Selenium-gene interactions involving microRNAs

Anabel Maciel Domínguez

Thesis submitted in partial fulfilment of the requirements
of the regulations for the degree of Doctor of Philosophy

Newcastle University

Faculty of Medical Sciences

Institute for Cell and Molecular Biosciences

July / 2012

Declaration

The material contained within this thesis has not been previously submitted for a degree at the University of Newcastle or any other university. The research here has been conducted by the author. All help by other has been acknowledged and all sources of information are indicated in the text.

Abstract

Selenium (Se) is an essential nutrient for health. In mammals Se is incorporated into ~25 selenoproteins in the form of the amino-acid selenocysteine encoded by the UGA codon through a complex interacting with selenocysteine insertion sequence (SECIS) in the 3'Unstralated Region. The selenoproteins have functions in antioxidant defence and redox control, thyroid hormone metabolism and mitochondrial metabolism. Previous scientific work has found that Se also affects a group of downstream targets. The aim of my work is to investigate whether expression of selenoproteins or the downstream targets affected by Se is regulated through epigenetic mechanisms involving microRNAs, a small non-coding RNA species that regulates a gene or groups of genes by binding to the mRNA 3'UTR. Gut epithelial Caco-2 cells were grown in either Se deficient or Se-supplemented medium for 72h. RNA extracted and miRNA expression analysed using a custom-designed human genome V2 Agilent 8x15K array. In addition, global mRNA transcriptome expression was analysed using an Illumina HumanRef-8 v3 microarray.

Se supply increased the expression of thirteen miRNAs and 53 mRNAs the observed differences were confirmed by real time PCR. miR-185 was selected as a further target of investigation because of its high sensitivity to Se. Bioinformatic analysis of the Se susceptible miRNAs and Se sensitive genes was carried out using miRWalk, MicroCosmo, microRNA.org, miRBase and microRNA.org algorithms and Ingenuity Pathway Analysis. This identified miR-185 recognition elements in the Se-sensitive mRNAs for Glutathione peroxidise 2 (*GPX2*), Glutathione peroxidise 3 (*GPX3*), and Selenophosphate Synthetase 2 (*SEPHS2*).

Expression of *GPX2* and *SEPHS2* was altered by miR-188 with a specific anti-miR and affected differently according to miRNA exposure. The data suggests that these mRNAs are targets of miR-185.

In conclusion, the experiment indicates that miRNA expression as regulation of target genes is regulated by Se supply in Caco-2 cells.

Acknowledgments

Foremost all I would like to express my thanks and appreciation to Prof Hesketh for his patient guidance, encouragement, and support especially during though times of my project. I am thankful to have had a successful professor whose enthusiasm towards science has made my research experience stimulating and to have an excellent person as my supervisor.

I would also like to show my gratitude to Prof Dianne Ford for her scientific contributions that pointed me in the right direction. I would also like to thank Dr Hannah Gautrey, Mr Chirs Blackwell, Dr Luisa Wakeling and Dr John Tyson for their patient and resourceful technical as well as personal contribution. This PhD thesis would not have been possible without the help of the Hesketh/Ford lab, where not only have I found good advice but also friendship.

I would also like to acknowledge Dr Daniel Swan for conducting the statistical microarray data analysis, initial bioinformatics analysis and useful advice; to Dr Allison Howard and Dr David Young for letting me use their QRT-PCR facilities and for their guidance with the data analysis; to Dr Andy Verner for his constructive criticism and suggestions; to Dr Tim Cheek for looking after me on bureaucratic matters. I would also like to thank my examiners Dr Ruth Valentine and Prof Bernard Corfe for the time they invested reading my thesis and the insightful questions of my *Viva voce*. I am thankful to the National Council of Science and Technology who provided financial support for this research.

Finally, I am truly in debt with my family, friends, and colleagues who contributed one way or another to this thesis, your love, support, and company have given me the strength to achieve this goal.

Thank you all.

"*If we knew* what it was *we were doing, it wouldn't be called* '*research*,' *would it*?"

Albert Einstein

Table of Contents

List of figures

List of tables

List of abbreviations

3'UTR	3' Untranslated Region
4SU	4-thiouridine
A	Adenine
ACBT	Beta Actin
Ago	Argonaute
APS	Ammonium persulfate
ARE	AU-rich elements
ATP	Adenosine Triphosphate
ATSCs	Adipose tissue–derived stem cells
Aub	Aubergine
BSA	Bovine serum albumin
C	Cytosine
Caco-2	Colonic Adenocarcinoma
CDK	Cyclin-dependent kinases
cDNA	copy DNA
C. elegans	*Caenorhabditis elegans*
CIP	Calf Intestine Alkaline Phosphatase
CLIP	Cross-linking and immunoprecipitation assays
Ct	Cycle threshold
Cys	Cysteine
D. melanogaster	*Drosophila melanogaster*
DEPC	Diethyl Pyrocarbonate
DGCR8	DiGEorge critical region 8
DHEA	Dehydroepiandrosterone
DIG	Digoxigenin
DMH	1,2-dimethylhydrazine
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
dNMP	Deoxyribonucleoside monophosphate
dNTP	Deoxynucleotide triphosphate
DTT	Dithiottreitol
EDTA	Ethylenediaminetetraacetic acid

