ increment of NO_2 exposure. Over the whole pregnancy, the relative expression of miR-17-3p rises with 13.72% (95% CI: 0.39% to 28.33%) and miR-18a-5p shows an increment of 19.65% (95% CI: 0.57% to 42.35%), both with a p-value of 0.04, for an increase of 5 $\mu g/m^3$ in NO_2 exposure. Moreover, borderline significant associations are observed for the third trimester of pregnancy and miRNAs, miR-17-3p, miR-18a-5p and miR-19a-3p (Figure 5F), as well as trimester 2 with miR-18a-5p.

The associations between prenatal BC exposure and the expression of the miR-17/92 cluster are also determined. For each increment of 0.5 $\mu g/m^3$ in BC exposure the relative expression of miR-17-3p (Figure 5A) will gain with 18.09% (95% CI: 2.74% to 35.72%, p= 0.02) and 24.94% (95% CI: 1.31% to 54.09%, p=0.04), during respectively trimester 1 and the whole pregnancy period. In addition, miR-18a-5p (Figure 5E) is also significant for trimester 1 and the entire pregnancy period with an increase of 32.49% (95% CI: 9.06% to 60.97%, p=0.005) and 43.14% (95% CI: 7.50% to 90.60%, p=0.02) respectively for an increment of 0.5 $\mu g/m^3$ in BC exposure. Moreover, a trend to significance is observed in the first trimester of pregnancy for both miR-17-5p (Figure 5B) and miR-20a-5p (Figure 5C) as well as trimester 2 and the whole pregnancy period for miR-18a-5p (Figure 5E) and miR-19a-3p (Figure 5F).

The paralogue cluster miR-106a-363 is significantly associated with NO_2 exposure during pregnancy. This effect is seen for miR-363-3p (Appendix 5, Figure 10C) in all trimester-specific windows as well as the entire pregnancy period. In Appendix 5, Table 22 and Table 23, the unadjusted models for respectively the miR-106a-363 and miR-106b-25 cluster are given. The estimated changes are given for each increase in $PM_{2.5}$ and NO_2 exposure with 5 $\mu g/m^3$ and the increment of 0.5 $\mu g/m^3$ in exposure with the corresponding CI and p-value. With each increase of 5 $\mu g/m^3$ in NO_2 exposure, the relative expression of miR-363-3p will increase with 20.53% (95%CI: 2.57% to 42.35%, p= 0.02), 21.31% (95% CI: 0.76% to 46.04%, p= 0.04), 20.17% (95% CI: 0.33% to 43.93%, p= 0.05) and 21.26% (95% CI: 0.87% to 45.53%, p=0.04), for the three trimester-specific time windows (trimester 1, 2 and 3) and the entire pregnancy respectively. Prenatal NO_2 exposure is also borderline significantly associated in all trimester-specific time windows and the entire pregnancy period for miR-20b-5p (Appendix 5, Figure 10B). Moreover, for each raise of 0.5 $\mu g/m^3$ in BC exposure, a significant increase is observed for the miR-106a-363 cluster in trimester 1 of pregnancy. miR-18b-5p (Appendix 5, Figure 10A), miR-20b-5p (Appendix 5, Figure 10B) and miR-363-3P (Appendix 5, Figure 10C) increased significantly with respectively 19.67% (95% CI: 0.24% to 42.87%, p= 0.05), 23.96% (95% CI: 1.86% to 50.86%, p= 0.03) and 29.64% (95% CI: 4.74% to 60.46%, p= 0.02).

Furthermore, the whole pregnancy period is shown to have a trend to significance for the association between BC and the relative expression of both miR-20b-5p and miR-363-3p. The second paralogue cluster (miR-106b-25, Appendix 5, Figure 11), shows no significant association for none of the exposures. However, miR-106b-5p (Appendix 5, Figure 11A) is borderline significant for trimester 2 and the entire pregnancy specific $PM_{2.5}$ exposure.

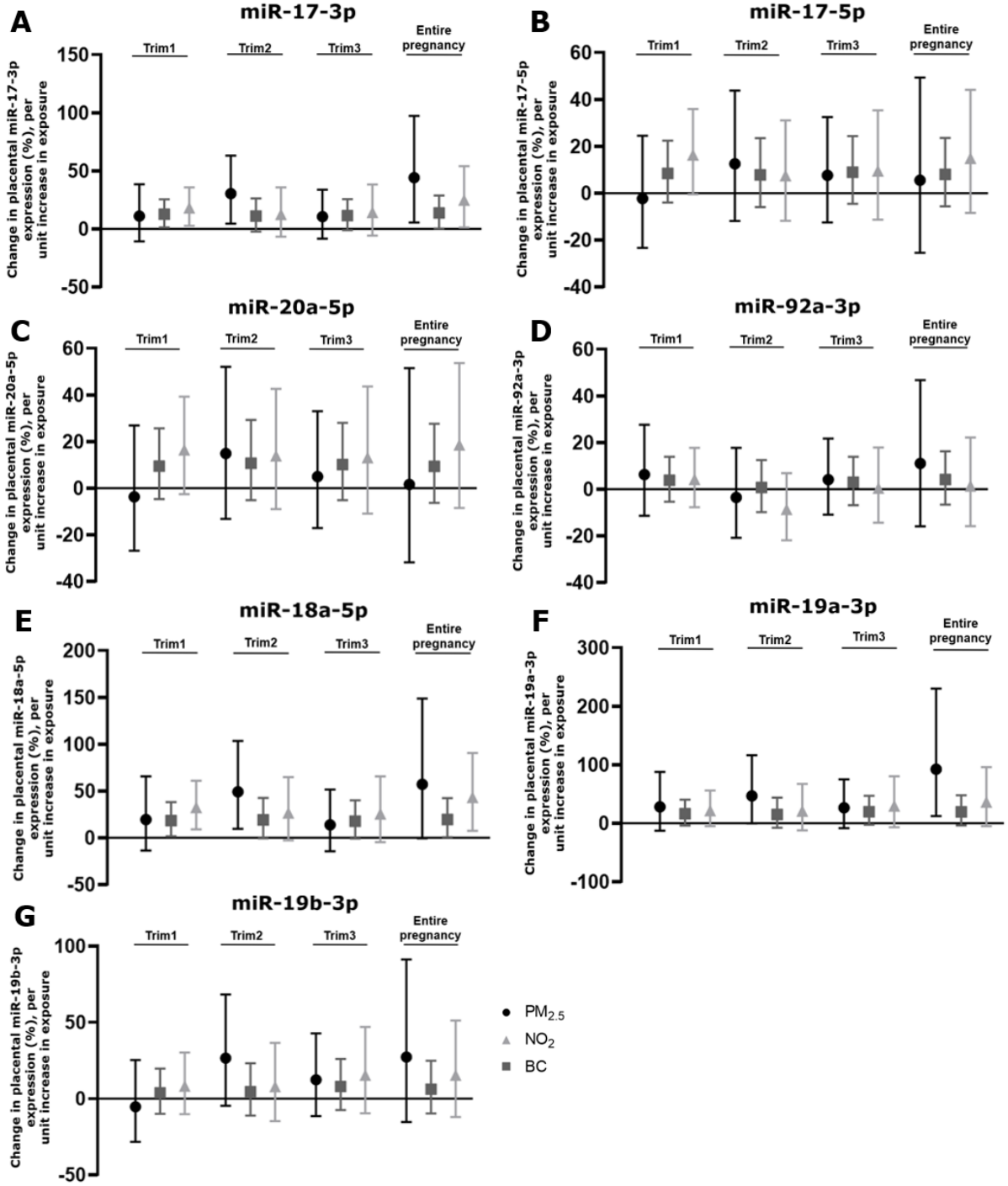

Figure 5. Estimated change (%) in relative placental miRNA expression for each increase of 5 µg/m³ of $PM_{2.5}$ and NO_2 or 0.5 µg/m³ increment in BC for the different time windows of pregnancy of the members of the miR-17/92 cluster, including (A) miR-17-3p, (B) miR-17-5p, (C) miR-20a-5p, (D) miR-92a-3p, (E) miR-18a-5p, (F) miR-19a-3p and (G) miR-19b-3p. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is adjusted for some covariates: date of delivery, maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

3.5 Cord blood gene-expression of both up- and downstream targets of the miR-17/92 cluster in association with air pollution exposure

3.5.1 Cord blood gene-expression of the upstream targets of the miR-17/92 cluster network, including TP53, STAT3, E2F1 and MYC, in association with air pollution exposure

In Figure 6, the estimated change (%), in the relative expression of upstream regulatory molecules of the miR-17/92 cluster, are shown for each increase of 5 $\mu g/m^3$ in $PM_{2.5}$ and NO_2 and an increase of 0.5 $\mu g/m^3$ in BC. These molecules are TP53, E2F1, C-MYC and STAT3 and they regulate the transcription of the miR-17/92 cluster. This model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m^2), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's gender, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy. In Appendix 6, Table 24, the estimated change is given for TP53, E2F1, C-MYC and STAT3 for the unadjusted model, with the corresponding 95% confidence interval and p-value.

Neither before (Appendix 6, Table 24) nor after correction TP53 (Figure 6A), significant associations are found for none of the time windows and none of the exposures.

The gene expression of E2F1 (Figure 6B) in cord blood is negatively associated with air pollution exposure. More specifically, the 2nd and 3rd trimester-specific estimated changes have respectively a decrease of 12.03% (95% CI, -21.26% to -1.72%) and 14.99% (95% CI, -24.18% to -4.68%), as well as the entire pregnancy period with a reduction of 25.51% of the expression (95% CI, -37.10% to -11.79%) for each increment of 5 $\mu g/m^3$ in $PM_{2.5}$ exposure. The associations with NO_2 exposure were also determined and show the same trend with significant inverse associations for the three individual pregnancy trimesters and the entire pregnancy period. Moreover, trimester-specific BC exposure windows during pregnancy show significant associations with E2F1 expression in cord blood with the first and second trimester as well as the entire pregnancy period for each increment of 0.5 $\mu g/m^3$, which shows similar patterns as $PM_{2.5}$ exposure.

Taken into account the trimester-specific exposure to air pollution, MYC expression in cord blood (Figure 6C) was only inversely associated (p-value: 0.003) with BC. In the first trimester, the expression of MYC decreases with 7.38% (95% CI, -12.00% to -2.51%) with each increment of 0.5 $\mu g/m^3$ in BC exposure. Moreover, the whole pregnancy period is borderline significant (p=0.06) with an estimated decrease of 12.76% (95% CI, -24.52% to 0.83%). No significance can be observed with prenatal $PM_{2.5}$ and NO_2 exposure for the trimester-specific exposure windows as well as the entire pregnancy. However, trimester 1 of pregnancy shows a trend to significance (-3;80%, 95% CI, -7.52% to 0.06%, p=0.05) .

Prenatal NO_2 and BC exposure show no significant associations with STAT3 expression in cord blood (Figure 6D). However, trimester 2 of pregnancy (10.40%, 95% CI: 2.89% to 18.45%, p=0.006) and the whole pregnancy (11.56%, 95% CI: 0.08% to 24.13%, p=0.05) is positively associated with $PM_{2.5}$ exposure for each increase of 5 $\mu g/m^3$ in exposure.

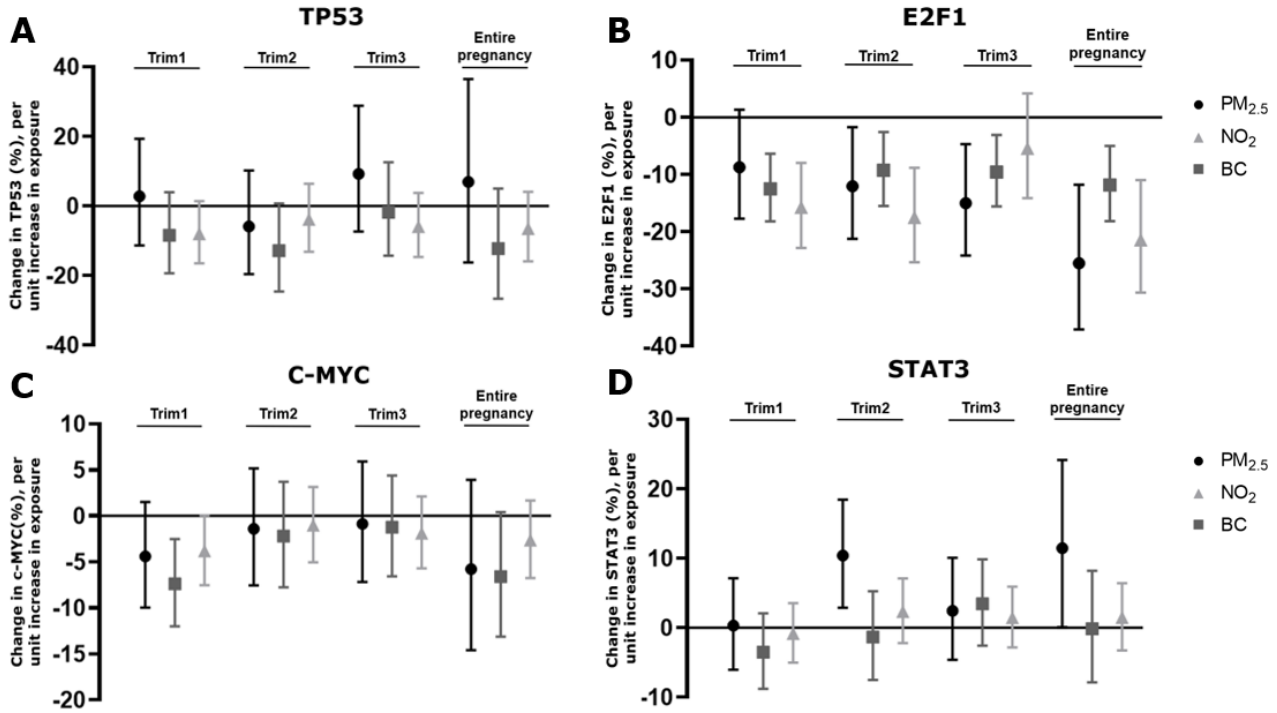

Figure 6. Estimated change (%) in relative gene expression of the upstream regulatory molecules in cord blood for each increase of 5 µg/m³ of $PM_{2.5}$ and NO_2 or 0.5 µg/m³ increment in BC for the different time windows of pregnancy. The effector molecules (A) TP53, (B) E2F1, (C) C-MYC and (D) STAT3 targets the expression of the miR-17/92 cluster. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), new-born's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

3.5.2 Cord blood gene expression of the downstream targets of the miR-17/92 cluster network, involved in cell proliferation (TGFBRII and MAPK14), in association with air pollution exposure

In Figure 7, the estimated change (%) in the relative expression of TGFBRII (Figure 7A) and MAPK14 (Figure 7B) is presented for an increase of 5 µg/m³ in $PM_{2.5}$ and NO_2 and an increase of 0.5 µg/m³ in BC. The model used to reproduce these data is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m²), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's gender, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy. In appendix 7, Table 25 the data of the unadjusted model for TGFBRII and MAPK14 are represented as the estimated change for each increment of 5 µg/m³ in $PM_{2.5}$ and NO_2 and an increase of 0.5 µg/m³ in BC, the 95% confidence interval and the p-value. Both genes, TGFBRII and MAPK14, are targeted by the miR-17/92 cluster and are involved in cell proliferation.

The expression of TGFBRII (Figure 7A) is negatively associated with $PM_{2.5}$, NO_2 and BC. The relative expression of TGFBRII decreases with 11.65% (95% CI: -19.94% to -2.51%) in association with trimester 1 specific exposure and in the third trimester, a decrease of -14.08% (95% CI: -22.81% to -4.36%) is seen related to $PM_{2.5}$ exposure. Moreover, the entire pregnancy period shows a significant decrease of 21.14% (95% CI: -32.84% to -7.41%) for each increase of 5 µg/m³ of $PM_{2.5}$ exposure. NO_2 and BC exposure in relation to the relative TGFBRII expression shows similar results.

Also the relative expression of MAPK14 (Figure 7B) is negatively correlated with the different trimester-specific $PM_{2.5}$ exposure, with significant results for trimester 2 (-21.23%, 95% CI: -32.13% to -8.58%, p=0.002) and the whole pregnancy (-27.38%, 95% CI: -19.47 to -9.60%, p=0.004). Prenatal exposure to NO_2 and BC appears to have a similar negative association with significance in trimester 1 and 2 as well the entire period of pregnancy.

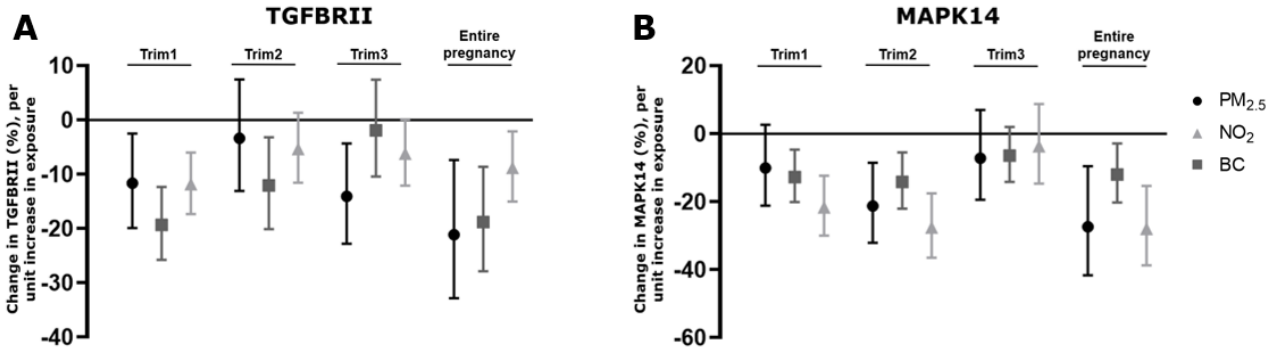

Figure 7. Estimated change (%) in relative gene expression in cord blood of the downstream (TGFBRII and MAPK14) molecules for each increase of 5 µg/m³ of $PM_{2.5}$ and NO_2 or 0.5 µg/m³ increment in BC for the different time windows of pregnancy. The downstream molecules (A) TGFBRII and (B) MAPK14 are targeted by the miR-17/92 cluster and are involved in cell proliferation. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), new-born's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

3.5.3 Cord blood gene expression of the downstream targets of the miR-17/92 cluster, involved in hypoxia and cell cycle arrest (HIF-1α and CDKN1A), in association with air pollution exposure

An adjusted and unadjusted model is used to determine the associations of the downstream targets, HIF-1α and CDKN1A, of the miR-17/92 cluster. These targets are involved in hypoxia and cell cycle arrest. Appendix 8, Table 26, represents the estimated change, of relative gene expression of both targets, with its corresponding 95% CI and p-value for the unadjusted model. The estimates are calculated for each increase of $PM_{2.5}$ and NO_2 exposure with 5 µg/m³ or BC exposure with 0.5 µg/m³. The adjusted model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m²), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's gender, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy and is presented in Figure 8. The estimated percentage change in the relative gene expression of both HIF-1α and CDKN1A and the 95% CI given for an increase in exposure, as previously mentioned.