EF-Sec	Selenocysteine-tRNA specific elongation factor
ERK	Extracellular signal-reductase kinase
FACT	Facilitates nucleosomal transcription
FC	Fold change
FDR	False Discovery Rate
FMRP	Fragile X Mental Retardation Protein
G	Guanine
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GPx	Glutathione peroxidase
GPx1	Ubiquitous cytosolic Glutathione peroxidase Glutathione peroxidase 1
GPX2	Glutathione peroxidase 2 (Gastrointestinal glutathione peroxidase)
GPX3	Glutathione peroxidase 3 (Plasma glutathione peroxidase)
h	Hour/s
HCl	Hydrochloric acid
IPA	Ingenuity Pathways Analysis
IRESs	Internal ribosome protein sites
KO	Knockout
LAMP	Labelled miRNA pull-down assay
LNA	Locked Nucleic Acid
MAPK	Mitogen-activated kinase
MCS	Multiple cloning site
MFE	Minimum Free Energy
min	Minute
mirAGE	miRNA serial analysis of gene expression
miRNA	MicroRNA
M. musculus	*Mus musculus*
mRNA	Messenger ribonucleic acid
MRPS36	Mitochondrial Ribosomal Protein S36
MSP	miRNA specific primer
MT1F	Metallothionein 1F
NaCl	Sodium chloride
NGS	Next Generation Sequencing

nt	Nucleotide
ORF	Open Reading Frame
P bodies	Processing bodies
PAGE	Polacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PIR	Protein Iron Binding
Pol I	RNA polymerase 1
QRT-PCR	Real Time Reverse Transcriptase-Polymerase Chain Reaction
RIN	RNA Integrity Number
RISC	RNA-induced silencing complex
RLC	RISC loading complex
RLM-RACE	RNA ligase-mediated amplification of cDNA end
RLP24	Ribosomal Protein Large 24
RPM	Revolutions per minute
RNA	Ribonucleic Acid
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
S. cerevisiae	*Sacharomises cerevisiae*
SAGE	Serial Analysis of Gene Expression
SDS	Sodium dodecyl sulphate
Se	Selenium
Sec	Selenocysteine
SECIS	Selenocysteine insertion sequence
SeCys	Selenocystine
SELK	Selenoprotein K
SELW	Selenoprotein W
SeMet	Selenomethionine
SBP2	SECIS Binding Protein 2
SEPHS2	Selenophosphate synthetases 2
Ser-tRNA^{Sec}	Serly-tRNASec
siRNA	Small interference RNA

snoRNA	Small nucleolar RNA
SNP	Single Nucleotide Polymorphisms
T	Thymine
T3	Triiodothyronine
T4	Tetraiodothyronine
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TEMED	Tetramethylethylenediamine
TF	Transcription factor
T_m	Melting temperature
Tris	Trimethylsilyl silane
tRNA-^{Sec}	Selenocysteyl-tRNA
TTFIIS	Transcription elongation factor IIS
U	Uracil
UCSC	University of California Santa Cruz
wk	Week
XPO5	Exportin-5

Chapter 1 Introduction

1.1 Introduction

This thesis addresses the broad hypothesis that gene expression can be modified at the post-transcriptional level by micronutrient intake mediated through small non-coding RNA (sncRNA) sequences. It has recently been found that mRNA degradation or reduced protein expression can be regulated by microRNAs (miRNAs) through perfect or imperfect target recognition of 3'UTR sequences. My hypothesis is that since miRNAs are able to target multiple genes there could be a common miRNA-controlled pathway that regulates genes sensitive to selenium supplementation. To address this proposal, I used both bioinformatics and molecular biological approaches to investigate if miRNA expression changes upon Se supplementation and if they are involved in regulation of this pathway. For the bioinformatics approach I have searched for predicted miRNAs targets to those genes using several algorithms: mirna.org, miRBase, and GeneSet2MiRNAand later miRWalk; and genomic data analysis (Ingenuity Pathway Analysis) and in the laboratory I have isolated RNA from cells in culture for miRNA and mRNA microarray analysis and further functional analysis of a miRNA. On the following pages, the role of selenium in biological processes, the mode of action of miRNAs, miRNA target recognition through bioinformatics approaches and the possible role of selenium (Se) responsive miRNAs on changes in expression of Se-sensitive genes will be discussed.

1.2 Regulation of gene expression

Cell function relies on proteins to undertake specific functions towards cell development and differentiation; therefore, the correct process of information transcription from DNA into RNA and translation within the ribosome to encode a polypeptide is paramount for cell function. The process by which the genetic code is used to synthesize proteins and produce the structure of cells is generally known as expression of genetic information and starts with the synthesis of RNA from DNA, and polypeptides from RNA. In general, nucleotides from DNA are decoded to produce linear sequences of nucleotides that can then be decoded into codons (3 nucleotides) that will then be decoded into amino acids forming a polypeptide (Strachan T and P, 1999); gene expression regulation varies from one organism to another; this brief introduction will focus on metazoans. The first step is known as transcription, it occurs mainly in the nucleus of eukaryotic cells where the double DNA helix unwinds and so one strand serves as a template for RNA synthesis in which AMP, CMP, GMP or UMP are added to the free hydroxyl group of the 3' end (transcription begins from 5'-3') (Strachan T and P, 1999). RNA polymerases I and II are responsible for this elongation. Pol I and Pol II elongate DNA by adding an 3' hydroxyl group through deoxyribonucleotide triphosphate (dNTP) downstream of the deoxyribonucleoside monophosphate (dNMP) growing chain to produce a leading strand that could elongate by the addition of nucleotides (Strachan T and P, 1999), but their requirements are different. Pol I has an intrinsic hydrolysis activity, promotes transcript cleavage thanks to its A12.2 subunit without a trans-acting factor and its transcription initiation step comprises: Rrn3, a monomer that interact with the A43 Pol I subunit; a transcription termination factor (TTF-1), a six subunit complex that directs transcription through a specific promoter know as core factor, an upstream binding factor (UBF). Its ability to transcribe nucleosomal DNA is facilitated by the FACT (facilitates nucleosomal transcription) complex that localizes and extracts rDNA during cell proliferation (Schneider, 2012). Cellular transcription by Pol I is the most active and energetically demanding

process of cellular translation and it involves 3 nuclear steps before the ribosome is assembled and transported outside the nucleus: transcription of ribosomal RNA (rRNA) 28S, 18S and 5.8S; the rRNA 5S synthesised by Pol III; and the synthesis of other 78 rRNAs from Pol II. Once these rRNA transcripts encode the rRNA 25S, 18S, 5.8S, 5S, and 78 rRNA proteins (Schneider, 2012).

Pol II requires a TFIIS (Transcription elongation factor IIS) to attenuate potential transcription arrest by activating the hydrolysis activity and RNA cleavage (Schneider, 2012); many protein coding genes are transcribed by Pol II have a TATA box at the promoter element, and undergo capping and polyadenylation after transcription (Strachan T and P, 1999).

During RNA processing, the RNA transcript is spliced, separating the gene coding sequences (exons) form the noncoding segments (introns). This mechanism is facilitated by snRNA U1, U2, U3, U4, U5, and U6. To identify an intron conserved GU and AG dinucleotides, as well as the branch site, need to be identified. In this way, the messenger RNA is formed by discarding the introns and keeping the protein coding region and relatively short untranslated regions (5’ and 3’ UTRs).

Once mRNA is synthetized, it migrates to the cytoplasm where the open reading frame (ORF)is translated into proteins; the 3’ and the 5’ UTRs function to help mRNA-ribosome binding and regulate expression of the gene stabilizing, modulating translation and localizing mRNA (Strachan T and P, 1999). Increasing evidence has discovered that these UTRs play a role in the regulation of gene expression by interacting with AU-rich elements (ARE) through ARE-binding proteins that could either increase or decrease RNA stability recruiting mRNA into processing bodies (P bodies), isolating mRNA from ribosomes.

More recently, the discovery of small non-coding RNA species that recognise have recognition elements commonly on the 3’UTR of the gene sequences have arisen as a new regulatory mechanism of regulation of

gene expression (Carthew and Sontheimer, 2009; Gu and Kay, 2010; Zhao *et al.*, 2011; Barrett *et al.*, 2012).

1.2.1 *Regulation of gene expression by nutrition*

The changes in nutrition that can influence changes in the genome have been studied to find the mechanism behind human evolution and the path towards a healthy lifestyle.

Events like increased protein and omega ω-6 and ω-3 fatty acids intake as food are suggested to be responsible for the development of the *Homo erectus* bigger brain. Another example of gene expression regulation by nutrition, is the Single Nucleotide Polymorphisms (SNP) caused by low essential nutrient intake; furthermore, the availability of methionine or choline food sources could have influenced the methylation of several genes (Babbitt *et al.*, 2011) leading to a biological programming of the organism through evolution to react to environmental changes -such as dietary change- and chromatin modifications so as to ensure an adequate gene expression (Francis, 2011).

Epigenetic mechanisms have been shown to have a role in gene expression through chromatin structures that carry gene expression patterns that can be heritable but do not incur changes to the DNA sequence. These include histone modification (Francis, 2011), DNA methylation, and miRNAs (McKay and Mathers, 2011). Research done in early post-natal nutrition suggests that the exposure to inadequate early nutrition is linked with epigenetics with the presence or absence of illness (Gabory *et al.*, 2011; Wiedmeier *et al.*, 2011).

Micronutrients like the methyl donors, and vitamins pyridoxine, folate, and B_{12}, as well as those nutrients involved in one-carbon metabolism influence DNA methylation: selenium requirements of homocysteine, changes in zinc metalloenzymes from the one-carbon cycle (McKay and Mathers, 2011).

Mammalian transcription has been shown to be also nutrient responsive. For example, effects of serum and growth factors after starvation can be explained by the fact that Pol I transcribes rRNA genes through a complex that is regulated through growth factors like MAPK-kinase and insulin-mTOR pathways (Stefanovsky *et al.*, 2006)

Due to the fact that epigenetic mechanisms linked to nutrition can be reversible (Gabory *et al.*, 2011), further is needed research in order to find the appropriate exogenous events that would lead to a healthy organism.

Environment changes like nutrient availability, trigger mechanisms where the cells have to respond to external stimuli by controlling different stages of the transcription so that gene activation responds to extracellular signals (Weake and Workman, 2010).

1.3 microRNAs (miRNAs)

miRNAs are part of a small non protein coding RNA family characterized by their ~20-30nt long strands and their regulation of gene expression through the formation of an RNA interference silencing complex (RISC). The RISC consists of a group of proteins from the argonaute (Ago) family (Kim *et al.*, 2009a) and RNAs and it interferes with genome function through mRNA degradation or mRNA translational repression (van den Berg *et al.*, 2008).