HIF-1α shows a decrease in the gene expression in association with prenatal exposure to $PM_{2.5}$, NO_2 and BC (Figure 8A). In trimester 2 of pregnancy, a significant decrease of 17.84% (95% CI: -26.13% to -8.62%) can be observed in relation to $PM_{2.5}$ exposure. The whole duration of pregnancy is also found to be significant (-23.15%, 95% CI: -34.73% to -9.51%, p=0.002). Moreover, In trimester 3 of pregnancy, a trend to significance is seen, in which the gene expression of HIF-1α decreases by 10.45% (95% CI: -19.83% to 0.03%, p=0.05). This reduced gene expression is seen per increase in $PM_{2.5}$ exposure with 5 µg/m³. Similar associations can be observed for NO_2 and BC perinatal exposure. However, only the first trimester of pregnancy is significant in association with BC exposure.

The trimester-specific and the overall estimated change in CDKN1A gene expression (Figure 8B) is negatively correlated with the different time-windows of exposure to $PM_{2.5}$, NO_2 and BC. However, no significant associations can be observed. Although, in trimester 3 of pregnancy, BC exposure is borderline significant with CDKN1A. Here, a decrease of 5.60% (95% CI: -11.39% to 0.56%, p=0.07) can be observed, with each increase of 0.5 $\mu g/m^3$ in BC exposure.

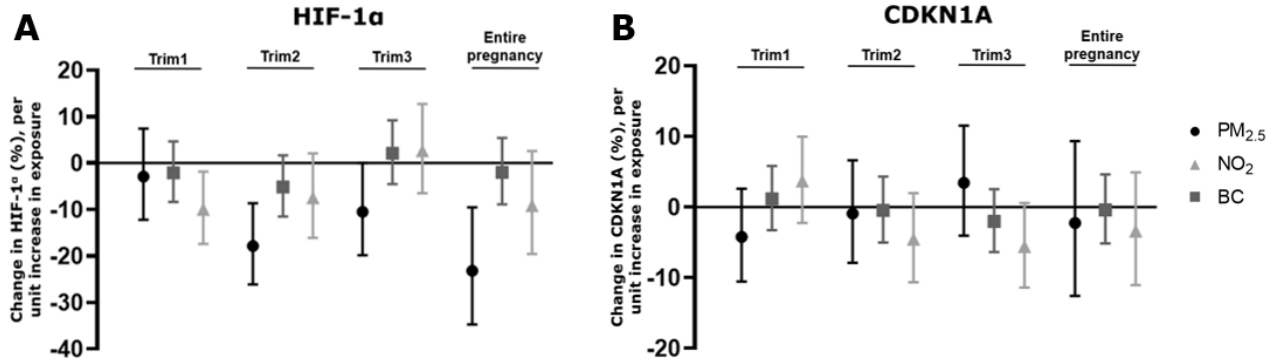

Figure 8. Estimated change (%) in relative gene expression in cord blood of the downstream (HIF-1α and CDKN1A) molecules for each increase of 5 $\mu g/m^3$ of $PM_{2.5}$ and NO_2 or 0.5 $\mu g/m^3$ increment in BC for the different time windows of pregnancy. The downstream molecules (A) HIF-1α and (B) CDKN1A are targeted by the miR-17/92 cluster and are involved respectively in hypoxia and cell cycle arrest. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), new-born's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

3.5.4 Cord blood gene expression of the downstream targets of the miR-17/92 cluster, involved in cell death (BCL2L11 and PTEN), in association with air pollution exposure

Figure 9 represents the estimated change (%) in the relative expression of (A) BCL2L11 and (B) PTEN, which are both downstream targets of the miR-17/92 cluster and are involved in the process of cell death. The estimated change is shown for each increase of 5 $\mu g/m^3$ in $PM_{2.5}$ and NO_2 and an increase of 0.5 $\mu g/m^3$ in BC. This model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m^2), parity (1,2 or ≥3), maternal educational status (low, middle or high), smoking status (never-smoked, past-smoker or current-smoker), newborn's gestational age (weeks), newborn's gender, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy. Appendix 9, Table 27 shows the estimated change for each increase of 5 $\mu g/m^3$ in $PM_{2.5}$ and NO_2 and an increase of 0.5 $\mu g/m^3$ in BC of the unadjusted model. The 95% confidence interval and p-value are also represented.

Taken into account the trimester-specific $PM_{2.5}$ exposure windows during pregnancy, the relative expression of BCL2L11 (Figure 9A) has a reverse association in trimester 1 versus trimester 3. The estimated change of the relative BCL2L11 expression is negatively associated (-15.11%, 95% CI: -24.97% to -3.96%, p=0.0095) with $PM_{2.5}$ exposure during the first trimester of pregnancy and borderline significant (-12.06%, 95% CI: -23.68% to 1.33%, P=0.08) in trimester 2 of pregnancy. However, $PM_{2.5}$ exposure during the 3rd trimester is positively associated with an increase of 15.50% (95% CI: 0.53% to 32.72%) in BCL2L11 expression with each increment 5 $\mu g/m^3$ in exposure. NO_2 and BC show, respectively, significant inverse associations in the first trimester of pregnancy and in the second trimester of pregnancy. Moreover, a trend to significance (p=0.07) is seen for the whole pregnancy period in association with BC exposure.

The relative expression of PTEN (Figure 9B) is negatively correlated with trimester 2 specific exposure of BC. For each increment of 0.5 µg/m^3 in BC exposure, the relative expression of PTEN decreases with 9.97% (95% CI: -18.69% to -0.32%) in de 2nd trimester. The associations with NO_2 and $PM_{2.5}$ are also observed. However, only borderline significance (p=0.1) can be observed in association with $PM_{2.5}$ in the 3rd trimester of pregnancy. Although, this association is positive. For each increase of 5 µg/m^3 in $PM_{2.5}$, the relative expression of PTEN will increase with 10.16% (95% CI: -1.79% to 23.56%).

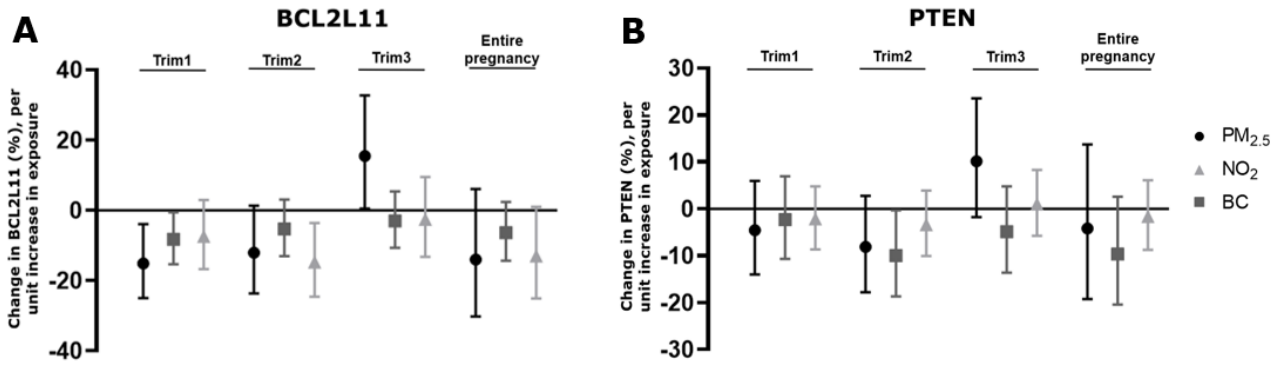

Figure 9. Estimated change (%) in relative gene expression in cord blood of the downstream (BCL2L11 and PTEN) molecules for each increase of 5 µg/m^3 of $PM_{2.5}$ and NO_2 or 0.5 µg/m^3 increment in BC for the different time windows of pregnancy. The downstream molecules (A) BCL2L11 and (B) PTEN are targeted by the miR-17/92 cluster and are involved in cell death. The changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is corrected for date of delivery and adjusted for some covariates: maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

4 Results – Whole miRNA microarray data in the placenta in association with telomere length

4.1 Characteristics of the study population

In Table 9, detailed information is given about demographic and lifestyle characteristics, both of mother/father and child. However, data are given for 2 subpopulations. In Figure 3, this part of the research is situated in panel B. To determine the association between placental miRNA expression (Figure 3, Panel B, part (2)) and TL at birth, we used a population of 60 individuals, which were further divided into 2 groups of 30 individuals. One group had long telomeres and the other group short telomeres. Within these two groups, the placental miRNA expression is measured of all known human miRNAs using miRNA microarray analysis.

In Table 8, the residential population-specific exposure characteristics are given for the different exposures ($PM_{2.5}$, NO_2 and BC) for each of the time-windows of pregnancy (µg/m^3). The $PM_{2.5}$ exposure, in the group of individuals with short telomeres at birth, ranged from 11.5 µg/m^3 (±4.1) in the third trimester of pregnancy and 12.6 µg/m^3 (±5.3) in the first trimester. The IQR was respectively 2.7 and 5.3. The residential exposure of BC during the different time-windows of pregnancy is 1.1 µg/m^3 in all trimester-specific time windows as well as the entire pregnancy period with an IQR of 0.5, 0.7 and 0.5 for the three trimesters consecutive and 0.4 for the entire pregnancy. The residential exposure of NO_2 is also determined for the different exposure periods and is the lowest exposure during the second trimester of pregnancy (15.77 µg/m^3, ±5.39, IQR=5.15) and the highest in the first trimester of pregnancy (16.31 µg/m^3, ±5.44, IQR= 6.97). The exposure characteristics are similar for the population group, including individuals with long TL at birth.

Table 8. Characteristics of air pollution exposure data (µg/m^3) for two study populations. Short telomeres (n=30) represents the population including the individuals selected based on their short relative placental telomere length. Long telomeres (n=30) symbolizes the population including the individuals with longer relative placental telomere length. The expression of the different exposures is given for each time-window of pregnancy (µg/m^3) as the mean (SD). SD= Standard deviation, IQR= Interquartile Range.

	Short telomeres (n=30)		***long telomeres (n=30)***	
Exposure data	***Mean (SD)***	***IQR***	***Mean (SD)***	***IQR***
$PM_{2.5}$				
Trimester 1	12.6 (3.9)	5.3	11.5 (4.0)	4.1
Trimester 2	11.6 (4.2)	6.3	11.2 (3.9)	3.9
Trimester 3	11.5 (4.1)	2.7	10.8 (3.2)	2.2
Entire pregnancy	11.9 (2.5)	4.6	11.1 (2.3)	3.7
Black carbon (BC)				
Trimester 1	1.1 (0.4)	0.5	1.1 (0.5)	0.6
Trimester 2	1.1 (0.4)	0.7	1.1 (0.4)	0.5
Trimester 3	1.1 (0.4)	0.5	1.1 (0.3)	0.4
Entire pregnancy	1.1 (0.3)	0.4	1.0 (0.3)	0.4
Nitrogen dioxide (NO_2)				
Trimester 1	16.3 (5.4)	7.0	15.9 (6.5)	9.2
Trimester 2	15.8 (5.4)	5.2	16.1 (5.7)	8.3
Trimester 3	16.1 (5.1)	6.6	16.0 (5.4)	3.5
Entire pregnancy	16.1 (4.3)	5.5	15.0 (5.0)	7.2

Table 9. Demographic characteristics of the study populations. The study population consisted of 60 individuals, selected on telomere length and are divided over 2 groups. The first population (short telomeres) includes 30 individuals with short telomeres. The second population (long telomeres) symbolizes individuals with longer telomere length and includes 30 subjects. Categorical data were presented as numbers or percentage and continuous data were presented as mean (±SD).

	Short telomeres (n=30)			***Long telomeres (n=30)***		
Characteristic	***Mean (SD)***	***No.***	***%***	***Mean (SD)***	***No.***	***%***
Parents						
Age Mother (years)	29.5 (3.6)			31.1 (5.2)		
Age Father (years)	31.6 (3.8)			32.9 (4.9)		
Pre-gestational body mass index (kg/m^3)	23.5 (3.4)			24.9 (5.9)		
Educational level (mother)						
Low		1	3.3		10	33.3
Middle		9	30.0		10	33.3
High		20	66.7		10	33.3
Parity (mother)						
1		15	50.0		16	53.3
2		12	40.0		11	36.7
≥3		3	10.0		3	10.0
Smoking status (mother)						
Never		30	100.0		30	100.0
Before pregnancy		0	0		0	0
During pregnancy		0	0		0	0
New-born						
Sex, Male		15	50.0		15	50.0
Ethnicity						
European-Caucasian		30	100.0		30	100.0
Non-European		0	0		0	0
Gestational age (weeks)	39.5 (1.2)			39.1 (1.4)		
Birth weight (g)	3506.2 (475.3)			3300.2(480.3)		
Season of birth						
Winter		9	30.0		14	46.7
Spring		6	20.0		6	20.0
Summer		7	23.3		4	13.3
Autumn		8	26.7		6	20.0
Relative placental telomere length	0.8 (0.1)			1.2 (0.2)		

In Table 9, demographic and lifestyle characteristics are given. Within the group, including individuals with a short telomere length at birth, the mothers were on average 29. 5 years old (±SD) ± 3.6 and had on average a pre-gestational BMI (±SD) of 23.5 ± 3.4 kg/m^2. The fathers were on average 31.6 years old at time of birth. None of the mothers had ever smoked (100 % never-smoked) and 66.7 % of the mothers had a high education grade. Half of the population consist of boys and all children were of European descent. The newborn's gestational age is on average 39.5 weeks (±1.2) and ranged between 38 and 40 weeks. The children had on average a birth weight (±SD) of 3506.2g ± 475.3. Moreover, 50% of the children were their mother's first child and the children were born equally divided over the different seasons of the year. These characteristics were similar for the population with a long TL. However, they have a significant difference in TL. The relative TL (±SD) was 0.8 (±0.1) Kb in the placenta of children with short telomeres and 1.2 (±0.2) Kb in the placenta of children with a long TL.

4.2 Placental miRNA expression in association with telomere length

In previous research, 18847 different genes were measured in the placenta of 200 individuals using mRNA microarray analysis in both boys and girls. In newborn boys, 1106 genes were found to be significantly associated with the placental TL at birth. In newborn girls, 543 genes were found to be significantly associated with TL. However, after FDR correction no significant results were found. This part of the research is found in Figure 3 in panel B part (1). These mRNA array data were already available and were used in the current research to select a miRNA subset. The miRNA subset is selected from the whole microRNA microarray dataset (situated in Figure 3 panel B part (2)), which were already available within the research group. Therefore, 19 GO terms were used, which were linked to telomeres and are shown in Appendix 1, Table 12

In appendix 2, Table 13(boys) and Table 14(girls) the top 50 of most significant genes of the whole mRNAarray dataset are given, for boys and girls separately. From the mRNA microarray dataset, a selection was made separately for boys and girls. In Table 10 a selection for both boys and girls is given, including the genes, which were significantly associated with TL at birth and belong to one or more of the selected GO terms. Ten genes were found for the new-born boys, such as excision repair cross-complementation (ERCC1), Exosome component 10 (EXOSC10), F-box protein 10 (FBXO4), Histone cluster 3 (HISTH3), Heterogeneous nuclear ribonucleoprotein F (HNRNPD), Protein phosphatase 1 catalytic subunit beta isozyme (PPP1C), SLX1 structure-specific endonuclease subunit beta isozyme (SLX1A), Structural maintenance of chromosome 5 (SMC5), TERF1-interacting nuclear factor 2 (TINF2), WD repeat containing antisense to TP53 (WRAP53). For the new-born girls, only two genes are belonging to this selection, such as Tuftelin interacting protein 11 (TFIP11) and TOX high mobility group box family member 4 (TOX4).

Based on these subsets of genes (Table 10), both for boys and girls separately, a selection of miRNAs was made. miRNAs were included in the selection if they were known to target one or more genes. In the newborn boys, 552 miRNAs were found to target minimal one of these genes (Appendix 3, Table 15). In the newborn girls, a total of 98 miRNAs was found to target one or both of the genes (Appendix 3, Table 16). Subsequently, the overlap was determined between the miRNA microarray data set and the selected miRNA lists. From the 552 selected miRNAs in the newborn boys, only 104 miRNAs had data available for the miRNA microarray analysis. In the newborn girl population, only 17 miRNAs were found in overlap with the miRNA microarray data.

Subsequently, statistical analysis was performed to determine the associations between placental expression of these miRNAs and telomere length in the placenta at time of birth. In the newborn girl population, none of the 17 miRNAs were found to be significantly associated with placental telomere length (data not shown). In the newborn boys (Table 11), two miRNAs were found to be significantly associated with the placental telomere length and eight miRNAs showed a trend to significance. Placental miRNA expression of hsa-miR-584-5p (β=3,38, SD=±0.95, p=0.002) and hsa-miR-4291 (β=1.51, SD=±0.68, p=0.04) were found to be positively associated with telomere length. Borderline significant association

were shown for hsa-miR-3646, hsa-miR200c-3p, hsa-miR-335-3p, hsa-miR-181d-5p, hsa-miR-1202, hsa-miR-126-5p, hsa-miR-181a-5p and hsa-miR-3125. Most of these miRNAs showed a positive trend. However, hsa-miR-1202 was inversely, but not significantly, associated with telomere length in the placenta. Although, after FDR correction for multiple testing, no significant associations were found.

Table 10. Overlap of genes found to be significantly associated with telomeres and belonging to one or more of the 19 selected GO terms. The Gene symbol, Gene name and the GO terms belonging to the genes were given. 10 genes were found for newborn boys and 2 genes for newborn girls in the overlap between GO terms and significantly associated genes of the placenta mRNA microarray analysis.

Gene Symbol	***Gene Name***	***GO terms***
Boys		
ERCC1	Excision repair cross-complementation group1	GO: 0032204, GO:0000781, GO:0032205, GO:0000784
EXOSC10	Exosome component 10	GO: 0032204, GO:0032211, GO:0032205
FBXO4	F-box protein 4	GO: 0032204
HIST3H3	Histone cluster 3, H3	GO:0000781, GO:0000784, GO:0016233
HNRNPD	Heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	GO: 0032204, GO:0042162, GO:0051973
PPP1CB	Protein phosphatase 1, catalytic subunit, beta isozyme	GO:0000781, GO:0000784
SLX1A	SLX1 structure-specific endonuclease subunit, beta isozyme	GO: 0032204, GO:0032205, GO:0010833
SMC5	Structural maintenance of chromosomes 5	GO:0000781
TINF2	TERF1 (TRF1)-interacting nuclear factor 2	GO: 0032204, GO:0032211, GO:0000781, GO:0032205, GO:0000783, GO:0000784, GO:0042162, GO:0016233, GO:0032202, GO:0070198
WRAP53	WD repeat containing, antisense to TP53	*GO:0007004,* GO:000078, GO:0010833, GO:0032202, GO:0051973
Girls		
TFIP11	Tuftelin interacting protein 11	GO:0000781, GO:0000784, GO:0031848, GO:0016233
TOX4	TOX high mobility group box family member 4	GO:0000781, GO:0000784

Table 11. Adjusted model of the association between placental miRNA expression and placental telomere length in newborn boys. The model is adjusted for date of hybridisation, the father's age (years), maternal education status (low, middle or high), parity (1,2 or ≥3), season of delivery, date of delivery and gestational age (weeks). Using the Benjamini-Hochberg false discovery rate (FDR), the p-value was adjusted for multiple testing, considering a significance level of 5%. For each miRNA, the relative change (β) in expression is given in association with telomere length with the corresponding standard deviation (SD) and p-value. The model is also corrected for multiple testing. The FDR corrected values were given as Q-values.