Small interfering RNAs (siRNAs), microRNAs (miRNAs), and piwi interacting RNAs (piRNA) are closely related categories of small non-coding RNAs (with a 21-31 nt-long sequences); from those, only siRNAs and miRNAs are dependent on Dicer enzymes for their biogenesis; in both cases as their single stranded versions are loaded into the RISC complex to repress mRNAs (Carthew and Sontheimer, 2009). Briefly, processing of piRNAs are through encoded PIWI proteins (Kim *et al.*, 2009a); a piRNA locus produces an antisense piRNA transcript that will bind to Aubergine (Aub) to form the Aub-piRNA complex; this complex will cleave to their sense targets to degraded them and then loaded onto Ago3 to produce the Ago3-piRNA complex. Once Ago3-piRNA complexes formed, it is capable of cleaving antisense transcripts to degrade and process sense mRNAs that will be loaded again onto Ago3; this way of transcript degradation, is called “ping-pong” (Choudhuri, 2010; Castañeda *et al.*, 2011; Röther and Meister, 2011). On the other hand, miRNA and siRNA biogenesis have a similar mode of action, but differ from one another as siRNAs originate from a double stranded RNA while miRNAs come from a single stranded RNA loop (Carthew and Sontheimer, 2009). mRNA repression by double stranded siRNAs processed by Dicer (Carthew and Sontheimer, 2009) involves, in most cases, an endonucleotic cleavage (Ago 2 action) (Jackson and Standart, 2007); in mRNA depletion by miRNA, the RISC is formed by Ago1, and m^7Gppp decapping follows deadenylation and leads low mRNA levels (Standart and Jackson, 2007).

A
siRNA
B
miRNA
C
piRNA
Pol II
Nucleus
Double stranded
RNA precursors
Single stranded
RNA loop
Single stranded
RNA precursors
Export
XPO5
Export
Export
Cytoplasm
Dicer
TRBP
Dicer
TRBP
Aub-piRNA
complex
Aub/Piwi
Ago3
"Ping pong"
cycle
Aub/Piwi
Ago3
Ago-3 piRNA
complex
Aub/Piwi
Ago3
AGO-2
RISC
AGO-2
RISC
3′
5′
3′
5′

Figure 1 Biogenesis of small RNA species.
A) siRNA biogenesis. Endogenous siRNA are transcribed by inverted DNA repeats to form double stranded hairpins. Once transported into the cytoplasm, this long double stranded RNA is processed by Dicer into a 22nt double strand which will then be loaded into the RISC by its antisense strand. B) miRNA biogenesis is explained in more detail in Figure 2; briefly miRNAs are transcribed as pri-miRNAs by Pol II, cropped so that they fold into a 70nt single strand (pre-miRNA), which is then transported outside the nucleus and cleaved by Dicer to form the miRNA-miRNA* duplex; the miRNA is then loaded into the RISC. C) piRNA precursors are transcribed from intragenic elements which are a mixture of single strand sense and anti-sense sequences. An Aub-piRNA complex is formed when anti-sense piRNAs bind to Aub/Piwi; the sense Aub/Piwi complex transcript will then bind to Ago3 a forming sense piRNA (Ago3 piRNA complex) that will recognise an antisense transcript, creating anti-sense piRNAs susceptible to bind to Aub/Piwi to cleave complementary transcript targets and continue the “Ping-pong” cycle.

MicroRNAs are highly conserved 20- 22 nucleotide (nt) long non-coding RNA which constitute 1-5% of the known genome (John *et al.*, 2004). In the majority of known cases, miRNAs regulate gene expression by imperfect base pairing to mRNAs through RISC and further reduction of protein translation due to increased mRNA degradation (Jackson and Standart, 2007; Cannell *et al.*, 2008). An exception of this is miR-10a which enhances translation of ribosomal proteins through a 5'UTR interaction (Orom *et al.*, 2008).

1.3.1 *miRNA biogenesis in mammals*

The biogenesis of miRNAs differs from other small endogenous RNAs that act through a silencing complex mechanism since they originate from a fold-back hairpin structure (Brodersen and Voinnet, 2009). There are eleven genes involved in miRNA biogenesis (DROSHA, DGCR8, RAN, XPO5, DICER, AGO1, AGO2, HIWI, GEMIN3, GEMIN4, and TRBP) (Kim *et al.*, 2010). Of these, Drosha and Dicer can be found almost in every cell. In tumour cells, Dicer can only be found in the cytoplasm unlike Drosha which can also be found in the nucleus although not as abundant as in the cytoplasm. Cytoplasmic Dicer and Drosha have a higher concentration in colon tumours cells compared with non-cancerous cells (Tchernitsa *et al.*, 2010b). Deletion of Dicer and DGCR8 has been linked to a deficiency of cellular self-renewal and differentiation (Shimono *et al.*, 2009), several pathological states are characterized by a decrease in mature miRNAs probably, because miRNAs processing components have been compromised (Suzuki *et al.*, 2009).

As shown in Figure 2, miRNAs are most commonly transcribed by RNA polymerase II (Pol II) to their precursors (pri-miRNAs), then they are capped by 7MGpppG and polyadenylated (Saito *et al.*, 2009). Hairpins are then processed in the nucleus by a drosha-DGCR8 (DiGEorge critical region 8) complex also called PASHA; here, the pri-miRNA is excised by Drosha once DGCR8 has localized the pri-miRNA stem to then bind at 11nt from it (Kohler and Hurt, 2007; Carthew and Sontheimer, 2009). This

complex involves enzymes that separate double stranded RNAs from the DEAD box: p68 and p72 (or DDX17) (Suzuki *et al.*, 2009).