Adjusted model			
miRNA	**β (±SD)**	**p-value**	**Q-value**
hsa-miR-584-5p	3.38 (±0.95)	**<0.01**	0.26
has-miR-4291	1.51 (±0.68)	**0.04**	0.73
has-miR-3646	2.36 (±1.17)	0.06	0.73
has-miR-200c-3p	1.26 (±0.62)	0.06	0.73
has-miR-335-3p	3.28 (±1.65)	0.06	0.73
has-miR-181d-5p	2.35 (±1.23)	0.08	0.73
has-miR-1202	-3.18 (±1.69)	0.08	0.73
has-miR-126-5p	2.81 (±1.57)	0.09	0.73
has-miR-181a-5p	0.93 (±0.52)	0.09	0.73
has-miR-3125	2.03 (±1.16)	0.1	0.73

5 Discussion

5.1 miR-17/92 network expression in cord blood in association with prenatal air pollution exposure

There is growing evidence that microRNAs could be the link between prenatal air pollution exposure and differences in the expression of genes. These alterations can result in foetal programming and adverse health outcomes later in life (57, 58). Developing children, are more susceptible for these carcinogenic pollutants, because of the presence of rapid proliferative and differentiating cells, the greater capability of absorption and reservation, their undeveloped immune system and the lower ability of detoxification, DNA repair systems or apoptosis (57, 59, 60). Nowadays, more evidence is obtained that cancer development during adult life can originate from epigenetic modifications, such as microRNAs, affected by *in utero* exposure to air pollution (58). However, knowledge of how microRNAs, such as the oncogenic cluster miR-17/92 and their network, will be affected by prenatal air pollution exposure is scarce.

Recently, data from the ENVIR*ON*AGE birth cohort have been shown that miR-17, miR-20a and miR-92a expression in cord blood is negatively associated with black carbon exposure during pregnancy. Moreover, miR-17 and miR-20a were also negatively associated with other exposures, namely NO_2 and $PM_{2.5}$ exposure. In line with these findings, miR-92a is found to be downregulated in MCF-7 cells, a human breast cancer cell line, after exposure to Dichlorodiphenyltrichloroethane. However, also particulate matter has already shown to be related to the expression of miR-17/92 cluster members (61). For example, in the model teleost fish, zebrafish, genome-wide transcriptional analysis revealed the toxicity effects of $PM_{2.5}$ by inducing altered expression of multiple microRNAs including downregulation of two cluster members, miR-19a-3p and miR-19b-3p (62). Moreover, in a population study, the impact of short-term air pollution exposure on the transcriptome and microRNA expression in the blood of healthy volunteers in London and Barcelona was studied. Here, a downregulation of miR-20a and miR-92a is found in association with PM exposure (63). However, little is known about the potential oncogenic mechanisms behind the miR-17/92 cluster.

According to the Barker hypothesis, changes in the *in utero* environment can have disadvantageous consequences in later life, which can be induced by posttranscriptional regulation by microRNAs. One of these consequences is cancer development. It has been shown that the individual's risk increases if the carcinogenic exposure happens early in life. Therefore, the ENVIR*ON*AGE birth cohort is a unique setting to study the associations between environmental exposure early in life and certain molecular targets. Within the ENVIR*ON*AGE birth cohort methylation changes occurring in genes, responding to DNA damage were shown to be associated with air pollution exposure during pregnancy. These alterations, in the expression of tumour suppressor or oncogenes, can be retained throughout lifetime and can eventually result in a predisposition for carcinogenesis (24). This may also be the case for the miR-17/92 cluster, as it regulates many important processes such as cell proliferation, cell cycle regulation, apoptosis, etc. (2, 30, 44). Moreover, the cluster is already found to be downregulated in several cancer types including leukaemia, breast cancer, hepatocellular carcinoma, ovarian cancer, melanomas and retinoblastoma (2, 3, 44, 45). Because of these reasons, the altered expression of the miR-17/92 cluster at birth could contribute to an increased carcinogenic risk in adult life. Therefore, it is important to determine the cluster's role in early life development, in association with prenatal air pollution exposure, and how its network of both the up- and downstream targets is affected by early-life exposure.

The present study found that the mRNA expression of some upstream targets of the miR-17/92 cluster, namely C-MYC and E2F1 were inversely associated with particulate matter, NO_2 and BC exposure during *in utero* life. Both targets have a regulatory function and will induce the activation of the cluster (1-3, 30). Therefore, keeping the observed downregulation of the cluster in mind, we expect a negative association of these targets with air pollution exposure during *in utero* life. These observations are consistent with the

findings of Chen and colleagues. They showed that downregulation of MYC induced downregulation of the miR-17/92 cluster (64). Moreover, Humphreys *et al.* showed, in a cell line derived from colon rectal cancers (HT29 and HCT116) treated with the histone deacetylase inhibitor butyrate, a downregulation of E2F1 in response to the butyrate treatment. The downregulation of E2F1 will subsequently result in a downregulation of the miR-17/92 cluster, which is also observed in this cell population (65). On the other hand, STAT3 is positively correlated with trimester 2 specific $PM_{2.5}$ exposure as well as the exposure during the entire pregnancy period. STAT3 is both an up- and downstream target of the cluster. STAT3 will induce the activation of the miR-17/92 cluster. However, downstream of the cluster, miR-17 and miR-20a will target and downregulate STAT3 expression (1, 3, 30). We assume that there is less inhibition of STAT3 downstream of the cluster in this study in association with prenatal air pollution exposure, which induces an upregulation of this target. In line with our results, Zhang *et al*. found that downregulation of two cluster members, including miR-17-5p and miR-20a, blocks the expression of STAT3 in myeloid-derived suppressor cells (66). Also, the associations with air pollution are consistent with previously performed research. Within the ENVIR*ON*AGE birth cohort, MYC was found to be significantly downregulated in relation to $PM_{2.5}$ long-term exposure. Moreover, in human umbilical vein endothelial cells (HUVEC) treated with $PM_{2.5}$ in a dose-dependent manner, they showed that STAT3 expression was significantly upregulated after exposure to ambient air pollution.

However, the transcription of the upstream targets of the cluster is tightly controlled. Beside the autoregulatory mechanism of E2F1, another positive feedback circuit will regulate the transcription of the miR-17/92 cluster. Within this regulatory network, two upstream targets were found to regulate each other's transcription. It has been shown that C-MYC is involved in the expression of E2F1. Vice versa, it has been shown that E2F1 induces the expression of C-MYC. Both upstream targets are therefore associated with each other. This is also observed in the current research, in which a Spearman correlation of <0.01 is found among the two activators of miR-17/92 (data not shown). This enforces our findings that both E2F1 and MYC are found to be downregulated in association with air pollution, and probably in the same direction (1). Therefore we assume that prenatal air pollution will affect the expression of E2F1. Subsequently, the changes in the E2F1 expression might result in the differences in the expression of MYC and additionally of both the miR-17/92 cluster and the targets within its network.

However, research to confirm this hypothesis is scarce. Nonetheless, as already explained, these changes in expression of the miR-17/92 cluster can result in a predisposition of cancer development later in life. This might be induced by changes in the network around the cluster. Within this research in line with the expectations, a negative association is seen between the expression of MYC and E2F1 in cord blood and prenatal air pollution exposure. Both activators of the cluster are known to be oncogenes, which would mean that a decreased expression of both genes will induce more apoptosis and less cancer proliferation. However, there are some paradoxical studies known in literature. In a mice model knocked-out for E2F1, there has been shown that, later in life of these mice, they developed hyperplasia and neoplasia (67). This demonstrates the tumour suppressor function of the gene and that predisposition, due to a downregulation of E2F1, will result in disease development later in life. Also, MYC has found to be dysregulated in cancer. In literature, it has been shown that MYC is downregulated in cancer cells, which increases cell survival under circumstances of limited energy sources (68). Within our research, this could be a plausible explanation for our findings. We assume that prenatal air pollution exposure could contribute to cancer predisposition later in life, by a downregulation of E2F1, which influence the expression of MYC. Thereby, E2F1 and MYC will downregulate the miR-17/92 cluster and influences its network, which has already been shown to be associated with some cancers.

Downstream of the miR-17/92 cluster, some genes involved in several critical cellular processes will be targeted. According to the literature, it is expected that the processes of cell proliferation, hypoxia, cell cycle arrest and cell death will be affected by air pollution exposure via the regulation of the miR-17/92 cluster (1, 3, 30). This can be indicated by studies performed within the ENVIR*ON*AGE birth cohort. They investigated recently the sex-specific responses in the transcriptome to both long- and short-term exposure to $PM_{2.5}$. In cord blood of mother-newborn pairs of the cohort, the whole genome gene expression was determined and by using overexpression analysis and gene set enrichment analysis, the association between $PM_{2.5}$ exposure and gene expression was analysed. Within this research, they showed that processes, such as DNA damage response and cell cycle regulation, are significantly altered for both sexes in response to long-term as well as short-term exposure to $PM_{2.5}$ (69). The current research indeed shows alteration in these processes of cell proliferation, hypoxia, cell cycle arrest and cell death. According to the literature, it is expected that downstream targets of the cluster will be increased in expression, due to less inhibition, in association with prenatal air pollution exposure. However, an inverse association is seen for BCL2L11, CDKN1A, HIF-1α, MAPK14, PTEN and TGFBRII with early-life air pollution exposure. Nonetheless, this can be explained by interference of other regulatory factors. The network around the miR-17/92 cluster, constructed in this research is incomplete. The miR-17/92 cluster is involved in the regulation of multiple other targets and the cluster itself is also regulated by other factors. Moreover, the chosen downstream targets are involved in frequent processes, which are probably regulated by other factors as well (1, 3, 30). In the present research, only MYC is correlated with the three cluster members of interest, miR-17, miR-20a and miR-92a. This strengthens the hypothesis that other molecules are also responsible for the regulations of the chosen targets. For example, it is already known that the paralogue clusters (miR-106a-363 and miR-106b-25) are involved in the transcriptional activation of the targets, which were observed in this research (3).

The miR-17/92 cluster is involved in the regulation of cell proliferation, by inhibiting TGFBRII and MAPK14 (1, 3, 30). The cluster is already observed to be downregulated in response to air pollution exposure. Therefore, a positive association of TGBFRII and MAPK14 might be plausible in relation to *in utero* air pollution exposure. However, our results indicate an inverse association with prenatal air pollution exposure at time of birth for both TGFBRII and MAPK14. However, we assume that both targets will be monitored by other factors. For example, both TGFBRII and MPK14 are found to be regulated by the paralogue clusters miR-106a-363 and miR-106b-25. More specifically, TGFBRII is targeted by miR-20b and miR-106a and MAPK14 is downregulated by miR-106b, which are respectively part of the miR-106a/363 and miR-106b/25 cluster (3). Moreover, other miRNAs, for example, can also affect the transcription of the genes. In glioblastoma cells, Mucaj and colleagues showed an increase in apoptosis due to the inhibition of MAPK14, which indicate the control of MAPK14 expression by miR-124 (70).

The second two processes investigated in response to air pollution exposure during pregnancy are hypoxia and cell cycle arrest. These processes were regulated by respectively HIF-1α and CDKN1A. We observed a significant negative association for HIF-1α in response to $PM_{2.5}$ exposure for the second trimester of pregnancy and the entire period as well as a borderline significant association for the third trimester. This is not in line with the expectations. However, the regulation of HIF-1α includes also the involvement of the miR-106a/363 cluster (miR-20b and miR-18b). Moreover, it has been shown that miR-22, when overexpressed, will inhibit HIF-1α expression (3, 71). On the other hand, CDKN1A is only borderline significant with a negative association in trimester 3 specific exposure to BC. This association can be explained by interference of the miRNAs miR-106a and miR-106b. Moreover, TRIM71, a target of Let7, is also involved in the regulation of the expression of CDKN1A and results in a downregulation of this target (72).

Within this research, we also observed the effects on BCL2L11 and PTEN expression, involved in cell death, in association with air pollution exposure during pregnancy. This downstream target, BCL2L11, is found to be negatively associated with $PM_{2.5}$ in the first trimester, as well as borderline significant in trimester 2 of pregnancy. PTEN shows only a significant inverse correlation with second-trimester specific exposure to BC. In literature, there has been shown that downregulation of the cluster should result in an increased expression of both PTEN and BCL2L11 (65). This is in contradiction to our results. However, the expression of both targets can also be affected by other factors and miRNAs. For example, the miR-106a-363 cluster, more specifically miR-106a, is also involved in the regulation of the BCL2L11 expression. Moreover, miR-339 was shown to negatively regulate BCL2L11. PTEN expression, on the other hand, was inversely correlated with miR-144. These regulatory miRNAs could contribute to the inverse association, which is seen in the current study.

Both BCL2L11 and PTEN shows a positive association in trimester 3 of pregnancy in relation to $PM_{2.5}$ exposure. This is in contradiction with the associations seen in the other trimesters of pregnancy. Within our research department, similar results were already seen. Tsamou *et al.* found a significant positive association of miR-20a and miR-21 with prenatal PM exposure in the first trimester of pregnancy. However, the same miRNAs were negatively associated with trimester 2 specific exposure (29). Nonetheless, these trimester-specific expression changes are not novel. Both in mice and fish, for example, they found decreased levels of the microRNA miR-9 after exposure to ethanol in early development (73, 74). However, later developmental stages showed an increased expression in mice and adult rats (75, 76). During different moments of pregnancy, it is possible that a similar mechanism could react differently to air pollutions. This because each trimester of pregnancy has its own hallmarks (77).

Another explanation might be that the expression of these genes is different among the different stages of pregnancy, because of the different needs in that stage of development, instead of reacting differently to the environmental exposures. Sitras *et al.* showed a differential expression of 25% of the 7519 tested genes in the placenta, in the first trimester compared with the third trimester of pregnancy. The differences in the expression were due to changes in the structure and function of the placenta throughout pregnancy (78). Moreover, specifically, the process of cell death incidence has already been shown to be differentially present over the different trimesters of pregnancy (79). This might also be the case for the expression patterns in cord blood. The different trimesters of pregnancy will be characterised by other developmental stages. For example, in the first trimester, the neural tube will be formed as well as hands, legs, fingers and toes, which are all characterised by apoptotic features (80). However, the third trimester is mainly characterised by foetal growth. The differences in stages of foetal development might be accompanied by other gene expression patterns.

Beside the trimester-specific expression of genes, it is possible that the direction of the association found in the 3rd trimester of pregnancy, between prenatal air pollution exposure and both BCL2L11 and PTEN, is dependent on other factors. It has already been shown that the gene expression profiles encounter seasonal differences (81). To overcome this problem, we correct for multiple covariates. However, there are still some unknown confounders. For example, population differences, sampling method, circadian rhythm, *etc*. Recently more evidence is obtained that circadian rhythm induces molecular differences in many peripheral organs. Martino *et al*. showed the differences in myocardial expression of 12.488 murine genes. 13% of the measured genes showed significant changes across a cycle of 24 hours. Moreover, genes, which were differentially expressed, were involved in important biological processes, such as cell growth, transcription, translation, *etc* (82).

Within the present study, the association between members of the miR-17/92 cluster and prenatal air pollution exposure was determined in cord blood as well as in the placenta. The expression of the paralogue clusters, miR-106a-363 and miR-106b-25, were also observed for associations with early-life air pollution exposure. It is interesting when observing the molecular effects of *in utero* environmental exposure, to study them both in cord blood and in the placenta. When pregnant females were exposed to air pollutants, ultra-fine particles can enter the bloodstream and will roam throughout the body. It might be plausible that during pregnancy, these particles will cross the placental barrier. In literature, it has already been shown that particles smaller than 240 nm can pass through the human placenta. Furthermore, in mice, particles smaller than 500 nm are even found to enter the foetal circulation. However, it might be possible that the transfer of some pollutants is blocked and will not enter the foetal compartment. These pollutants will not show an effect in cord blood as it is never got there. However, indirectly they can have effects on the foetal development by interfering with the nutrient exchange and the development of the vascular network. For example, PM exposure may be cause complex vascular modifications in the second and third trimester of pregnancy and this can lead to placental abnormalities, which will affect the foetus. Meanwhile, we assume that the molecular effects seen in cord blood will give a good impression of the direct effect on the foetus. The cord is a direct link between the placenta and the foetal circulation and is important for the exchange of oxygen and nutrients to the foetus during pregnancy. However, also environmental pollutants can enter the foetal circulation via this way. Therefore the effects observed in cord blood might give a good impression about the effects in de foetal circulation and therefore directly on the foetus. Another advantage is that we will have both an impression on circulation markers and tissue-specific markers, respectively observed in cord blood and the placenta. This is because the circulation of cord blood is separated, it may not be representative for specific tissues. Another explanation could be the differences between short-term and long-term exposure. Cord blood is a very dynamic system with a continuous exchange, for example for nutrients between mother and child. Therefore, we assume that cord blood will reveal the effects of short-term environmental pollution. However, the placenta is present during the entire pregnancy and develops throughout this period. Therefore, we assume that the cumulative effects of air pollution exposure or long-term effects, might rather be visible in the placenta (4, 83).

In contradiction with the association in cord blood, the placental expression of members of the miR-17/92 and miR-106a-363 clusters are found to be positively associated with prenatal air pollution exposure. Like previously explained, the miR-17/92 cluster has already been found to be related to pollution, more specifically DDT. However, in this study, they showed an inverse association. Meanwhile, the association between environmental pollution and the miR-106a-363 has never been shown, to our knowledge. However, both the miR-17/92 and miR-106a-363 cluster have been shown to be associated with gastric cancer. The miR-17/92 cluster's expression is significantly higher in the mucosa of patients suffering from gastric cancer (84). Moreover, miR-106a-363 knock-down has been shown to be protective against carcinogenic processes via upregulation of MBP-1, which is important for the transition from G1 to the S-phase during cell cycle progression (85). Consequently, upregulation of both clusters is associated with gastric cancer risk. On the other hand, air pollution more specifically $PM_{2.5}$, PM_{10} and NO_2 have been related to an increase in cancer risk (86). Therefore, we assume that prenatal air pollution exposure will predispose children to develop cancer later in life by interfering with the molecular mechanisms around the miR-17/92 and miR-106a-363 cluster, which will explain our findings. However, these associations in the placenta are still contradictory to the findings shown in cord blood. Nonetheless, it is possible that the expression patterns and its reaction, induced by pollutants, are cell-specific. The cell composition of placental tissue is completely different from cord blood. The placenta is, for example, composed of a complex mixture of cell types and the cell-specific effects of miRNA expression could be influenced by the placental composition (83).

The current research has to deal with some limitation and results should be interpreted taking these into account. First, the observed alterations in the gene expression in cord blood are only indirect evidence of the effects on the child's health and the child's risk for disease development in adult life. A second limitation is that we have only measured the mRNA expression of the target genes and miRNAs at time of birth. Therefore, we cannot observe the temporal changes in the regulation of the miRNAs of the miR-17/92, miR-106a-363 and miR-106b-25 clusters and the miR-17/92 cluster network. Within this research, we measured the miRNA expression levels of the miR-17/92 cluster and the gene expression levels of its up- and downstream targets in plasma of cord blood. Moreover, the expression of the members of the miR-17/92, miR-106a-363 and miR-106b-25 cluster are measured in the placenta. However, the exact cellular content of these samples, both placental tissue and cord blood, is not known. Another limitation can be attributed to the fact that we estimated the exposures based on the maternal address of the participants. The contrast of exposure is relatively narrow, for example, an IQR of 2.85 for $PM_{2.5}$ exposure in the entire pregnancy period. However, this narrow exposure contrast gives the opportunity to investigate the effects of a realistic, daily exposure in Belgium. A limitation of the measurements in the placenta (Figure 3, Panel A, part (3)) is that we have a low sample size, only 60 individuals. Due to the low power, other associations could possibly stay unrevealed. The last limitation within this study is the inability of the study to correct for residual confounding by unknown factors, which could interfere with the miRNA expression, both in cord blood and placental tissue, the mRNA expression of the targets in our network and with air pollution.

5.2 Whole microarray data in the placenta in association with telomere

Telomere length reflects the cellular turnover of the individual and is found to be a biomarker of biological ageing. From birth onwards, there is a large variability of the telomere length among individuals. In an animal-based study in zebra-finches, they showed that despite the same age, the TL in the animals was very different. Therefore, not only age but also other factors should contribute to the variability in TL (10). This variability can partly be explained by genetic factors. However, more evidence is obtained that environmental factors could accelerate telomere attrition (38). Epigenetic mechanisms could play a role in the process of ageing. Therefore, miRNAs could have a critical function in the regulation of ageing and telomere functioning. We hypothesised that in the placenta of newborns, some miRNAs will be differentially expressed in association with telomere length.

Recently, within ENVIR*ON*AGE birth cohort, they investigated the association of all known miRNAs and telomere length in the placenta of newborns. However, no significant results were found after correction for multiple testing (data not published). To overcome this problem, the input list of miRNAs will be reduced. A selection was made based on the link of their target genes with several GO terms that have to done with telomeres. In the present study, the relative placental telomere length was found to be significantly associated with two specific miRNAs, namely hsa-miR-584-5p and hsa-miR-4291, in newborn boys. Moreover, eight other miRNAs were found to be borderline significant. However, in newborn girls, no significant neither borderline significant associations could be found. In line with our findings, Bilsland *et al.* found that miR-584 is involved in the process of cell senescence, which is known to be related to telomere shortening (87).

From birth onwards, telomere length is highly variable with longer telomeres seen in girls compared with boys. This is due to the activity of telomerase and the oestrogen-induced inhibition of oxidative stress, which cause inhibition of telomere shortening (88). However, the variability in TL from birth onwards will have consequences later in life. At birth, the cellular capacity of individuals with longer telomeres is higher to manage processes influencing telomere shortening, such as oxidative stress and inflammation (38). Telomere shortening is involved in "normal" ageing. However, accelerated telomere attrition is also seen in various diseases, such as cancer, due to its role in cell homeostasis (38). Throughout the years, more evidence is obtained that shorter telomere length is associated with higher cancer risk. However, the exact mechanism behind this association is not known. However, miRNAs could be involved (89). miR-584 for example, which is found to be associated with TL in the present research, has a tumour-suppressive

function and is involved in cell proliferation in many cancer types. During cell proliferation, telomeres will shorten gradually and this process is important in many cancer types. In non-small cell lung cancer (NSCLC) miR-584 is found to be downregulated, which will inhibit invasion and proliferation of the lung cancer cells (90). Furthermore, Sun *et al* showed that shorter telomeres were associated with a higher risk of developing lung cancer (91). Moreover, miR-584 is found to be upregulated in melanoma cells after treatment with metformin, which is an antihyperglycemic agent used in diabetic patients to improve glucose intolerance. Metformin inhibits cell growth by increasing miRNA expression, including miR-584, which will inhibit cell motility and induce apoptosis. To summarise, overexpression of miR-584 will inhibit cancer progression (90). miR-4291, which is also found to be positively associated with telomere length in the current research. Its function is not known yet, because this miRNA is not discovered for a long time, which is shown by the high number. miR-4291 is also found to be dysregulated in melanoma patients. Sand *et al* showed, in cutaneous malignant melanoma cell lines, that miR-4291 was significantly downregulated (92). Longer telomeres, associated with a higher expression of both miRNAs will be associated with a higher risk of developing cancer later in life, which is in line with our results (89). Contradictory, in melanoma cells, it has been shown that shorter telomeres will have a protective role to carcinogenesis (89). But overall, cancer risk will increase with shorter telomeres. However, it might be that TL is not a good parameter to assess cancer risk. The reason for this is that TL only estimates the plenty of telomere sequences in one cell. However, there is no information about the distribution over other chromosomal ends (91).

TL is not only associated with cancer, but also other disorders such as preeclampsia (PE) have been shown to be related to telomere length. PE is a vascular pathology, which occurs in two to ten percent of the pregnancies and is characterised by maternal hypertension and proteinuria. PE complications are: diminished implantation, endothelial dysfunction and systemic inflammation. Due to hypoxic/oxidative stress and senescence, telomeres will be significantly shorter in the placenta of preeclamptic pregnancies (93). It is hypothesised that miRNAs are involved in the pathophysiological processes of PE development, because of their role in several critical cell functions such as proliferation, apoptosis *etc*. Also, miR-584, which has been shown to be significantly associated with telomere length in the present study, is found to be significantly downregulated in the placenta of preeclamptic pregnancies (94). These findings are in line with the associations found in the present study. To summarize, we assume that a higher placental miR-584 expression, which is related with longer telomere length in the placenta, might contribute to a higher risk for PE development during pregnancy, which is characterised by shorter telomeres.

The current study has to deal with some limitations. First of all, the sample size was relatively small. The expression of all known human miRNAs is only measured in the placenta of 60 individuals. This is rather small, which is disadvantageous for the power of the study. A low power can ensure that small associations will be uncovered. However, by reducing the number of miRNAs in the analysis, the power will increase. A second disadvantage is found in the selection of the sample tissue. Placental tissue is composed of a complex mixture of cells and the cell composition can be variable over the different places in the placenta. Therefore, the results can be driven by a certain cell type in the placenta. Meanwhile, the placenta will give the chronical epigenetic effects of air pollution. Another limitation is the method by which TL is determined. We used a real-time PCR method, which is known to have a higher assay variability compared with other methods. However, the variation coefficients were smaller than 7%.

6 Conclusion

In the first panel (Figure 3, panel A) of this research, the aim was to investigate the expression of the miR-17/92 cluster and its network of both up- and downstream targets known to be involved in cancer initiation in association with *in utero* air pollution exposure in cord blood. As well as the expression of the oncogenic cluster miR-17/92 and its paralogue clusters miR-106a-363 and miR-106b-25 in the placenta. To determine these associations, cord blood and placental samples were collected at time of birth within the unique setting of the longitudinal population-based ENVIR*ON*AGE birth cohort.

In order to determine the association between prenatal air pollution exposure and the expression of the oncogenic cluster miR-17/92 in cord blood and placental tissue (Figure 3 , panel A, part (1) and (3)), the relative expression of the miR-17/92 cluster was determined in cord blood by RT-PCR and by miRNA microarray analysis in placental tissue. This might give an impression about the effects of respectively short-term and long-term prenatal exposure to environmental air pollution, as well as the differences between the effects seen in the circulation and tissue-specific effects. Our results show the association of early-life air pollution exposure and three specific miRNAs (miR-17, miR-20a and miR-92a), in cord blood and three miRNAs (miR-17-3p, miR-18a-5p and miR-19a-3p), in placental tissue. However, in these two tissues, the association for the miR-17/92 cluster and air pollution was paradoxical and cell-specific. In literature, the miR-17/92 cluster have been found to be related to cancer development. Therefore, it is important to determine how the network of both activators and targets will be affected by *in utero* exposure.

The mRNA expression of both up- and downstream targets, known to be involved in cancer initiation, of three specific miR-17/92 cluster members (miR-17, miR-20a and miR-92a) was determined with RT-PCR in cord blood. This network included four upstream targets, such as E2F1, MYC, STAT3 and TP53 and six downstream targets, including BCL2L11, CDKN1A, HIF-1α, MAPK14, PTEN and TGFBRII. Within this network, an autoregulatory feedback loop exists, in which both E2F1 and STAT3 were activators, but also downstream targets by the cluster. Results showed that upstream of the cluster, E2F1 and MYC were downregulated in association with early life air pollution exposure, which induces less transcription of the cluster. In addition, STAT3 was found to be positively associated with air pollution exposure during pregnancy, which is explained by less inhibition downstream of the cluster. Moreover, downstream targets BCL2L11, HIF-1α, MAPK14, PTEN and TGFBRII were determined to be downregulated in association to *in utero* air pollution exposure. However, this is in contradiction to the expectations in which a downregulation of the miR-17/92 cluster will induce an increased expression of its downstream targets.

It can be concluded that in association with early life prenatal air pollution exposure, the cluster miR-17/92 and its network is differentially expressed in cord blood, which confirms the original hypothesis. A downregulation of the cluster as well as its upstream activators E2F1 and MYC is observed in association with *in utero* exposure. These data suggest that prenatal air pollution exposure might downregulate the activators of the miR-17/92 cluster, which will induce less expression of the miRNAs of the cluster. This might results in upregulation in the expression of the downstream, oncogenic and tumour suppressor, targets of the cluster, which is seen for STAT3. In contradiction with an expected upregulation of the targets, BCL2L11, HIF-1α, MAPK14, PTEN and TGFBRII were found to be downregulated. This might be explained that other factors, in association with air pollution, also contribute to the regulation of these targets. In literature, both the cluster and air pollution has already been shown to be associated with increased cancer risk. Based on the results in this research, it is assumed that prenatal air pollution exposure, via interfering with the upstream activators (E2F1 and subsequently MYC), will downregulate the miR-17/92 cluster. This might lead to differences in the expression of its downstream targets. By interfering with these oncogenic and tumour suppressor genes, prenatal air pollution exposure can predispose these newborn children to disease development and tumorigenesis later in life.

Nonetheless, more evidence should be obtained in the future. In order to do so, it might be interesting to determine pathways of E2F1 and to validate some regulatory targets of this gene to determine via which mechanism air pollution can affect E2F1 expression. For example, it has been shown that E2F1 can be deregulated via disruption of NF-κB, which has already been shown to be deregulated by $PM_{2.5}$ (95, 96). By investigating the mechanisms around the inhibition of E2F1 and MYC, it might be possible to determine the effects of prenatal air pollution on the activators of the cluster. On the other hand, both the up- and downstream targets selected in this research are known to be involved in carcinogenesis. Accordingly, it is alluring to investigate some pathways involved in critical processes, which include our downstream targets. For example, the pathway of apoptosis, in which BCL2L1, PTEN and TP53 are involved, is already known to be associated with PM exposure. Therefore, other members of this pathway, for example, part of the mitochondrial pathway, such as caspase 9 and cytochrome C, are useful targets to validate in this cohort setting to determine its association with prenatal PM exposure. However, the expression of the downstream targets of the cluster are in contradiction with the expectation. Consequently, there should be investigated if other miRNAs, for example the miR-106a-363 cluster, are found to be differentially expressed in cord blood in association with prenatal air pollution exposure. Further investigations on the expression of our targets and the cluster itself should be determined in function of temporal changes. In order to do so, it is interesting to measure the expression of the miRNA cluster and its network in blood of the children in a later stage of life. Therefore, the advantage of this cohort is that we have a follow-up meeting at 4-6 years.

In the second panel of this research, situated in Figure 3, panel B, our goal was to determine possible alterations in the placental expression of some candidate miRNAs in relation with TL at time of birth. In order to do so, available miRNA microarray data, including data about the placental expression of all known human miRNAs, were used. Based on *in silico* work, miRNAs were selected for this research. These miRNAs were known to target genes, shown to be associated with TL within our research department and belong to GO terms including the word "telomere". This research showed, in newborn boys but not in girls, a positive association between TL at birth and 2 specific miRNAs, specifically miR-584 and miR-4291. However, the significant association disappears after FDR correction. Therefore, the initial hypothesis can be confirmed. miR-584 and miR-4291 are differentially expressed in association with telomere length at birth in newborn boys.

However, our associations cannot be observed after multiple testing. Therefore, a future perspective might be to validate the expression of these miRNAs, with RT-PCR, in placental tissue of a larger population. This will increase the power and therefore the ability to determine significant associations. Moreover, if positive associations can be found in a larger population, it might be interesting to investigate the network around these miRNAs to determine possible mechanism by which they affect TL. Therefore, target genes of these miRNAs can be measured, using RT-PCR, in placental tissue of participants of the ENVIR*ON*AGE birth cohort. Also, TERT, known to be a possible mechanisms by which miRNAs can affect TL and telomerase activity, are interesting targets to include in further research. If more knowledge is obtained about the mechanisms by which miRNAs could affect TL, it might be interesting to determine the association between these miRNAs and some age-related diseases, such as cardiovascular diseases. Therefore, the microvasculature is an interesting clinical outcome as well as the thickness and arterial stiffness of the carotid intima-media. Both parameters are investigated at the follow-up meeting at the age of 4-6 years old.

To conclude, this research showed the association between both the miRNA expression of the miR-17/92 cluster as well as the mRNA expression of the targets in its network in association with prenatal air pollution exposure. Moreover, the second part of this research found the association between telomere length at birth and the expression of two miRNAs (miR-584 and miR-4291). Therefore, this research provides biological insight in the underlying epigenetics mechanisms, by microRNA expression, on the regulation of telomere length and the oncogenic mechanisms of the miR-17/92 cluster in early life, in response to *in utero* air pollution exposure during pregnancy.

References

1. Xiang J, Wu J. Feud or friend? The role of the miR-17-92 cluster in tumorigenesis. Current genomics. 2010;11(2):129-35.
2. Grillari J, Hackl M, Grillari-Voglauer R. miR-17–92 cluster: ups and downs in cancer and aging. Biogerontology. 2010;11(4):501-6.
3. Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell death and differentiation. 2013;20(12):1603.
4. Janssen BG, Madhloum N, Gyselaers W, Bijnens E, Clemente DB, Cox B, et al. Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIR ON AGE): a birth cohort study. International journal of epidemiology. 2017;46(5):1386-7m.
5. Barker DJ. The fetal and infant origins of adult disease. BMJ: British Medical Journal. 1990;301(6761):1111.
6. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. New England Journal of Medicine. 2008;359(1):61-73.
7. Harries L. MicroRNAs as mediators of the ageing process. Genes. 2014;5(3):656-70.
8. Blackburn EH. Structure and function of telomeres. Nature. 1991;350(6319):569.
9. Hayflick L. The limited in vitro lifetime of human diploid cell strains. Experimental cell research. 1965;37(3):614-36.
10. Heidinger BJ, Blount JD, Boner W, Griffiths K, Metcalfe NB, Monaghan P. Telomere length in early life predicts lifespan. Proceedings of the National Academy of Sciences. 2012;109(5):1743-8.
11. Hjelmborg JB, Dalgård C, Möller S, Steenstrup T, Kimura M, Christensen K, et al. The heritability of leucocyte telomere length dynamics. Journal of medical genetics. 2015;52(5):297-302.
12. Pieters N, Janssen BG, Dewitte H, Cox B, Cuypers A, Lefebvre W, et al. Biomolecular markers within the core axis of aging and particulate air pollution exposure in the elderly: a cross-sectional study. Environmental health perspectives. 2015;124(7):943-50.
13. Farmer SA, Nelin TD, Falvo MJ, Wold LE. Ambient and household air pollution: complex triggers of disease. American Journal of Physiology-Heart and Circulatory Physiology. 2014;307(4):H467-H76.
14. Kampa M, Castanas E. Human health effects of air pollution. Environmental pollution. 2008;151(2):362-7.
15. Organisation WH. Ambient air pollution - a major threat to health and climate: World Helath Organisation 2018 [updated 2018. Available from: https://www.who.int/airpollution/ambient/en/.
16. Pope III CA, Dockery DW. Health effects of fine particulate air pollution: lines that connect. Journal of the air & waste management association. 2006;56(6):709-42.
17. IRCEL-CELINE. High resolution maps particulate matter (PM2.5) [updated Wednesday 21 February 2018. Available from: http://www.irceline.be/en/air-quality/measurements/particulate-matter/history/pm25_anmean_rioifdm.
18. Schroeder H. Developmental brain and behavior toxicity of air pollutants: a focus on the effects of polycyclic aromatic hydrocarbons (PAHs). Critical reviews in environmental science and technology. 2011;41(22):2026-47.
19. Kelly FJ, Fussell JC. Air pollution and public health: emerging hazards and improved understanding of risk. Environmental geochemistry and health. 2015;37(4):631-49.
20. Loomis D, Huang W, Chen G. The International Agency for Research on Cancer (IARC) evaluation of the carcinogenicity of outdoor air pollution: focus on China. Chinese journal of cancer. 2014;33(4):189.
21. Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). The lancet oncology. 2013;14(9):813-22.
22. Blasco MA. The epigenetic regulation of mammalian telomeres. Nature Reviews Genetics. 2007;8(4):299.
23. Hou L, Zhang X, Wang D, Baccarelli A. Environmental chemical exposures and human epigenetics. International journal of epidemiology. 2011;41(1):79-105.
24. Janssen BG, Godderis L, Pieters N, Poels K, Kiciński M, Cuypers A, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. Particle and fibre toxicology. 2013;10(1):22.
25. Breton CV, Marutani AN. Air pollution and epigenetics: recent findings. Current environmental health reports. 2014;1(1):35-45.
26. Hamilton JP. Epigenetics: principles and practice. Digestive diseases. 2011;29(2):130-5.
27. Felekkis K, Touvana E, Stefanou C, Deltas C. microRNAs: a newly described class of encoded molecules that play a role in health and disease. Hippokratia. 2010;14(4):236.
28. Lu M, Zhang Q, Deng M, Miao J, Guo Y, Gao W, et al. An analysis of human microRNA and disease associations. PloS one. 2008;3(10):e3420.
29. Tsamou M, Vrijens K, Madhloum N, Lefebvre W, Vanpoucke C, Nawrot TS. Air pollution-induced placental epigenetic alterations in early life: a candidate miRNA approach. Epigenetics. 2018;13(2):135-46.
30. Concepcion CP, Bonetti C, Ventura A. The miR-17-92 family of microRNA clusters in development and disease. Cancer journal (Sudbury, Mass). 2012;18(3):262.
31. Visone R, Croce CM. MiRNAs and cancer. The American journal of pathology. 2009;174(4):1131-8.
32. Hou L, Wang D, Baccarelli A. Environmental chemicals and microRNAs. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2011;714(1-2):105-12.
33. Jardim MJ, Fry RC, Jaspers I, Dailey L, Diaz-Sanchez D. Disruption of microRNA expression in human airway cells by diesel exhaust particles is linked to tumorigenesis-associated pathways. Environmental health perspectives. 2009;117(11):1745.
34. Bollati V, Marinelli B, Apostoli P, Bonzini M, Nordio F, Hoxha M, et al. Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. Environmental health perspectives. 2010;118(6):763.
35. Jansson MD, Lund AH. MicroRNA and cancer. Molecular oncology. 2012;6(6):590-610.
36. Malumbres M. miRNAs and cancer: an epigenetics view. Molecular aspects of medicine. 2013;34(4):863-74.