Processed pri-miRNAs are cleaved 11 base pairs from the loop by Drosha which is positioned by DGCR8 to cut the pri-miRNA into ~70nt hairpins to form the pre-miRNA. These pre-miRNAs are next transported to the cytoplasm by a karyopherin exportin cargo, exportin-5 (XPO5), which recognizes a nucleotide motif (Kohler and Hurt, 2007) of the pre-miRNAs. Once in the cytoplasm, an RNA III enzyme, Dicer, cuts the pre-miRNA into two single strands, the miRNA (the guide strand) and miRNA*(Bartel, 2004; Filipowicz *et al.*, 2008; Ghildiyal and Zamore, 2009; Winter *et al.*, 2009). This cytoplasmic cleavage leaves 2 unpaired nucleotides at the 3' end and these are the target of the argonaute (Ago) PAZ domain; alternatively, the phosphates from the 5' are targeted by the MID domain of the same protein (Hock and Meister, 2008). *MID has also a protein-protein interaction function.

For maturation of metazoan miRNA, a single stranded miRNA has to bind to an RNA induced silencing complex (RISC). To this end, it has been suggested that for thermodynamic stability, the strand with unstable base pairs (GU rather than GC) at the 5' is loaded into an Ago to become a mature miRNA while the other single stranded miRNA (miRNA*) is degraded (Brodersen and Voinnet, 2009) by nucleases. The 5' strand is usually the more conserved, thus likely to bind the 3'UTR mRNA. In fact, a perfect 5'end miRNA-3'UTR of mRNA pairing is found in *Drosophila* (Lai, 2002) and plants. The RISC is formed by Dicer and Argonaute 2 (Ago 2) - the only endonuclease protein from the Argonaute family located on a different chromosome (8 rather than 1) (Hock and Meister, 2008) with an incision capacity (Li *et al.*, 2006; Hock and Meister, 2008).

Dicer, Ago 2, and a double stranded RNA binding protein (TRBP or PACT) (Kim *et al.*, 2009a) form the RISC loading complex (RLC) that will assemble the single stranded miRNA (Hock and Meister, 2008; Ghildiyal and Zamore, 2009) after Dicer has digested the miRNAs into the

RISC in which mRNA is degraded by miRNA-mRNA binding, deadenylation, or translational repression (Winter *et al.*, 2009).

The origin of miRNA promoters is not well understood. However, chromatin immunoprecipitation chip (chiP-chip) assays have suggested that miRNA transcript start sites may be shared by the introns of their coding genes or that they might have their own promoter (Corcoran *et al.*, 2009) and behave like a protein coding Pol II transcript as transcription factors, transcription co activators and chromatin modifications regulate these miRNAs (Cannell and Bushell, 2010). Other miRNAs are originated from intronic regions (Hebert and De Strooper, 2009); when miRNAs are processed from short intronic hairpins called mirtrons, these function as pre-miRNAs; splicing takes the place of nuclear processing by Drosha, from miRNA transport to the cytoplasm onwards, mirtrons follow the miRNA processing pathway (Slezak-Prochazka *et al.*, 2010).