37. ElSharawy A, Keller A, Flachsbart F, Wendschlag A, Jacobs G, Kefer N, et al. Genome-wide miRNA signatures of human longevity. Aging cell. 2012;11(4):607-16.
38. Tsamou M, Martens DS, Cox B, Madhloum N, Vrijens K, Nawrot TS. Sex-specific associations between telomere length and candidate miRNA expression in placenta. Journal of translational medicine. 2018;16(1):254.
39. Farooqi A, Mansoor Q, Alaaeddine N, Xu B. Microrna regulation of telomerase reverse transcriptase (tert): Micro machines pull strings of papier-mache puppets. International journal of molecular sciences. 2018;19(4):1051.
40. Slattery ML, Herrick JS, Pellatt AJ, Wolff RK, Mullany LE. Telomere length, TERT, and miRNA expression. PloS one. 2016;11(9):e0162077.
41. (WHO) WHO. Cancer Switzerland2018 [updated 2018. Available from: https://www.who.int/cancer/en/.
42. National Cancer Institute (NCI). What Is Cancer? 2015 [updated February 9, 2015. Available from: https://www.cancer.gov/about-cancer/understanding/what-is-cancer.
43. Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Heerema N, et al. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in Eμ-miR155 transgenic mice. Proceedings of the National Academy of Sciences. 2006;103(18):7024-9.
44. Olive V, Jiang I, He L. mir-17-92, a cluster of miRNAs in the midst of the cancer network. The international journal of biochemistry & cell biology. 2010;42(8):1348-54.
45. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. Cell. 2008;133(2):217-22.
46. Walter BAAJJLMRKRP. Molecular biology of THE CELL. 5te ed. ed. 711 Third Avenue, 8th floor, New York, NY 10017, USA: Garland Science, Taylor & Francis Group, LLC; 2008.
47. Database resources of the National Center for Biotechnology Information [Internet]. Available from: https://www.ncbi.nlm.nih.gov/.
48. Bustin SA, Mueller R. Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. Clinical Science. 2005;109(4):365-79.
49. Cox B, Martens E, Nemery B, Vangronsveld J, Nawrot TS. Impact of a stepwise introduction of smoke-free legislation on the rate of preterm births: analysis of routinely collected birth data. Bmj. 2013;346:f441.
50. Janssen S, Dumont G, Fierens F, Mensink C. Spatial interpolation of air pollution measurements using CORINE land cover data. Atmospheric Environment. 2008;42(20):4884-903.
51. Lefebvre W, Degrawe B, Beckx C, Vanhulsel M, Kochan B, Bellemans T, et al. Presentation and evaluation of an integrated model chain to respond to traffic-and health-related policy questions. Environmental modelling & software. 2013;40:160-70.
52. Lefebvre W, Vercauteren J, Schrooten L, Janssen S, Degraeuwe B, Maenhaut W, et al. Validation of the MIMOSA-AURORA-IFDM model chain for policy support: modeling concentrations of elemental carbon in Flanders. Atmospheric environment. 2011;45(37):6705-13.
53. Bino Maiheu NV, Peter Viaene, Koen De Ridder, Dirk Lauwaet, Nele Smeets, Felix Deutsch, Stijn Janssen. Bepaling van de best beschikbare grootschalige concentratiekaarten luchtkwaliteit voor België. Belgium 2012 December 2012.
54. Martens DS, Plusquin M, Gyselaers W, De Vivo I, Nawrot TS. Maternal pre-pregnancy body mass index and newborn telomere length. BMC medicine. 2016;14(1):148.
55. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic acids research. 2009;37(3):e21-e.
56. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome biology. 2019;20(1):18.
57. Jansson T, Powell TL. Role of the placenta in fetal programming: underlying mechanisms and potential interventional approaches. Clinical science. 2007;113(1):1-13.
58. Motta V, Angelici L, Nordio F, Bollati V, Fossati S, Frascati F, et al. Integrative Analysis of miRNA and inflammatory gene expression after acute particulate matter exposure. Toxicological sciences. 2013;132(2):307-16.
59. Selevan SG, Kimmel CA, Mendola P. Identifying critical windows of exposure for children's health. Environmental health perspectives. 2000;108(suppl 3):451-5.
60. WRIGHT RO. Genetic Susceptibility to Environmental Chemicals in Children's Health and Development. Textbook of Children's Environmental Health. 2013:89.
61. Tilghman SL, Bratton MR, Segar HC, Martin EC, Rhodes LV, Li M, et al. Endocrine disruptor regulation of microRNA expression in breast carcinoma cells. PloS one. 2012;7(3):e32754.
62. Duan J, Yu Y, Li Y, Jing L, Yang M, Wang J, et al. Comprehensive understanding of PM2. 5 on gene and microRNA expression patterns in zebrafish (Danio rerio) model. Science of the Total Environment. 2017;586:666-74.
63. Espin-Perez A, Krauskopf J, Chadeau-Hyam M, van Veldhoven K, Chung F, Cullinan P, et al. Short-term transcriptome and microRNAs responses to exposure to different air pollutants in two population studies. Environmental pollution (Barking, Essex : 1987). 2018;242(Pt A):182-90.
64. Chen L, Li C, Zhang R, Gao X, Qu X, Zhao M, et al. miR-17-92 cluster microRNAs confers tumorigenicity in multiple myeloma. Cancer letters. 2011;309(1):62-70.
65. Humphreys KJ, Cobiac L, Le Leu RK, Van der Hoek MB, Michael MZ. Histone deacetylase inhibition in colorectal cancer cells reveals competing roles for members of the oncogenic miR-17-92 cluster. Molecular Carcinogenesis. 2013;52(6):459-74.
66. Zhang M, Liu Q, Mi S, Liang X, Zhang Z, Su X, et al. Both miR-17-5p and miR-20a alleviate suppressive potential of myeloid-derived suppressor cells by modulating STAT3 expression. The Journal of Immunology. 2011;186(8):4716-24.
67. Johnson DG. The paradox of E2F1: oncogene and tumor suppressor gene. Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center. 2000;27(3):151-7.
68. Okuyama H, Endo H, Akashika T, Kato K, Inoue M. Downregulation of c-MYC protein levels contributes to cancer cell survival under dual deficiency of oxygen and glucose. Cancer research. 2010;70(24):10213-23.
69. Winckelmans E, Vrijens K, Tsamou M, Janssen BG, Saenen ND, Roels HA, et al. Newborn sex-specific transcriptome signatures and gestational exposure to fine particles: findings from the ENVIR ON AGE birth cohort. Environmental Health. 2017;16(1):52.
70. Mucaj V, Lee SS, Skuli N, Giannoukos DN, Qiu B, Eisinger-Mathason TK, et al. MicroRNA-124 expression counteracts pro-survival stress responses in glioblastoma. Oncogene. 2015;34(17):2204.

71. Yamakuchi M, Yagi S, Ito T, Lowenstein CJ. MicroRNA-22 regulates hypoxia signaling in colon cancer cells. PloS one. 2011;6(5):e20291.
72. Chang H-M, Martinez NJ, Thornton JE, Hagan JP, Nguyen KD, Gregory RI. Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nature communications. 2012;3:923.
73. Tal TL, Franzosa JA, Tilton SC, Philbrick KA, Iwaniec UT, Turner RT, et al. MicroRNAs control neurobehavioral development and function in zebrafish. The FASEB Journal. 2012;26(4):1452-61.
74. Sathyan P, Golden HB, Miranda RC. Competing interactions between micro-RNAs determine neural progenitor survival and proliferation after ethanol exposure: evidence from an ex vivo model of the fetal cerebral cortical neuroepithelium. Journal of Neuroscience. 2007;27(32):8546-57.
75. Wang L-L, Zhang Z, Li Q, Yang R, Pei X, Xu Y, et al. Ethanol exposure induces differential microRNA and target gene expression and teratogenic effects which can be suppressed by folic acid supplementation. Human Reproduction. 2008;24(3):562-79.
76. Pietrzykowski AZ, Friesen RM, Martin GE, Puig SI, Nowak CL, Wynne PM, et al. Posttranscriptional regulation of BK channel splice variant stability by miR-9 underlies neuroadaptation to alcohol. Neuron. 2008;59(2):274-87.
77. Costantine M. Physiologic and pharmacokinetic changes in pregnancy. Frontiers in pharmacology. 2014;5:65.
78. Sitras V, Fenton C, Paulssen R, Vårtun Å, Acharya G. Differences in gene expression between first and third trimester human placenta: a microarray study. PloS one. 2012;7(3):e33294-e.
79. Halperin R, Peller S, Rotschild M, Bukovsky I, Schneider D. Placental apoptosis in normal and abnormal pregnancies. Gynecologic and obstetric investigation. 2000;50(2):84-7.
80. Haanen C, Vermes I. Apoptosis: programmed cell death in fetal development. European Journal of Obstetrics & Gynecology and Reproductive Biology. 1996;64(1):129-33.
81. Dopico XC, Evangelou M, Ferreira RC, Guo H, Pekalski ML, Smyth DJ, et al. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nature communications. 2015;6:7000.
82. Martino T, Arab S, Straume M, Belsham DD, Tata N, Cai F, et al. Day/night rhythms in gene expression of the normal murine heart. Journal of Molecular Medicine. 2004;82(4):256-64.
83. Luyten LJ, Saenen ND, Janssen BG, Vrijens K, Plusquin M, Roels HA, et al. Air pollution and the fetal origin of disease: A systematic review of the molecular signatures of air pollution exposure in human placenta. Environmental research. 2018;166:310-23.
84. Shiotani A, Uedo N, Iishi H, Murao T, Kanzaki T, Kimura Y, et al. H. pylori eradication did not improve dysregulation of specific oncogenic miRNAs in intestinal metaplastic glands. Journal of gastroenterology. 2012;47(9):988-98.
85. Khuu C, Utheim TP, Sehic A. The three paralogous microRNA clusters in development and disease, miR-17-92, miR-106a-363, and miR-106b-25. Scientifica. 2016;2016.
86. Nagel G, Stafoggia M, Pedersen M, Andersen ZJ, Galassi C, Munkenast J, et al. Air pollution and incidence of cancers of the stomach and the upper aerodigestive tract in the European Study of Cohorts for Air Pollution Effects (ESCAPE). International journal of cancer. 2018;143(7):1632-43.
87. Bilsland AE, Revie J, Keith W. MicroRNA and senescence: the senectome, integration and distributed control. Crit Rev Oncog. 2013;18(4):373-90.
88. Lin J, Kroenke CH, Epel E, Kenna HA, Wolkowitz OM, Blackburn E, et al. Greater endogenous estrogen exposure is associated with longer telomeres in postmenopausal women at risk for cognitive decline. Brain research. 2011;1379:224-31.
89. Smith L, Luchini C, Demurtas J, Soysal P, Stubbs B, Hamer M, et al. Telomere length and health outcomes: An umbrella review of systematic reviews and meta-analyses of observational studies. Ageing research reviews. 2019.
90. Zhang Y, Wang Y, Wang J. MicroRNA-584 inhibits cell proliferation and invasion in non-small cell lung cancer by directly targeting MTDH. Experimental and therapeutic medicine. 2018;15(2):2203-11.
91. Sun B, Wang Y, Kota K, Shi Y, Motlak S, Makambi K, et al. Telomere length variation: A potential new telomere biomarker for lung cancer risk. Lung cancer. 2015;88(3):297-303.
92. Sand M, Skrygan M, Sand D, Georgas D, Gambichler T, Hahn SA, et al. Comparative microarray analysis of microRNA expression profiles in primary cutaneous malignant melanoma, cutaneous malignant melanoma metastases, and benign melanocytic nevi. Cell and tissue research. 2013;351(1):85-98.
93. Biron-Shental T, Sukenik-Halevy R, Sharon Y, Goldberg-Bittman L, Kidron D, Fejgin MD, et al. Short telomeres may play a role in placental dysfunction in preeclampsia and intrauterine growth restriction. American journal of obstetrics and gynecology. 2010;202(4):381. e1-. e7.
94. Enquobahrie DA, Abetew DF, Sorensen TK, Willoughby D, Chidambaram K, Williams MA. Placental microRNA expression in pregnancies complicated by preeclampsia. American journal of obstetrics and gynecology. 2011;204(2):178. e12-. e21.
95. Ginsberg D. E2F1 pathways to apoptosis. FEBS letters. 2002;529(1):122-5.
96. Dagher Z, Garçon G, Billet S, Verdin A, Ledoux F, Courcot D, et al. Role of nuclear factor-kappa B activation in the adverse effects induced by air pollution particulate matter (PM2. 5) in human epithelial lung cells (L132) in culture. Journal of Applied Toxicology: An International Journal. 2007;27(3):284-90.

Appendices

Appendix 1:List of Gene Ontology terms

Table 12. Gene Ontology (GO) terms, used for *in silico* work and selection of miRNAs, containing the term 'telomere'

Gene Ontology (GO) terms	
GO:0007004	Telomere maintenance via telomerase
GO:0010521	Telomerase inhibitory activity
GO: 0032204	Regulation of telomere maintenance
GO:0032211	Negative regulation of telomere maintenance via telomerase
GO:0000781	Chromosome, telomeric region
GO:0032205	Negative regulation of telomere maintenance
GO:0010833	Telomere maintenance via telomere lengthening
GO:0000783	Nuclear telomere cap complex
GO:0003691	Double-stranded telomeric DNA binding
GO:0051974	Negative regulation of telomerase activity
GO:0000784	Nuclear chromosome, telomeric region
GO:0031848	Protection from non-homologous end joining at telomere
GO:0042162	Telomeric DNA binding
GO:0045141	Meiotic telomere clustering
GO:0016233	Telomere capping
GO:0032202	Telomere assembly
GO:0051973	Positive regulation of telomerase activity
GO:0060381	Positive regulation of single-stranded telomeric DNA binding
GO:0070198	Protein localization to chromosome, telomeric region

Appendix 2: mRNA microarray data list of the top 50 most significant genes in the placenta associated with Telomere length at birth for boys and girls

Table 13. mRNA microarray data list of the top 50 most significant genes, in the placenta, associated with telomere length (TL) at birth, in newborn boys. In 200 placenta samples, 18847 genes were measured, using mRNA microarray analysis. The association between the placental expression of these genes and placental TL at birth were determined using mixed linear models and corrected for multiple testing (q-value). For each gene, the Gene symbols, Gene name, the estimate (β) ± SD, p-value and q-value were given. SD= standard deviation.

Gene symbol	Gene name	β ± SD	p-value	q-value
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	0.95 ± 0.20	1.70E-05	0.310597
KRCC1	lysine-rich coiled-coil 1	0.93 ±0.21	3.30E-05	0.310597
OPRL1	opiate receptor-like 1	2.30 ±0.54	6.39E-05	0.401497
CAPRIN1	cell cycle associated protein 1	1.46 ±0.36	0.000119	0.560286
ZBTB14	zinc finger and BTB domain containing 14	1.11 ±0.28	0.000156	0.560286
LOC101928273	uncharacterized LOC101928273	1.75 ±0.44	0.000202	0.560286
PACRG	PARK2 co-regulated	1.66 ±0.42	0.000226	0.560286
LOC101927856	uncharacterized LOC101927856	2.21 ±0.57	0.000257	0.560286
STARD5	StAR-related lipid transfer (START) domain containing 5	1.64 ± 0.43	0.000297	0.560286
TP53TG3C	TP53 target 3C	2.86 ±0.75	0.000319	0.560286
ELAC1	elaC ribonuclease Z 1	0.92 ±0.24	0.000336	0.560286
MAEL	maelstrom spermatogenic transposon silencer	2.13 ±0.57	0.000406	0.560286
LOC100128386	uncharacterized LOC100128386	3.60 ±0.96	0.000407	0.560286
SMKR1	small lysine-rich protein 1	3.73 ±1.00	0.000416	0.560286
XLOC_013728		2.17±0.60	0.00055	0.680144
TTC30B	tetratricopeptide repeat domain 30B	1.39 ±0.39	0.000618	0.680144
NAV2-AS5	NAV2 antisense RNA 5	3.82 ±1.06	0.00064	0.680144
HIST1H1A	histone cluster 1, H1a	-3.2 ±0.89	0.00065	0.680144
LMCD1-AS1	LMCD1 antisense RNA 1 (head to head)	2.77 ±0.78	0.00073	0.68843
HN1L	hematological and neurological expressed 1-like	-1.21 ±0.34	0.000731	0.68843
ANTXRL	anthrax toxin receptor-like	3.87 ±1.09	0.000782	0.701613
NXPH3	neurexophilin 3	-1.96 ±0.56	0.000907	0.761708
LINC01023	long intergenic non-protein coding RNA 1023	1.00 ±0.29	0.000955	0.761708
C14orf28	chromosome 14 open reading frame 28	1.87 ±0.54	0.00097	0.761708
ETFB	electron-transfer-flavoprotein, beta polypeptide	0.75 ±0.22	0.001166	0.797362
PCTP	phosphatidylcholine transfer protein	1.02 ±0.33	0.001166	0.797362
XLOC_012020		1.79 ±0.53	0.001313	0.797362
LINC00959	long intergenic non-protein coding RNA 959	1.20 ±0.36	0.001324	0.797362
TMX4	thioredoxin-related transmembrane protein 4	1.74 ±0.52	0.001388	0.797362
KLHL3	kelch-like family member 3	2.19 ±0.66	0.001435	0.797362
SPDYA	speedy/RINGO cell cycle regulator family member A	2.68 ±0.80	0.001444	0.797362
LOC101927285	uncharacterized LOC101927285	0.98 ±0.29	0.001453	0.797362
LMCD1	LIM and cysteine-rich domains 1	1.38 ±0.41	0.001478	0.797362
FUT8-AS1	FUT8 antisense RNA 1	2.12 ±0.64	0.001494	0.797362
COMMD8	COMM domain containing 8	0.82 ±0.25	0.001529	0.797362
C21orf119	chromosome 21 open reading frame 119	1.10 ±0.33	0.001591	0.797362
ANXA1	annexin A1	1.23 ±0.37	0.001663	0.797362
LOC100422737	uncharacterized LOC100422737	3.39 ±1.03	0.001688	0.797362
MXI1	MAX interactor 1, dimerization protein	-2.16 ±0.66	0.001697	0.797362
SUMO2	small ubiquitin-like modifier 2	0.51 ±0.15	0.001713	0.797362
TTTY16	testis-specific transcript, Y-linked 16 (non-protein coding)	1.54 ±0.47	0.001735	0.797362
WWOX	WW domain containing oxidoreductase	2.06 ±0.63	0.00182	0.79793
ZNF462	zinc finger protein 462	1.07 ±0.33	0.00185	0.79793
ERCC1	excision repair cross-complementation group 1	0.77 ±0.24	0.001944	0.79793
DSC2	desmocollin 2	-1.17 ±0.36	0.001953	0.79793
XLOC_002675		1.41 ±0.44	0.001982	0.79793
ZNF688	zinc finger protein 688	0.80 ±0.25	0.002003	0.79793

ZHX1	zinc fingers and homeoboxes 1	0.52 ±0.16	0.002074	0.79793
METTL18	methyltransferase like 18	1.27 ±0.40	0.002136	0.79793
ZNF3	zinc finger protein 3	1.65 ±0.52	0.002156	0.79793

Table 14. mRNA microarray data list of the top 50 most significant genes, in the placenta, associated with telomere length (TL) at birth, in newborn girls . In 200 placenta samples, 18847 genes were measured, using mRNA microarray analysis. The association between the placental expression of these genes and placental TL at birth were determined using mixed linear models and corrected for multiple testing (q-value). For each gene, the Gene symbols, Gene name, the estimate (β) ± SD, p-value and q-value were given. SD= standard deviation.

Gene symbol	Gene name	β ± SD	p-value	q-value
LOC100133032	uncharacterized LOC100133032	-3.05 ± 0.77	1.65E-04	0.999920043
NR1I3	nuclear receptor subfamily 1, group I, member 3	-2.27 ± 0.61	4.33E-04	0.999920043
XLOC_009631		-2.05 ± 0.57	5.55E-04	0.999920043
LOC100507501	uncharacterized LOC100507501	-1.84 ± 0.51	6.18E-04	0.999920043
AGRP	agouti related protein homolog (mouse)	-1.98 ± 0.56	7.29E-04	0.999920043
JPH2	junctophilin 2	-2.62 ± 0.74	7.32E-04	0.999920043
NLRC5	NLR family, CARD domain containing 5	-1.68 ± 0.48	9.27E-04	0.999920043
RBMY1B	RNA binding motif protein, Y-linked, family 1, member B	-2.84 ± 0.83	1.05E-03	0.999920043
TBATA	thymus, brain and testes associated	-1.57 ± 0.47	1.37E-03	0.999920043
XLOC_013124		-2.77 ± 0.83	1.37E-03	0.999920043
LOC100131150	uncharacterized LOC100131150	-1.92 ± 0.59	1.61E-03	0.999920043
FLJ31713	uncharacterized protein FLJ31713	-2.12 ± 0.65	1.67E-03	0.999920043
CYP2D6	cytochrome P450, family 2, subfamily D, polypeptide 6	-2.19 ± 0.67	1.72E-03	0.999920043
RNF222	ring finger protein 222	-3.85 ± 1.18	1.78E-03	0.999920043
FAN1	FANCD2/FANCI-associated nuclease 1	-1.57 ± 0.49	1.96E-03	0.999920043
FAM118A	family with sequence similarity 118, member A	-2.65 ± 0.83	2.07E-03	0.999920043
KHDC1	KH homology domain containing 1	-2.61 ± 0.83	2.37E-03	0.999920043
LOC101929918	uncharacterized LOC101929918	-2.06 ± 0.66	2.55E-03	0.999920043
XLOC_006688		-2.06± 0.66	2.63E-03	0.999920043
LOC401317	uncharacterized LOC401317	-3.27 ± 1.05	2.72E-03	0.999920043
XLOC_013766		-1.51 ± 0.49	2.87E-03	0.999920043
SFR1	SWI5-dependent recombination repair 1	-2.34 ± 0.75	2.89E-03	0.999920043
XLOC_006971		-1.65 ± 0.54	3.12E-03	0.999920043
LCN8	lipocalin 8	1.58 ± 0.52	3.16E-03	0.999920043
XLOC_012183		-2.11 ± 0.69	3.17E-03	0.999920043
XLOC_013546		-2.29 ± 0.75	3.18E-03	0.999920043
XLOC_013773		-1.29 ± 0.42	3.20E-03	0.999920043
ARSA	arylsulfatase A	-1.49 ± 0.49	3.20E-03	0.999920043
LOC102724572	uncharacterized LOC102724572	-2.40 ± 0.78	3.21E-03	0.999920043
AZGP1	alpha-2-glycoprotein 1, zinc-binding	-2.37 ± 0.78	3.32E-03	0.999920043
LOC100132188	LP7097	-1.81 ± 0.60	3.37E-03	0.999920043
CHST6	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6	2.41 ± 0.79	3.50E-03	0.999920043
TAAR2	trace amine associated receptor 2	-2.80 ± 0.92	3.52E-03	0.999920043
XLOC_013161		-1.32 ± 0.43	3.53E-03	0.999920043
LOC100130539	uncharacterized LOC100130539	-0.92 ± 0.31	3.55E-03	0.999920043
PRO1804	uncharacterized LOC100133319	-1.74 ± 0.58	3.70E-03	0.999920043
ANKRD36BP2	ankyrin repeat domain 36B pseudogene 2	1.49 ± 0.50	3.79E-03	0.999920043
LINC00200	long intergenic non-protein coding RNA 200	-2.08 ± 0.69	3.80E-03	0.999920043
TMEM179	transmembrane protein 179	-2.09 ± 0.70	3.90E-03	0.999920043
PTPN2	protein tyrosine phosphatase, non-receptor type 2	0.91 ± 0.30	3.90E-03	0.999920043
LTB4R2	leukotriene B4 receptor 2	-1.02 ± 0.34	3.97E-03	0.999920043
PROB1	proline-rich basic protein 1	-2.11 ± 0.71	4.03E-03	0.999920043
PHF12	PHD finger protein 12	-1.27 ± 0.43	4.09E-03	0.999920043
CDRT8	CMT1A duplicated region transcript 8	-2.68 ± 0.90	4.10E-03	0.999920043
SIRPG	signal-regulatory protein gamma	-1.99 ± 0.67	4.26E-03	0.999920043
XLOC_003613		-2.28 ± 0.77	4.26E-03	0.999920043
CASC3	cancer susceptibility candidate 3	-0.68 ± 0.23	4.53E-03	0.999920043
PBOV1	prostate and breast cancer overexpressed 1	-2.84 ± 0.97	4.54E-03	0.999920043
ZNF717	zinc finger protein 717	-1.96 ± 0.67	4.61E-03	0.999920043
CIZ1	CDKN1A interacting zinc finger protein 1	-1.07 ± 0.37	4.83E-03	0.999920043

Appendix 3: miRNA input list for new-born boys and girls– miRNAs selected based on Gene Ontology and mRNA microarray data

Table 15. Input list of 552 miRNAs for statistical analysis for the newborn boys. Based on the boys mRNA microarray data, a selection of genes was made, which were found to be significantly associated with telomere length (TL) at birth and belonging to Gene Ontology (GO) terms including the term "telomere". The miRNAs targeting these genes were predicted and used for further analysis.

miRNA selection						
hsa- miR-3154	hsa-miR-3167	hsa-miR-4729	hsa-miR-6079	hsa-miR-12136	hsa-miR-3680-5p	hsa-miR-505-5p
hsa- miR-4691-3p	hsa-miR-3168	hsa-miR-4731-3p	hsa-miR-6085	hsa-miR-122-5p	hsa-miR-3682-3p	hsa-miR-506-5p
hsa-let-7a-3p	hsa-miR-3169	hsa-miR-4744	hsa-miR-612	hsa-miR-1226-3p	hsa-miR-3685	hsa-miR-507
hsa-let-7b-3p	hsa-miR-3176	hsa-miR-4745-5p	hsa-miR-6134	hsa-miR-1250-3p	hsa-miR-3686	hsa-miR-5089-3p
hsa-let-7c-3p	hsa-miR-3180-5p	hsa-miR-4760-3p	hsa-miR-6165	hsa-miR-1251-3p	hsa-miR-3688-3p	hsa-miR-5092
hsa-let-7f-1-3p	hsa-miR-3182	hsa-miR-4761-3p	hsa-miR-616-5p	hsa-miR-1252-3p	hsa-miR-3689a-3p	hsa-miR-513a-3p
hsa-let-7f-2-3p	hsa-miR-3185	hsa-miR-4766-3p	hsa-miR-625-5p	hsa-miR-1257	hsa-miR-3689b-3p	hsa-miR-513a-5p
hsa-miR-10393-3p	hsa-miR-3187-3p	hsa-miR-4775	hsa-miR-627-3p	hsa-miR-125a-3p	hsa-miR-3689c	hsa-miR-513b-3p
hsa-miR-10393-5p	hsa-miR-3187-5p	hsa-miR-4776-3p	hsa-miR-631	hsa-miR-1261	hsa-miR-3689d	hsa-miR-513c-3p
hsa-miR-10394-5p	hsa-miR-3198	hsa-miR-4777-3p	hsa-miR-633	hsa-miR-126-5p	hsa-miR-371b-5p	hsa-miR-5189-5p
hsa-miR-10398-5p	hsa-miR-323b-3p	hsa-miR-4778-3p	hsa-miR-637	hsa-miR-1271-3p	hsa-miR-373-5p	hsa-miR-518a-5p
hsa-miR-103a-3p	hsa-miR-32-3p	hsa-miR-4778-5p	hsa-miR-642a-5p	hsa-miR-1273h-5p	hsa-miR-374a-5p	hsa-miR-5191
hsa-miR-10522-5p	hsa-miR-335-3p	hsa-miR-4779	hsa-miR-6499-3p	hsa-miR-1276	hsa-miR-374b-3p	hsa-miR-5197-3p
hsa-miR-10523-5p	hsa-miR-337-3p	hsa-miR-4780	hsa-miR-6502-5p	hsa-miR-1277-5p	hsa-miR-374b-5p	hsa-miR-519a-2-5p
hsa-miR-10527-5p	hsa-miR-338-3p	hsa-miR-4786-3p	hsa-miR-6507-5p	hsa-miR-1284	hsa-miR-374c-5p	hsa-miR-520a-5p
hsa-miR-105-5p	hsa-miR-338-5p	hsa-miR-4789-5p	hsa-miR-6513-5p	hsa-miR-1285-3p	hsa-miR-377-3p	hsa-miR-520b-5p
hsa-miR-107	hsa-miR-339-5p	hsa-miR-4795-3p	hsa-miR-651-3p	hsa-miR-1286	hsa-miR-378a-5p	hsa-miR-520d-5p
hsa-miR-1183	hsa-miR-33a-3p	hsa-miR-4796-5p	hsa-miR-6516-5p	hsa-miR-1288-5p	hsa-miR-379-3p	hsa-miR-524-5p
hsa-miR-1185-1-3p	hsa-miR-340-3p	hsa-miR-4797-5p	hsa-miR-6529-3p	hsa-miR-1301-5p	hsa-miR-380-3p	hsa-miR-525-5p
hsa-miR-1185-2-3p	hsa-miR-340-5p	hsa-miR-4801	hsa-miR-653-3p	hsa-miR-1303	hsa-miR-382-5p	hsa-miR-526b-5p
hsa-miR-1193	hsa-miR-34a-3p	hsa-miR-484	hsa-miR-655-3p	hsa-miR-1304-3p	hsa-miR-3908	hsa-miR-527
hsa-miR-1200	hsa-miR-3606-3p	hsa-miR-489-3p	hsa-miR-656-3p	hsa-miR-1305	hsa-miR-3913-3p	hsa-miR-539-5p
hsa-miR-1202	hsa-miR-3613-3p	hsa-miR-495-3p	hsa-miR-656-5p	hsa-miR-137-3p	hsa-miR-3913-5p	hsa-miR-541-5p
hsa-miR-1205	hsa-miR-3619-5p	hsa-miR-495-5p	hsa-miR-659-3p	hsa-miR-138-1-3p	hsa-miR-3916	hsa-miR-544b
hsa-miR-12114	hsa-miR-3646	hsa-miR-497-5p	hsa-miR-661	hsa-miR-140-3p	hsa-miR-3918	hsa-miR-548a-5p
hsa-miR-12116	hsa-miR-3652	hsa-miR-498-3p	hsa-miR-664b-5p	hsa-miR-140-5p	hsa-miR-3922-3p	hsa-miR-548aa
hsa-miR-12120	hsa-miR-3658	hsa-miR-499b-5p	hsa-miR-670-3p	hsa-miR-141-3p	hsa-miR-3924	hsa-miR-548ab

hsa-miR-12122	hsa-miR-3661	hsa-miR-5004-3p	hsa-miR-6715a-3p	hsa-miR-1468-3p	hsa-miR-3925-5p	hsa-miR-548ac
hsa-miR-12123	hsa-miR-3662	hsa-miR-5008-3p	hsa-miR-6715b-3p	hsa-miR-146a-5p	hsa-miR-3928-3p	hsa-miR-548ad-5p
hsa-miR-12128	hsa-miR-3665	hsa-miR-5011-5p	hsa-miR-6732-5p	hsa-miR-146b-5p	hsa-miR-3941	hsa-miR-548ae-3p
hsa-miR-12129	hsa-miR-3671	hsa-miR-501-5p	hsa-miR-6733-3p	hsa-miR-148a-3p	hsa-miR-3942-5p	hsa-miR-548ae-5p
hsa-miR-12135	hsa-miR-3678-3p	hsa-miR-503-3p	hsa-miR-6733-5p	hsa-miR-148a-5p	hsa-miR-3945	hsa-miR-548ag
hsa-miR-150-5p	hsa-miR-3974	hsa-miR-548ai	hsa-miR-6734-3p	hsa-miR-148b-3p	hsa-miR-3972	hsa-miR-4475
hsa-miR-152-3p	hsa-miR-410-3p	hsa-miR-548aj-3p	hsa-miR-6736-3p	hsa-miR-6822-5p	hsa-miR-7110-5p	hsa-miR-4477a
hsa-miR-153-3p	hsa-miR-411-3p	hsa-miR-548ak	hsa-miR-6737-3p	hsa-miR-6823-3p	hsa-miR-215-3p	hsa-miR-4477b
hsa-miR-153-5p	hsa-miR-4251	hsa-miR-548am-3p	hsa-miR-6739-3p	hsa-miR-6827-3p	hsa-miR-216b-3p	hsa-miR-4481
hsa-miR-1-5p	hsa-miR-4255	hsa-miR-548am-5p	hsa-miR-6739-5p	hsa-miR-6828-5p	hsa-miR-223-5p	hsa-miR-4482-3p
hsa-miR-181a-5p	hsa-miR-4259	hsa-miR-548ap-3p	hsa-miR-6740-3p	hsa-miR-6830-5p	hsa-miR-24-2-5p	hsa-miR-4489
hsa-miR-181b-5p	hsa-miR-4260	hsa-miR-548ap-5p	hsa-miR-6750-5p	hsa-miR-6832-5p	hsa-miR-2467-3p	hsa-miR-4495
hsa-miR-181c-5p	hsa-miR-4262	hsa-miR-548aq-3p	hsa-miR-6752-3p	hsa-miR-6834-3p	hsa-miR-2681-5p	hsa-miR-4499
hsa-miR-181d-5p	hsa-miR-4263	hsa-miR-548aq-5p	hsa-miR-6752-5p	hsa-miR-6842-3p	hsa-miR-28-3p	hsa-miR-4500
hsa-miR-1827	hsa-miR-4264	hsa-miR-548ar-5p	hsa-miR-6757-3p	hsa-miR-6842-5p	hsa-miR-297	hsa-miR-4502
hsa-miR-186-5p	hsa-miR-4266	hsa-miR-548as-5p	hsa-miR-6759-3p	hsa-miR-6846-5p	hsa-miR-299-5p	hsa-miR-4503
hsa-miR-190a-3p	hsa-miR-4273	hsa-miR-548au-5p	hsa-miR-6763-5p	hsa-miR-6847-5p	hsa-miR-29b-2-5p	hsa-miR-4505
hsa-miR-1912-5p	hsa-miR-4275	hsa-miR-548ay-5p	hsa-miR-6765-5p	hsa-miR-6848-5p	hsa-miR-302c-5p	hsa-miR-4506
hsa-miR-1913	hsa-miR-4282	hsa-miR-548az-5p	hsa-miR-676-5p	hsa-miR-6851-5p	hsa-miR-30a-3p	hsa-miR-4509
hsa-miR-196a-1-3p	hsa-miR-4283	hsa-miR-548b-5p	hsa-miR-6768-3p	hsa-miR-6853-3p	hsa-miR-30b-3p	hsa-miR-450a-1-3p
hsa-miR-1972	hsa-miR-429	hsa-miR-548ba	hsa-miR-6770-5p	hsa-miR-6857-3p	hsa-miR-30c-1-3p	hsa-miR-450a-2-3p
hsa-miR-197-3p	hsa-miR-4291	hsa-miR-548bb-3p	hsa-miR-6772-3p	hsa-miR-6857-5p	hsa-miR-30c-2-3p	hsa-miR-4519
hsa-miR-199a-5p	hsa-miR-4293	hsa-miR-548bb-5p	hsa-miR-6773-3p	hsa-miR-6859-5p	hsa-miR-30d-3p	hsa-miR-4520-3p
hsa-miR-199b-5p	hsa-miR-4294	hsa-miR-548c-3p	hsa-miR-6779-3p	hsa-miR-6860	hsa-miR-30e-3p	hsa-miR-452-5p
hsa-miR-19a-3p	hsa-miR-4309	hsa-miR-548c-5p	hsa-miR-6779-5p	hsa-miR-6865-5p	hsa-miR-3117-5p	hsa-miR-4528
hsa-miR-19b-3p	hsa-miR-4310	hsa-miR-548d-3p	hsa-miR-6780a-5p	hsa-miR-6871-3p	hsa-miR-3122	hsa-miR-4534
hsa-miR-200a-3p	hsa-miR-4311	hsa-miR-548d-5p	hsa-miR-6782-3p	hsa-miR-6873-3p	hsa-miR-3123	hsa-miR-4656
hsa-miR-200b-3p	hsa-miR-4313	hsa-miR-548e-5p	hsa-miR-6788-5p	hsa-miR-6878-3p	hsa-miR-3124-3p	hsa-miR-4659a-5p
hsa-miR-200c-3p	hsa-miR-4317	hsa-miR-548h-3p	hsa-miR-6801-5p	hsa-miR-6878-5p	hsa-miR-3125	hsa-miR-4659b-5p
hsa-miR-203a-3p	hsa-miR-4318	hsa-miR-548h-5p	hsa-miR-6807-5p	hsa-miR-6879-3p	hsa-miR-3133	hsa-miR-4661-3p
hsa-miR-203b-3p	hsa-miR-4425	hsa-miR-548i	hsa-miR-6808-5p	hsa-miR-6885-3p	hsa-miR-3136-3p	hsa-miR-4666a-3p
hsa-miR-2053	hsa-miR-4430	hsa-miR-548j-3p	hsa-miR-6810-5p	hsa-miR-6889-3p	hsa-miR-3137	hsa-miR-4668-3p
hsa-miR-205-3p	hsa-miR-4433a-5p	hsa-miR-548j-5p	hsa-miR-6813-5p	hsa-miR-6891-3p	hsa-miR-3140-3p	hsa-miR-4668-5p