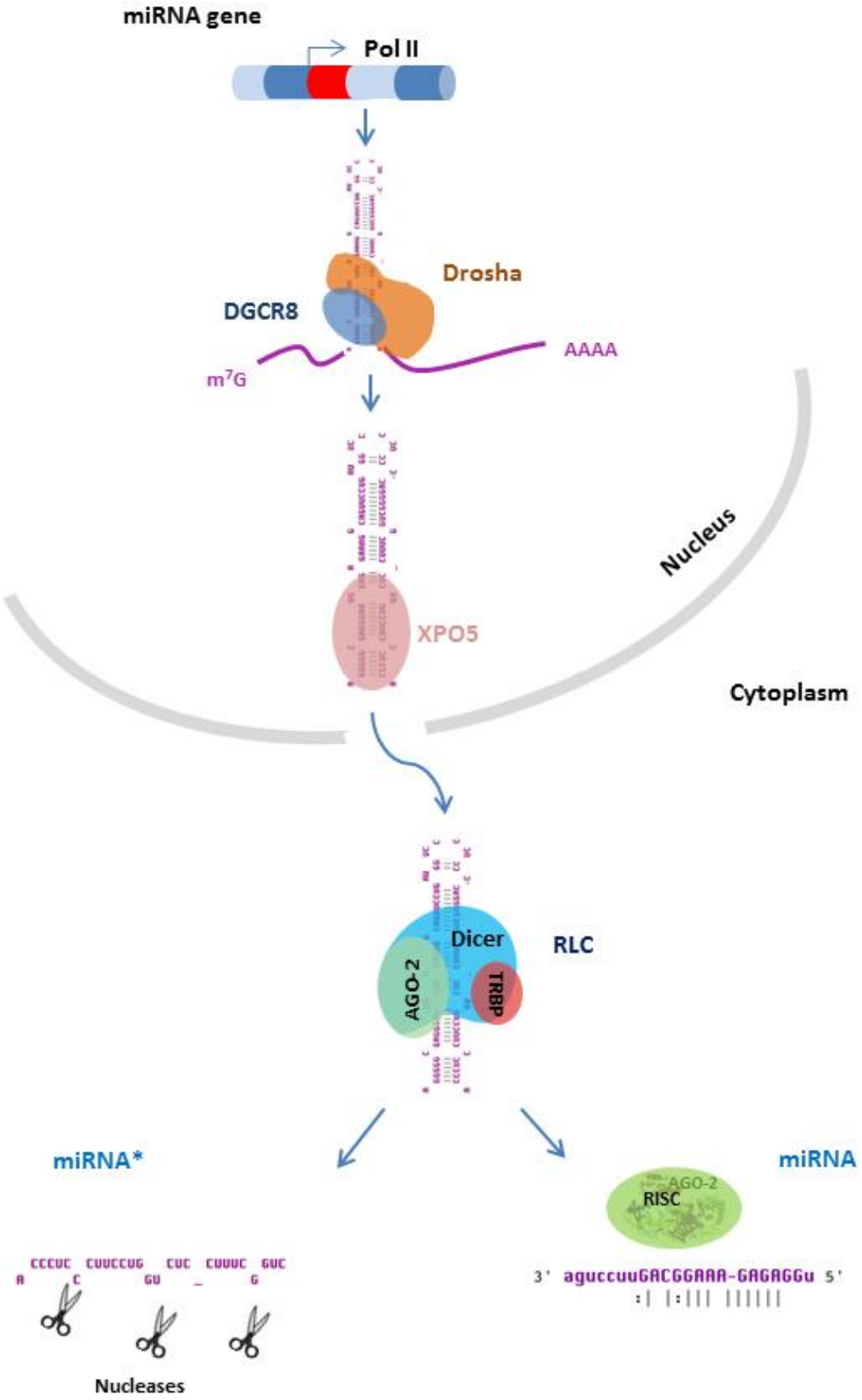


Figure 2 miRNA biogenesis.
miRNA biogenesis from RNA polymerase II starts with transcription into a hairpin structure, pri-miRNA. This is followed by capping and polyadenylation, and then the double stranded pre-miRNA is excised by Drosha The ~70nt strand is then transported outside the nucleus by XPO5. Once in the cytoplasm, Dicer will process the pre-miRNAs into 2 strands, the guide strand and the * strand. The former strand is loaded into RISC whereas the * strand is generally degraded.

1.3.2 *Gene expression regulated by miRNAs*

For miRNAs to regulate gene expression, they need a ribonuclear protein complex, RISC. As described in the miRNA biogenesis section, for miRNA maturation, the5' strand of the miRNA is recognised by Ago2 and loaded into the RISC; this protein is responsible for mRNA binding (Liu *et al.*, 2012) and the slicer activity that leads to RNA interference (Kawamata and Tomari, 2010). The effects of miRNAs on mRNAs, once the former becomes mature by loading into RISC, are on mRNA degradation or translational repression.

Their mechanism of action is still under investigation but most researchers agree that miRNAs affects are brought about by these mechanisms one of the following (Gu and Kay, 2010):

a) Site specific cleavage. In this case a miRNA cleaves its target to a nearly perfect or perfect complementarity resulting on RNA interference; Ago2 is responsible for mRNA slicer activity.

b) mRNA degradation, associated with mismatches on the miRNA-mRNA binding. This mechanism has been suggested to work thorough ARE, a previously documented decay element (Zhao *et al.*, 2011; Vasudevan, 2012) that mediate an accelerated mRNA degradation.

c) Translational inhibition that happens when a an imperfect match occurs in the miRNA-mRNA interaction, and this can potentially be triggered by deadenylation (Djuranovic *et al.*, 2011). Another factor that could facilitate mRNA degradation is disintegrating the poly A binding protein (PABPC) and removing the polyA tail leading to exonucleolytic degradation (Pasquinelli, 2012).

miRNAs can inhibit mRNA translation by interacting with eIF4F complex, an eukaryotic translation-initiation factor complex formed by eIF4E, eIF4G, and eIF4A, that cap the mRNA structure protecting the

mRNA from degradation (Huntzinger and Izaurralde, 2011). mRNA deadenylation is another proposed mechanism of action of miRNAs, and this can occur either at initiation (Liu, 2008)of or post-initiation (Cannell *et al.*, 2008). Reduced protein synthesis may result from mRNA decapping, deadenylation, or degradation in P Bodies. Alternatively, elongation distortion by active mRNA translocation, blocked polypeptide synthesis (ribosome “drop off”), or other undefined mechanisms may be involved in miRNA-mediated reduction in protein expression post-initiation (Standart and Jackson, 2007; Cannell *et al.*, 2008; Kong *et al.*, 2008).

Elongation repression is suggested to be by RISC interactions. Proposed models include the competition between RISC and elF4E, in which case RISC would bind the mRNA 5’ cap structure while Apo2 replaces elF4E tryptophans with phenylalanines (Carthew and Sontheimer, 2009). Translational repression is thought to be due to mRNA destabilization by poly (A) tail deadenylation initiated by either RISC or GW182.

miRNAs interact with cis-regulatory elements as transcriptions factors (TF) do and this might be why the more transcription factors gen, the higher chance a miRNA can have of controlling that gene. (Su *et al.*, 2010).

Although miRNAs are thought to down-regulate the expression of genes (due to the RISC) overexpression of genes has also been registered. It has been hypothesized that a saturation of the sh/miRNA nuclear export machinery is responsible; cell processes in which small RNA transport machinery or RISC saturation is suggested to occur are cell cycle and oncogenesis (Khan *et al.*, 2009).

Apart from their role in RNA interference, miRNAs are also thought to postransciptionally upregulate their targets (Krol *et al.*, 2010; Vasudevan, 2012) by the relief of repression from the miRNA which occurs when Ago2 and GW182 cannot interact properly (Vasudevan, 2012).