hsa-miR-2054	hsa-miR-4433b-5p	hsa-miR-548l	hsa-miR-6815-5p	hsa-miR-6892-3p	hsa-miR-3143	hsa-miR-4672
hsa-miR-2110	hsa-miR-4436b-5p	hsa-miR-548m	hsa-miR-6817-3p	hsa-miR-6893-5p	hsa-miR-3144-3p	hsa-miR-4676-3p
hsa-miR-211-3p	hsa-miR-4450	hsa-miR-548n	hsa-miR-6818-3p	hsa-miR-7106-5p	hsa-miR-3145-3p	hsa-miR-4677-5p
hsa-miR-2114-3p	hsa-miR-4452	hsa-miR-548o-5p	hsa-miR-6819-3p	hsa-miR-711	hsa-miR-3147	hsa-miR-4684-3p
hsa-miR-212-5p	hsa-miR-4457	hsa-miR-548p	hsa-miR-6822-3p	hsa-miR-7110-3p	hsa-miR-3148	hsa-miR-4695-5p
hsa-miR-548t-3p	hsa-miR-7-1-3p	hsa-miR-3150a-3p	hsa-miR-570-5p	hsa-miR-595	hsa-miR-8064	hsa-miR-5590-5p
hsa-miR-548t-5p	hsa-miR-7152-5p	hsa-miR-3153	hsa-miR-577	hsa-miR-598-3p	hsa-miR-8069	hsa-miR-5680
hsa-miR-548w	hsa-miR-7153-5p	hsa-miR-3154	hsa-miR-578	hsa-miR-603	hsa-miR-8071	hsa-miR-5688
hsa-miR-548x-3p	hsa-miR-7154-5p	hsa-miR-3155a	hsa-miR-5787	hsa-miR-607	hsa-miR-8076	hsa-miR-5692a
hsa-miR-548y	hsa-miR-7155-3p	hsa-miR-3155b	hsa-miR-584-3p	hsa-miR-6072	hsa-miR-8082	hsa-miR-5692b
hsa-miR-548z	hsa-miR-7156-5p	hsa-miR-3157-5p	hsa-miR-584-5p	hsa-miR-892c-3p	hsa-miR-8084	hsa-miR-5692c
hsa-miR-550a-3-5p	hsa-miR-7157-3p	hsa-miR-3160-5p	hsa-miR-586	hsa-miR-892c-5p	hsa-miR-8485	hsa-miR-5696
hsa-miR-550a-5p	hsa-miR-7158-3p	hsa-miR-3163	hsa-miR-875-5p	hsa-miR-940	hsa-miR-873-5p	hsa-miR-570-3p
hsa-miR-550b-2-5p	hsa-miR-7158-5p	hsa-miR-3166	hsa-miR-876-5p	hsa-miR-5579-3p	hsa-miR-7844-5p	hsa-miR-4714-3p
hsa-miR-551b-5p	hsa-miR-7-2-3p	hsa-miR-4703-5p	hsa-miR-887-5p	hsa-miR-558	hsa-miR-7853-5p	hsa-miR-4717-3p
hsa-miR-556-3p	hsa-miR-7515	hsa-miR-4704-3p	hsa-miR-888-5p	hsa-miR-5582-3p	hsa-miR-7856-5p	hsa-miR-4719
hsa-miR-557	hsa-miR-7705	hsa-miR-4711-5p	hsa-miR-889-3p	hsa-miR-5584-5p	hsa-miR-7977	hsa-miR-4728-5p
hsa-miR-9500	hsa-miR-944	hsa-miR-98-3p	hsa-miR-889-5p	hsa-miR-5585-3p	hsa-miR-8060	hsa-miR-589-5p
hsa-miR-95-5p	hsa-miR-590-3p	hsa-miR-8062	hsa-miR-891a-3p	hsa-miR-559		

Table 16. Input list of 98 miRNAs for statistical analysis for the newborn girls. Based on the girls mRNA microarray data, a selection of genes was made, which were found to be significantly associated with telomere length (TL) at birth and belonging to Gene Ontology (GO) terms including the term "telomere". The miRNAs targeting these genes were predicted and used for further analysis

miRNA selection						
hsa-miR-100-3p	hsa-miR-4489	hsa-miR-586	hsa-miR-3194-3p	hsa-miR-509-3-5p	hsa-miR-6796-3p	hsa-miR-6728-3p
hsa-miR-12113	hsa-miR-4490	hsa-miR-590-3p	hsa-miR-34b-5p	hsa-miR-509-5p	hsa-miR-6802-3p	hsa-miR-6742-3p
hsa-miR-12130	hsa-miR-452-5p	hsa-miR-6074	hsa-miR-3613-3p	hsa-miR-513a-5p	hsa-miR-6813-5p	hsa-miR-6744-3p
hsa-miR-12131	hsa-miR-4539	hsa-miR-6085	hsa-miR-3618	hsa-miR-548e-5p	hsa-miR-6815-3p	hsa-miR-6782-5p
hsa-miR-12136	hsa-miR-4641	hsa-miR-616-5p	hsa-miR-371b-5p	hsa-miR-548g-3p	hsa-miR-6885-3p	hsa-miR-494-3p
hsa-miR-1250-3p	hsa-miR-4659a-3p	hsa-miR-621	hsa-miR-373-5p	hsa-miR-554	hsa-miR-7154-5p	hsa-miR-498-5p
hsa-miR-1277-5p	hsa-miR-4659b-3p	hsa-miR-646	hsa-miR-376a-5p	hsa-miR-5571-3p	hsa-miR-758-5p	hsa-miR-5011-5p
hsa-miR-1283	hsa-miR-4676-3p	hsa-miR-6499-3p	hsa-miR-384	hsa-miR-5692a	hsa-miR-760	hsa-miR-5089-5p
hsa-miR-1295b-5p	hsa-miR-4679	hsa-miR-6507-5p	hsa-miR-4259	hsa-miR-5692b	hsa-miR-766-3p	hsa-miR-202-5p
hsa-miR-153-5p	hsa-miR-4749-3p	hsa-miR-6512-5p	hsa-miR-4313	hsa-miR-5692c	hsa-miR-7702	hsa-miR-26a-1-3p
hsa-miR-16-1-3p	hsa-miR-4757-5p	hsa-miR-6515-3p	hsa-miR-432-5p	hsa-miR-5693	hsa-miR-7977	hsa-miR-26a-2-3p
hsa-miR-190a-3p	hsa-miR-4760-5p	hsa-miR-653-3p	hsa-miR-4330	hsa-miR-5699-3p	hsa-miR-8061	hsa-miR-3120-3p
hsa-miR-1912-3p	hsa-miR-4777-3p	hsa-miR-653-5p	hsa-miR-4418	hsa-miR-582-3p	hsa-miR-876-3p	hsa-miR-3158-5p
hsa-miR-1913	hsa-miR-483-3p	hsa-miR-664a-3p	hsa-miR-4421	hsa-miR-584-5p	hsa-miR-892c-3p	hsa-miR-93-3p

Appendix 4: miRNA expression of the miR-17/92 cluster in association with air pollution exposure

Table 17. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative miRNA expression of miR-17, miR-20a and miR-92a in cord blood. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 $\mu g/m^3$ for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 $\mu g/m^2$ for BC.

Unadjusted model	**miR-17**		**miR-20a**		**miR-92a**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure						
1st trimester	-0.20 (-0.32 to -0.05)	**0.01**	-0.11 (-0.25 to 0.06)	0.19	0.12 (-0.02 to 0.29)	0.10
2nd trimester	-0.18 (-0.30 to -0.04)	**0.01**	-0.10 (-0.23 to 0.05)	0.17	-0.12 (-0.22 to 0.00002)	0.05
3rd trimester	-0.20 (-0.32 to -0.06)	**<0.01**	-0.31 (-0.41 to -0.19)	**<0.01**	-0.14 (-0.24 to -0.01)	**0.03**
entire pregnancy	-0.46 (-0.59 to -0.29)	**<0.01**	-0.43 (-0.57 to -0.25)	**<0.01**	-0.15 (-0.32 to 0.07)	0.17
NO_2 exposure						
1st trimester	-0.14 (-0.23 to -0.03)	**0.01**	-0.09 (-0.19 to 0.03)	0.13	0.04 (-0.05 to 0.15)	0.41
2nd trimester	-0.21 (-0.30 to -0.11)	**<0.01**	-0.17 (-0.26 to -0.07)	**<0.01**	-0.14 (-0.22 to -0.06)	**<0.01**
3rd trimester	-0.09 (-0.20 to 0.04)	0.16	-0.14 (-0.24 to -0.02)	**0.02**	-0.06 (-0.66 to 0.04)	0.20
entire pregnancy	-0.23 (-0.34 to -0.10)	**<0.01**	-0.20 (-0.31 to -0.07)	**<0.01**	-0.09 (-0.20 to 0.03)	0.14
BC exposure						
1st trimester	-0.22 (-0.33 to -0.08)	**<0.01**	-0.18 (-0.30 to -0.04)	**0.02**	0.04 (-0.09 to 0.18)	0.56
2nd trimester	-0.25 (-0.37 to -0.11)	**<0.01**	-0.26 (-0.38 to -0.12)	**<0.01**	-0.22 (-0.32 to -0.10)	**<0.01**
3rd trimester	0.14 (-0.06 to 0.38)	0.17	-0.05 (-0.21 to 0.15)	0.59	0.10 (-0.05 to 0.29)	0.21
entire pregnancy	-0.26 (-0.42 to -0.07)	**0.01**	-0.31 (-0.45 to -0.13)	**<0.01**	-0.08 (-0.24 to 0.11)	0.38

Table 18. Spearman correlations for the different miRNAs (miR-17, miR-20a and miR-92a) and the different up-and downstream targets in cord blood (n=432) The Spearman correlations are given with the corresponding p-value.

	miR-17		miR-20a		miR-92a	
Up-or downstream targets	**Spearman Correlation Coefficients**	**p-value**	**Spearman Correlation Coefficients**	**p-value**	**Spearman Correlation Coefficients**	**p-value**
BCL2L11	0.01	0.80	-0.04	0.47	-0.04	0.46
CDKN1A	-0.08	0.12	-0.02	0.63	-0.07	0.17
E2F1	0.03	0.55	-0.0006	0.99	0.04	0.42
HIF-1α	-0.01	0.77	0.09	0.08	-0.10	**0.04**
MAPK14	0.06	0.28	-0.005	0.92	0.07	0.21
MYC	-0.16	**<0.01**	-0.16	**<0.01**	-0.08	0.12
PTEN	0.03	0.51	0.1	0.06	-0.14	**<0.01**
STAT3	-0.02	0.75	-0.003	0.95	0.03	0.62
TGFBRII	-0.002	0.97	0.01	0.81	0.04	0.39
TP53	-0.06	0.26	-0.11	**0.03**	0.05	0.37

Table 19. Spearman correlations among the different miRNAs (miR-17, miR-20a and miR-92a) in cord blood (n=432). The Spearman correlations are given with the corresponding p-value.

	miR-17		miR-20a		miR-92a	
	Spearman Correlation Coefficients	**p-value**	**Spearman Correlation Coefficients**	**p-value**	**Spearman Correlation Coefficients**	**p-value**
miR-17	-	-	0.86	**<0.01**	0.71	**<0.01**
miR-20a	0.86	**<0.01**	-	-	0.57	**<0.01**
miR-92a	0.71	**<0.01**	0.57	**<0.01**	-	-

Appendix 5: The placental expression of the miR-17/92 cluster and its paralogue clusters in association with prenatal air pollution exposure

Table 20. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative miRNA expression of miR-17-3p, miR-17-5p, miR-20a-5p and miR-92a-3p of the miR-17/92 cluster in the placenta. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 $\mu g/m^3$ for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 $\mu g/m^2$ for BC.

Unadjusted model	**miR-17-3p**		**miR-17-5p**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	0.08 (-0.05 to 0.23)	0.21	-0.002 (-0.12 to 0.13)	0.98
2nd trimester	0.08 (-0.05 to 0.22)	0.25	0.08 (-0.05 to 0.22)	0.22
3rd trimester	-0.07 (-0.19 to 0.07)	0.32	0.02 (-0.11 to 0.17)	0.74
entire pregnancy	0.10 (-0.11 to 0.37)	0.38	0.09 (-0.11 to 0.34)	0.40
NO_2 exposure				
1st trimester	0.09 (-0.002 to 0.18)	0.06	0.03 (-0.05 to 0.12)	0.49
2nd trimester	0.01 (-0.08 to 0.11)	0.82	0.03 (-0.06 to 0.13)	0.50
3rd trimester	-0.02 (-0.11 to 0.08)	0.68	0.03 (-0.06 to 0.13)	0.55
entire pregnancy	0.04 (-0.07 to 0.16)	0.46	0.04 (-0.06 to 0.16)	0.44
BC exposure				
1st trimester	0.09 (-0.03 to 0.22)	0.13	0.06 (-0.05 to 0.18)	0.32
2nd trimester	-0.03 (-0.14 to 0.10)	0.66	0.03 (-0.09 to 0.16)	0.68
3rd trimester	-0.06 (-0.20 to 0.10)	0.42	0.02 (-0.13 to 0.19)	0.81
entire pregnancy	0.02 (-0.13 to 0.20)	0.81	0.06 (-0.10 to 0.23)	0.48
Unadjusted model	**miR-20a-5p**		**miR-92a-3p**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	-0.04 (-0.16 to 0.11)	0.60	0.06 (-0.05 to 0.19)	0.28
2nd trimester	0.09 (-0.05 to 0.24)	0.23	-0.05 (-0.15 to 0.06)	0.33
3rd trimester	0.05 (-0.10 to 0.22)	0.53	-0.09 (-0.19 to 0.02)	0.12
entire pregnancy	0.09 (-0.14 to 0.37)	0.48	-0.07 (-0.23 to 0.12)	0.44
NO_2 exposure				
1st trimester	0.02 (-0.07 to 0.12)	0.62	0.01 (-0.06 to 0.09)	0.72
2nd trimester	0.05 (-0.05 to 0.16)	0.34	-0.05 (-0.12 to 0.03)	0.22
3rd trimester	0.06 (-0.05 to 0.18)	0.30	-0.05 (-0.13 to 0.03)	0.19
entire pregnancy	0.06 (-0.06 to 0.19)	0.34	-0.04 (-0.13 to 0.06)	0.41
BC exposure				
1st trimester	0.05 (-0.08 to 0.19)	0.46	-0.01 (-0.10 to 0.10)	0.89
2nd trimester	0.06 (-0.08 to 0.21)	0.40	-0.10 (-0.19 to -0.01)	**0.04**
3rd trimester	0.07 (-0.10 to 0.26à	0.45	-0.11 (-0.22 to 0.02)	0.09
entire pregnancy	0.09 (-0.09 to 0.29)	0.35	-0.10 (-0.22 to 0.03)	0.13

Table 21. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative miRNA expression of miR-18a-5p, miR-19-3p and miR-19b-3p of the miR-17/92 cluster in the placenta. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 $\mu g/m^3$ for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 $\mu g/m^2$ for BC.

Unadjusted model	**miR-18a-5p**		**miR-19a-3p**		**miR-19b-3p**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure						
1st trimester	0.10 (-0.08 to 0.33)	0.28	0.19 (-0.04 to 0.46)	0.11	0.05 (-0.09 to 0.21)	0.50
2nd trimester	0.14 (-0.04 to 0.36)	0.12	0.10 (-0.10 to 0.35)	0.33	0.09 (-0.05 to 0.25)	0.20
3rd trimester	0.01 (-0.16 to 0.22)	0.91	0.03 (-0.18 to 0.28)	0.81	0.02 (-0.12 to 0.19)	0.76
entire pregnancy	0.27 (-0.06 to 0.70)	0.11	0.32 (-0.06 to 0.85)	0.11	0.16 (-0.08 to 0.47)	0.19
NO_2 exposure						
1st trimester	0.11 (-0.01 to 0.25)	0.08	0.12 (-0.02 to 0.29)	0.09	0.05 (-0.04 to 0.16)	0.28
2nd trimester	0.06 (-0.07 to 0.20)	0.37	0.03 (-0.11 to 0.20)	0.70	0.03 (-0.07 to 0.14)	0.61
3rd trimester	0.04 (-0.09 to 0.19)	0.54	0.05 (-0.10 to 0.23)	0.55	0.02 (-0.08 to 0.14)	0.65
entire pregnancy	0.11 (-0.05 to 0.28)	0.18	0.10 (-0.08 to 0.31)	0.28	0.05 (-0.07 to 0.19)	0.41
BC exposure						
1st trimester	0.16 (-0.01 to 0.35)	0.06	0.14 (-0.05 to 0.36)	0.17	0.08 (-0.05 to 0.22)	0.23
2nd trimester	0.05 (-0.11 to 0.25)	0.55	0.02 (-0.16 to 0.25)	0.81	0.03 (-0.10 to 0.19)	0.62
3rd trimester	0.03 (-0.17 to 0.27)	0.81	0.08 (-0.16 to 0.39)	0.53	0.05 (-0.011 to 0.25)	0.54
entire pregnancy	0.14 (-0.08 to 0.43)	0.23	0.13 (-0.12 to 0.46)	0.34	0.09 (-0.08 to 0.30)	0.33

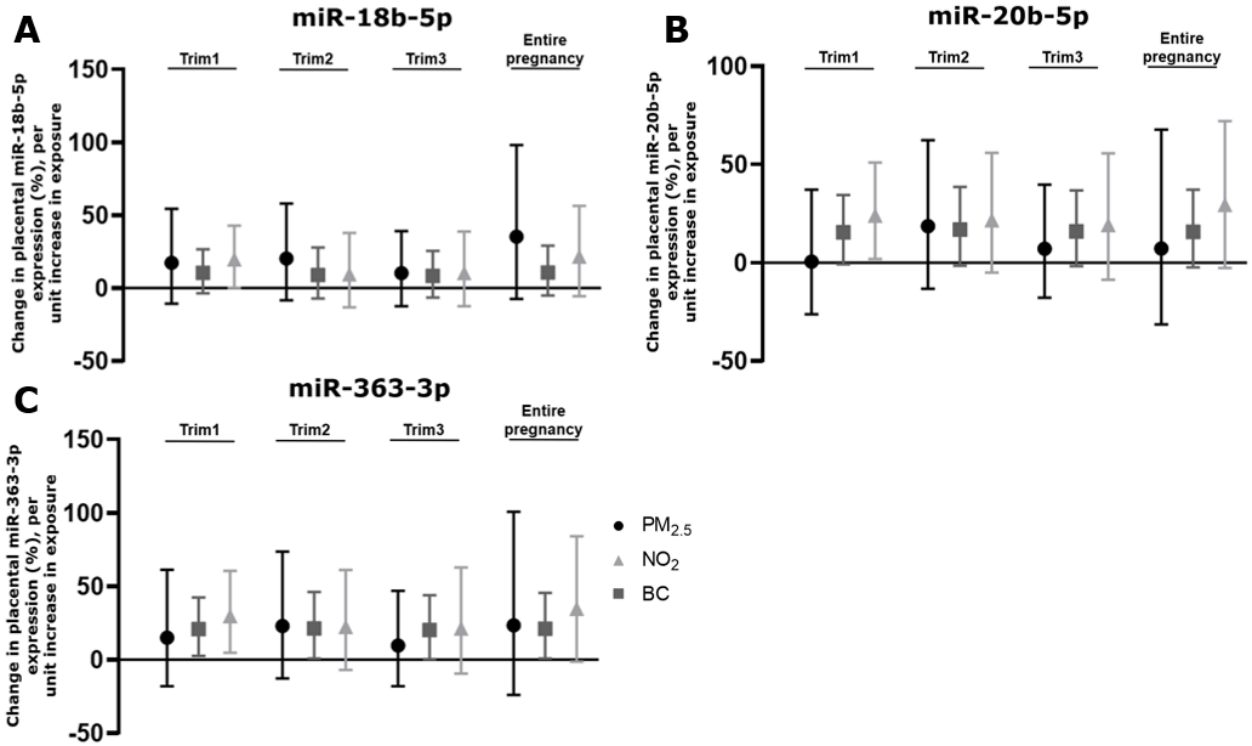

Figure 10. Estimated change (%) in relative miRNA expression for each increase of 5 µg/m³ of $PM_{2.5}$ and NO_2 or 0.5 µg/m³ increment in BC for the different time windows of pregnancy of members of the miR-106a-363 cluster , including (A) miR-18b-5p, (B) miR-20b-5p and (C) miR-363-3p. The percentual changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is adjusted for some covariates: date of delivery, maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy**.**

Table 22. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative miRNA expression of miR-18-5p, miR-20b-5p and miR-363-3p in the placenta. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 $\mu g/m^3$ for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 $\mu g/m^2$ for BC.