1.3.3 *miRNA target recognition*

miRNA-mRNA interactions are due to base pairing. In plants, this interaction includes perfect or almost perfect base pairing to the mRNA (Meng *et al.*, 2011), whereas in animals, this interaction permits mismatches. Even when mismatches or wobbles are found (mostly in the central region of a miRNA), protein expression is downregulated suggesting that miRNAs can anneal to their target without the need of the 100% complementarity required by other small RNA species. More research is required to understand miRNA mechanisms of action and roles over gene expression regulation, but it has been suggested that protein degradation rather than mRNA degradation can be the predominant mechanism, (Kong *et al.*, 2008) as protein synthesis is reduced at a translational level. miRNA mediated cleavage usually involves targeting the 3'UTR of the mRNA (Filipowicz *et al.*, 2008), which is the suggested miRNA mechanisms of action is due to the 3'UTR since it is the most common binding site for this small RNA species in fact this statement is often within the first paragraph on any miRNA publication. However, these interactions have also been found within the 5' UTR or ORF. It is likely that 3' or 5' UTR are more functional sites since the energy demanded for these particular sites requires a weak base pairing (Lee *et al.*, 2009), but experiments in HeLa cells have also indicated the mRNA 5'UTR as a cleavage site for mediating mRNA degradation. On the other hand, ORF target interactions offer repression when the mRNA (susceptible to miRNA cleavage) has been translated.(Lytle *et al.*, 2007). miRNA Recognition Elements (MREs) on the ORF are reported as less effective than the 3'UTR MRE; however, genes that are most commonly repressed have several ORF MRE (Schnall-Levin *et al.*, 2011). In fact, RISC that targets mRNAs at the ORF or the 5'UTR might ensure message repression after ribosomal translation, as these sites are less affected by decodification. Either way, a MRE where an mRNA has a complementary Watson-Crick base pairing to the seed region of an miRNA is needed for mRNA-miRNA interaction.

The understanding of mRNA target recognition began with the observation of multiple complementary sequences at the mRNA 3'UTR from lin-4 RNA and lin-7 RNA (Brodersen and Voinnet, 2009). From that fact, it has been thought that an expressed miRNA is able to regulate several proteins during translation.

As several mechanisms of miRNA target recognition have been identified, it is important to spot the features that allow mRNA-miRNA cleavage. For instance, when the mRNA is in a secondary structure by means of intermolecular base pairing, miRNA cleavage will result in energy consumption to open the secondary structure. (Jackson and Standart, 2007). Duplex formation, however, represents energy gain.

miRNA binding within the 3'UTR is more efficient if the binding site is located at least 15nt far from the stop codon. If the UTR binding site is long, the binding site is more likely to bind away from the centre (Grimson *et al.*, 2007). Another factor that would facilitate miRNA annealing to at mRNA is the presence of AU-rich sites located next to the binding site (Sun *et al.*, 2010),although RNA-binding proteins that interact with AU-rich element of the target mRNA protect it from degradation (Pasquinelli, 2012).

1.3.4 *Seed region*

This 2-7nt long region is located at the 5' end of the miRNA sequence. It was first discovered in *D. melanogaster* and is the section of the strand that is responsible for the most conserved binds (binding site similar among human, mouse, rat, dog, and chicken genome). However, non-conserved sites can still function as a result of competition between a more highly-expressed but less conserved miRNA end with a higher fidelity match to the target but expressed at lower level. miRNA binding beyond the seed region have still been recognized at the 5'UTR or codon mRNA region (Lytle *et al.*, 2007).

Evidence of conserved interactions (and moreover, miRNAs' faculty to regulate multiple mRNAs) matching the seed region was supported by Lim, L.P., et at., 2005 (Lim *et al.*, 2005), who transfected Hela cells with miRNA-124 and a mutated 5-6 region version and obtained different sets of downregulated genes from the two assays.

A conserved seed region is often recognized by mRNA that contains adenosines (Lewis *et al.*, 2005). In fact, when multiple seed sites are found is the same 3'UTR, the one with the highest AU content has the highest cleavage possibility (Long *et al.*, 2008).

A suggested A-form helix conformation has been proposed by Bartel, et al (Bartel, 2009). This model suggests that nucleotides arranged in an A-form helix would need to be 7-8nt long miRNA segments to ensure effectiveness in mRNA targeting; a longer strand will result in large loops while a shorter single strand would not have sufficient specificity. Different miRNA-mRNA binding models have been suggested due to miRNAs flexibility to match imperfectly to their targets, there are different types of cleavage sites:

a) Canonical. Specific Watson-Crick match of 7-8 nucleotides (nt) in the seed region. Ranked by microarray analysis as the most efficient interaction for miRNA mediated mRNA degradation (Bartel, 2009) potentially in the 5-6 positions (Lim *et al.*, 2005).

b) Dominant seed sites. Complementary at the 5' site, but allows mismatches at the 3' end.

c) Atypical sites. Sites at 3' miRNA could match its target if at least 3-4 contiguous short stretches of nucleotides from a 13-14 nt long strand, matching the mRNA poly (A) tail. A loop between the matches and mismatches is formed.

d) Conserved sites. Sites with perfect base pairing at the 3' miRNA but have mismatches or loops at the 5'.

On a miRNA seed pairing, minimum energy is lost during cleavage according to the minimum free energy model (MFE) which is based on the

total energy that a miRNA needs to bind an mRNA; and is measured as an energy-based score which relates the gained energy as an mRNA is bound and the lost energy result of mismatches (Chen *et al.*, 2008).

1.3.5 *Non-seeded sites*

Although, as noted, most conserved miRNAs bind to the 3'UTR of their target genes, there are known instances of miRNAs targeting the protein coding sequence. This view is supported by computational miRNA predictions in *Drosophila melanogaster* (Stark *et al.*, 2007) and experimentally validated data from HeLa Cells (Lytle *et al.*, 2007), and let-7 (showing three repressing human mRNA- miRNA) (Brodersen and Voinnet, 2009), that interact in the open reading frame (ORF) (Mazière and Enright, 2007).