Unadjusted model	**miR-18b-5p**		**miR-20b-5p**		**miR-363-3p**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure						
1st trimester	0.10 (-0.07 to 0.30)	0.27	-0.03 (-0.17 to 0.14)	0.72	0.04 (-0.13 to 0.25)	0.63
2nd trimester	0.03 (-0.12 to 0.20)	0.74	0.10 (-0.05 to 0.29)	0.19	0.11 (-0.07 to 0.32)	0.24
3rd trimester	0.02 (-0.15 to 0.48)	0.86	0.05 (-0.11 o 0.24)	0.58	0.04 (-0.14 to 0.26)	0.66
entire pregnancy	0.12 (-0.15 to 0.48)	0.41	0.11 (-0.14 to 0.44)	0.42	0.19 (-0.12 to 0.60)	0.25
NO_2 exposure						
1st trimester	0.06 (-0.05 to0.18)	0.28	0.05 (-0.06 to 0.16)	0.39	0.08 (-0.04 to 0.22)	0.19
2nd trimester	0.002 (-0.11 to 0.13)	0.97	0.07 (-0.04 to 0.20)	0.20	0.09 (-0.04 to 0.24)	0.18
3rd trimester	0.03 (-0.09 to 0.16)	0.68	0.08 (-0.04 to 0.21)	0.21	0.10 (-0.14 to 0.26)	0.17
entire pregnancy	0.04 (-0.09 to 0.19)	0.57	0.09 (-0.04 to 0.25)	0.19	0.13 (-0.03 to 0.31)	0.11
BC exposure						
1st trimester	0.09 (-0.06 to 0.26)	0.26	0.08 (-0.06 to 0.24)	0.30	0.11 (-0.06 to 0.29)	0.21
2nd trimester	-0.03 (-0.17 to 0.14)	0.70	0.09 (-0.06 to 0.26)	0.26	0.09 (-0.08 to 0.30)	0.31
3rd trimester	0.01 (-0.17 to 0.22)	0.96	0.08 (-0.10 to 0.31)	0.40	0.10 (-0.12 to 0.36)	0.41
entire pregnancy	0.03 (-0.16 to 0.27)	0.74	0.13 (-0.07 to 0.37)	0.22	0.15 (-0.07 to 0.44)	0.20

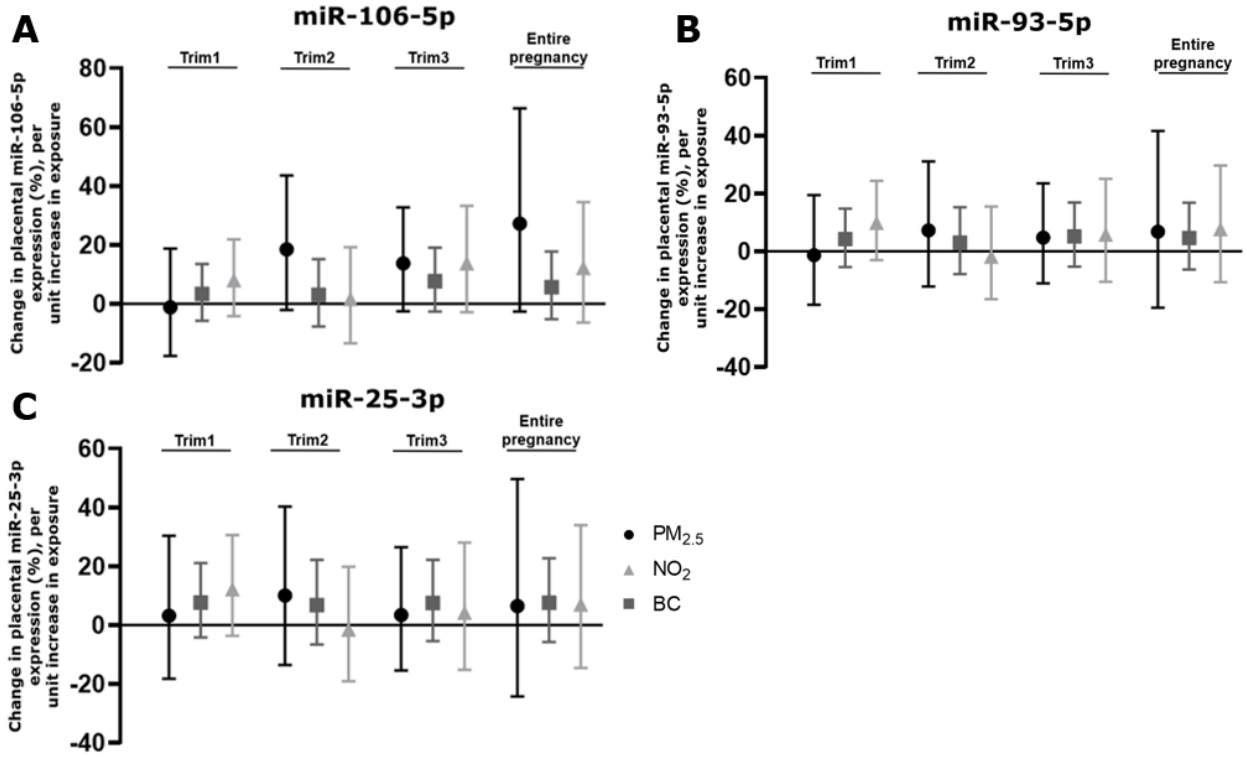

Figure 11. Estimated change (%) in relative miRNA expression for each increase of 5 µg/m^3 of $PM_{2.5}$ and NO_2 or 0.5 µg/m^3 increment in BC for the different time windows of pregnancy of members of the miR-106b-25 cluster , including (A) miR-106-5p, (B) miR-93-5p and (C) miR-25-3p. The percentual changes are presented with a 95% confidence interval. In black, the circle indicates $PM_{2.5}$ exposure, the triangle (light grey) represents BC exposure and in dark grey (square) shows the results associated with NO_2. The model is adjusted for some covariates: date of delivery, maternal age (years), pre-gestational BMI (kg/m2), parity (1,2 or ≥3), educational status (low, middle or high), smoking status (never-smoker, past-smoker or current-smoker), new-born's gestational age (weeks), newborn's sex, ethnicity (European and non-European) and the apparent temperature for each specific time window of pregnancy.

Table 23. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative miRNA expression of miR-106b-5p, miR-93-5p and miR-25-3p in the placenta. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 $\mu g/m^3$ for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 $\mu g/m^2$ for BC.

Unadjusted model	**miR-106b-5p**		**miR-93-5p**		**miR-25-3p**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure						
1st trimester	0.06 (-0.04 to 0.17)	0.22	0.02 (-0.07 to 0.13)	0.64	0.02 (-0.10 to 0.17)	0.71
2nd trimester	0.04 (-0.05 to 0.15)	0.38	0.02 (-0.08 to 0.12)	0.75	0.04 (-0.09 to 0.18)	0.56
3rd trimester	-0.01 (-0.11 to 0.10)	0.84	-0.03 (-0.13 to 0.08)	0.61	-0.02 (-0.15 to 0.12)	0.74
entire pregnancy	0.09 (-0.07 to 0.29)	0.28	0.01 (-0.14 to 0.20)	0.88	0.04 (-0.16 to 0.29)	0.72
NO_2 exposure						
1st trimester	0.04 (-0.03 to 0.11)	0.24	0.02 (-0.05 to 0.09)	0.55	0.03 (-0.06 to 0.12)	0.55
2nd trimester	-0.01 (-0.08 to 0.07)	0.84	-0.01 (-0.08 to 0.06)	0.78	0.02 (-0.08 to 0.11)	0.74
3rd trimester	-0.001 (-0.07 to 0.08)	0.98	-0.01 (-0.08 to 0.07)	0.86	0.01 (-0.08 to 0.12)	0.79
entire pregnancy	0.02 (-0.07 to 0.11)	0.70	0.002 (-0.08 to 0.09)	0.96	0.03 (-0.08 to 0.15)	0.63
BC exposure						
1st trimester	0.05 (-0.03 to 0.15)	0.23	0.04 (-0.05 to 0.13)	0.45	0.03 (-0.08 to 0.15)	0.63
2nd trimester	-0.03 (-0.11 to 0.07)	0.58	-0.04 (-0.13 to 0.06)	0.44	-0.02 (-0.13 to 0.12)	0.80
3rd trimester	0.01 (-0.11 to 0.14)	0.88	-0.02 (-0.14 to 0.10)	0.70	-0.03 (-0.17 to 0.14)	0.74
entire pregnancy	0.02 (-0.10 to 0.16)	0.73	-0.01 (-0.13 to 0.13)	0.89	-0.002 (-0.15 to 0.17)	0.98

Appendix 6: The expression of the upstream targets of the miR-17:92 cluster network, including TP53, STAT3, E2F1 and MYC, in association with prenatal air pollution exposure

Table 24. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative mRNA expression of the upstream targets STAT3, TP53, C-MYC and E2F1 in cord blood. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 μg/m^3 for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 μg/m^2 for BC.

Unadjusted model	**STAT3**		**TP53**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	0.07 (0.01 to 0.12)	**0.01**	0.09 (-0.02 to 0.22)	0.12
2nd trimester	0.02 (-0.02 to 0.07)	0.33	-0.03 (-0.12 to 0.08)	0.56
3rd trimester	-0.03 (-0.07 to 0.02)	0.24	0.06 (-0.05 to 0.18)	0.27
entire pregnancy	0.07 (-0.02 to 0.16)	0.11	0.13 (-0.06 to 0.36)	0.18
NO_2 exposure				
1st trimester	0.03 (-0.01 to 0.06)	0.11	-0.007 (-0.08 to 0.07)	0.85
2nd trimester	-0.001 (-0.03 to 0.03)	0.95	-0.02 (-0.09 to 0.06)	0.66
3rd trimester	-0.03 (-0.06 to 0.01)	0.15	-0.04 (-0.11 to 0.04)	0.29
entire pregnancy	0.004 (-0.04 to 0.05)	0.86	-0.03 (-0.12 to 0.07)	0.53
BC exposure				
1st trimester	0.04 (-0.01 to 0.08)	0.15	0.04 (-0.06 to 0.15)	0.47
2nd trimester	0.008 (-0.04 to 0.06)	0.76	-0.02 (-0.12 to 0.10)	0.78
3rd trimester	0.02 (-0.04 to 0.07)	0.58	0.02 (-0.09 to 0.14)	0.73
entire pregnancy	0.04 (-0.03 to 0.11)	0.30	0.03 (-0.11 to 0.19)	0.71

Unadjusted model	**C-MYC**		**E2F1**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	-0.03 (-0.07 to 0.01)	0.19	-0.001 (-0.08 to 0.08)	0.98
2nd trimester	-0.04 (-0.08 to 0.004)	0.08	-0.10 (-0.17 to -0.04)	**<0.01**
3rd trimester	-0.03 (-0.07 to 0.01)	0.18	-0.14 (-0.20 to -0.07)	**<0.01**
entire pregnancy	-0.09 (-0.15 to -0.02)	**0.01**	-0.23 (-0.32 to -0.12)	**<0.01**
NO_2 exposure				
1st trimester	-0.02 (-0.05 to 0.01)	0.15	-0.05 (-0.10 to 0.004)	0.07
2nd trimester	-0.02 (-0.05 to 0.01)	0.13	-0.08 (-0.13 to -0.03)	**<0.01**
3rd trimester	-0.02 (-0.05 to 0.01)	0.13	-0.10 (-0.14 to -0.05)	**<0.01**
entire pregnancy	-0.04 (-0.07 to 0.002)	0.06	-0.11 (-0.17 to -0.05)	**<0.01**
BC exposure				
1st trimester	-0.06 (-0.09 to -0.02)	**<0.01**	-0.10 (-0.16 to -0.03)	**<0.01**
2nd trimester	-0.04 (-0.08 to 0.006)	0.09	-0.15 (-0.21 to -0.09)	**<0.01**
3rd trimester	-0.03 (-0.08 to 0.01)	0.16	-0.10 (-0.17 to -0.03)	**<0.01**
entire pregnancy	-0.07 (-0.13to -0.02)	**<0.01**	-0.20 (-0.28 to -0.12)	**<0.01**

Appendix 7: The expression of the downstream targets of the miR-17/92 cluster, including TGFBRII and MAPK14, involved in cell proliferation, in association with prenatal air pollution exposure

Table 25 . Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative mRNA expression of the downstream targets TGFBRII and MAPK14, involved in cell proliferation in cord blood. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 μg/m^3 for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 μg/m^2 for BC.

Unadjusted model	**TGFBRII**		**MAPK14**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	-0.02 (-0.09 to 0.06)	0.60	0.08 (-0.03 to 0.19)	0.16
2nd trimester	-0.11 (-0.17 to -0.04)	**<0.01**	-0.16 (-0.24 to -0.08)	**<0.01**
3rd trimester	-0.17 (-0.22 to -0.11)	**<0.01**	-0.18 (-0.25 to -0.10)	**<0.01**
entire pregnancy	-0.26 (-0.34 to -0.17)	**<0.01**	-0.25 (-0.36 to -0.12)	**<0.01**
NO_2 exposure				
1st trimester	-0.04 (-0.09 to 0.01)	0.10	0.004 (-0.07 to 0.08)	0.92
2nd trimester	-0.09 (-0.13 to -0.04)	**<0.01**	-0.13 (-0.18 to -0.07)	**<0.01**
3rd trimester	-0.09 (-0.14 to -0.05)	**<0.01**	-0.13 (-0.19 to -0.07)	**<0.01**
entire pregnancy	-0.11 (-0.16 to -0.05)	**<0.01**	-0.13 (-0.20 to -0.05)	**<0.01**
BC exposure				
1st trimester	-0.13 (-0.19 to –0.07)	**<0.01**	-0.08 (-0.16 to 0.01)	0.09
2nd trimester	-0.15 (-0.21 to -0.08)	**<0.01**	-0.21 (-0.28 to -0.13)	**<0.01**
3rd trimester	-0.09 (-0.16 to -0.02)	**0.01**	-0.12 (-0.21 to -0.03)	**0.01**
entire pregnancy	-0.21 (-0.28 to -0.13)	**<0.01**	-0.22 (-0.32 to -0.12)	**<0.01**

Appendix 8: The expression of the downstream targets of the miR-17/92 cluster, including HIF-1α and CDKN1A, involved in hypoxia and cell cycle arrest, in association with prenatal air pollution exposure

Table 26. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative mRNA expression of the downstream targets HIF-1α and CDKN1A, involved in hypoxia and cell cycle arrest in cord blood. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 μg/m^3 for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 μg/m^2 for BC.

Unadjusted model	**HIF-1α**		**CDKN1A**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	-0.06 (-0.13 to 0.02)	0.11	-0.002 (-0.05 to 0.05)	0.94
2nd trimester	-0.17 (-0.23 to -0.11)	**<0.01**	0.05 (0.01 to 0.10)	**0.03**
3rd trimester	-0.14 (-0.20 to -0.07)	**<0.01**	0.01 (-0.04 to 0.06)	0.70
entire pregnancy	-0.33 (-0.41 to -0.25)	**<0.01**	0.07 (-0.01 to 0.16)	0.10
NO_2 exposure				
1st trimester	-0.04 (-0.09 to 0.02)	0.19	0.020 (-0.01 to 0.05)	0.27
2nd trimester	-0.09 (-0.14 to -0.05)	**<0.01**	0.03 (0.0005 to 0.07)	0.05
3rd trimester	-0.02 (-0.07 to 0.04)	0.51	-0.02 (-0.05 to 0.02)	0.30
entire pregnancy	-0.08 (-0.14 to -0.02)	**0.02**	0.02 (-0.02 to 0.07)	0.33
BC exposure				
1st trimester	-0.16 (-0.22 to -0.10)	**<0.01**	0.05 (-0.001 to 0.10)	0.06
2nd trimester	-0.18 (-0.24 to -0.12)	**<0.01**	0.03 (-0.02 to 0.08)	0.23
3rd trimester	-0.08 (-0.15 to -0.001)	0.05	-0.04 (-0.08 to 0.02)	0.16
entire pregnancy	-0.24 (-0.31 to -0.17)	**<0.01**	0.03 (-0.03 to 0.10)	0.37

Appendix 9: The expression of the downstream targets of the miR-17/92 cluster, including BCL2L11 and PTEN involved in cell death, in association with prenatal air pollution exposure

Table 27. Unadjusted model of the association between $PM_{2.5}$, NO_2, or BC exposure in different time windows of pregnancy and the relative mRNA expression of the downstream targets BCL2L11 and PTEN, involved in cell death in cord blood. The estimates (95% CI) represents the unadjusted change in relative miRNA expression for an increase of exposure of 5 μg/m^3 for $PM_{2.5}$ and NO_2 and an increment of exposure of 0.5 μg/m^2 for BC.

Unadjusted model	**BCL2L11**		**PTEN**	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
$PM_{2.5}$ exposure				
1st trimester	0.01 (-0.08 to 0.11)	0.85	0.007 (-0.07 to 0.09)	0.87
2nd trimester	0.001 (-0.08 to 0.10)	0.98	-0.02 (-0.09 to 0.05)	0.54
3rd trimester	0.04 (-0.05 to 0.14)	0.40	-0.01 (-0.08 to 0.07)	0.80
entire pregnancy	0.06 (-0.09 to 0.24)	0.44	-0.03 (-0.14 to 0.10)	0.68
NO_2 exposure				
1st trimester	0.01 (-0.06 to 0.07)	0.84	0.008 (-0.04 to 0.06)	0.77
2nd trimester	0.003 (-0.06 to 0.07)	0.93	-0.01 (-0.06 to 0.04)	0.60
3rd trimester	-0.05 (-0.11 to 0.02)	0.15	-0.02 (-0.07 to 0.03)	0.36
entire pregnancy	-0.01 (-0.09 to 0.07)	0.72	-0.02 (-0.08 to 0.05)	0.64
BC exposure				
1st trimester	0.06 (-0.03 to 0.16)	0.19	-0.007 (-0.08 to 0.07)	0.84
2nd trimester	-0.01 (-0.10 to 0.09)	0.83	-0.07 (-0.14 to 0.00004)	**0.05**
3rd trimester	-0.004 (-0.10 to 0.10)	0.93	-0.08 (-0.15 to -0.009)	**0.03**
entire pregnancy	0.03 (-0.09 to 0.16)	0.62	-0.09 (-0.18 to 0.002)	0.06