Among other species, exemplified in *C. elegans*, lin-4, let-7, and lys-6, miRNAs miR-10 in humans and AGO1-miRNA complexes in *D. melanogaster* (Brodersen and Voinnet, 2009) functional cleavage sites outside the seed region are commonly found.

Also, pairing at the 13-16nt of the miRNA 3' still confirmed alignment efficiency after microarray analysis; efficient target recognition is shown at 8nt site. This sites lose efficacy if located 15nt far from the stop codon (Grimson *et al.*, 2007).

1.3.6 *miRNA conservation*

miRNAs that occur in multiple species are referred to as conserved miRNA. Conserved miRNAs generally bind to their gene target sequence outside the ORF, and generally in the 3'UTR. This characteristic is used when predicting miRNA target recognition as comparison among species is used to retain the most positive matches as it is thought that

orthologous genes could lead to anticipation of false negatives in miRNA predictions (Norden-Krichmar *et al.*, 2007). However, miRNA conservation should not be taken as criteria of exclusion of possible mRNA target; is so, valuable miRNAs cleavage could be left out in cases of species that not necessarily present similar genes to other species (not conserved).

miRNA mediated cleavage usually involves targeting the 3'UTR mRNA (Filipowicz *et al.*, 2008) but experiments in HeLa cells have also indicated the mRNA 5'UTR as a cleavage site for mediating mRNA degradation. On the other hand, ORF target interactions offer repression when the mRNA (susceptible to miRNA cleavage) has been translated (Lytle *et al.*, 2007). In fact, RISC that targets mRNAs at the ORF or the 5'UTR might ensure message repression after ribosomal translation, as these sites are less affected by decodification.

1.3.7 *mRNA target multiplicity*

The percentage of the interactions between a single miRNA interacting with an mRNA is a roughly 10-20% (Cannell and Bushell, 2010). However, translational silencing can be enhance when the mRNA is cleaved by several miRNAs (Jackson and Standart, 2007; Brodersen and Voinnet, 2009; Schnall-Levin *et al.*, 2011; Subramanyam and Blelloch, 2011). Moreover, this target multiplicity is maximized when the miRNA target sites are close to each other; this phenomenon is called "cooperativity"(Grimson *et al.*, 2007). In miR-124, an interspace of 34nt between the two miRNA signals still permits this cooperativity, while 19 to 56nt does not. According to this principle, a stronger downregulation could come from different miRNAs closely targeting the same mRNA (Grimson *et al.*, 2007; Saetrom *et al.*, 2007). The idea of different silencing complexes (belonging to the same or multiple miRNAs) confers a stronger downregulation relationship. This is because during translation, the mRNAs cleaved by miRNAs could be decoded by the ribosome and therefore dissociates the silencing complex. Disrupted silencing

complexes are frequent in the 15nt 3'UTR target sites (Saetrom *et al.*, 2007).

1.3.8 *mRNA network regulation by a single miRNA*

miRNA target prediction observes target multiplicity even within the same mRNA sequence which has been previously documented where a miRNA or miRNA network able to control several genes (Grimson *et al.*, 2007; Saetrom *et al.*, 2007; Flynt and Lai, 2008; Pitto *et al.*, 2008; Subramanyam and Blelloch, 2011). Single miRNAs that regulate several genes have been suggested to have roles in cancer processes (miR-19a; apoptosis, miR-14; developmental timing, lin-4, lin-28, miR-172 (Yeung *et al.*, 2005)) by targeting one or several genes. On the other side of the argument, stronger mRNA repression is observed when multiple miRNAs regulate a single target gene. Moreover, groups of miRNAs that as a whole target group of genes can be identified, consistent with an overall model of regulation at the level of pathways.

The existence of families or clusters of miRNA that regulate genes; example of this is the mir-17-92 family which have MRE for different protein coding genes (Juhila *et al.*, 2011) and has been identified as oncogenes (Olive *et al.*, 2010) or regulator of the of is has also been proposed particularly in relation to cellular processes and cancer as miRNA clusters seems to be needed in the regulation of cellular complexes.

Members of an miRNA family are identified with same number but differing with a different lower case letter, they share a common seed region or a location in gene clusters present in different chromosomes (Park *et al.*, 2008).

Examples include the miR-200 family (found to be expressed in the carcinogen epithelium tissue), which represses ZEB1 and ZEB2 mRNA (responsible for interleukin 2 expression repression) mRNA; the miR-302 family responsible for stem cell growth (Lin *et al.*, 2008), and the

miR-16 family, which regulates cellular proteins responsible for cell division and cell growth (Linsley *et al.*, 2007; Zhang *et al.*, 2009).

The miR-183-96-182 cluster has been shown to control the entire insulin signalling pathway whereas miR-375 alone, regulates only insulinemia (Zhang *et al.*, 2009) providing an example of an effect of a miRNA network that contrasts with the more limited effect of a single miRNA with the network.

Transfection of miR-124 and miR-1 in HeLa cells demonstrated that a single miRNA downregulates several genes.(Lim *et al.*, 2005). miR-24, expressed in brain tissues, down regulated 174 genes transfected HeLa cells. The response correlated with those expressed in lower levels in the brain than any other tissue. In agreement, miR-1, found in heart and skeletal muscles, preferably down regulated genes present in those tissues.

1.4 Bioinformatics approaches

The complexity of miRNA gene interactions introduces the opportunity to use interdisciplinary approaches involving computer science, mathematics, and statistics to identify possible mRNA targets.

The prediction of miRNA targets has been possible using computational approaches in which miRNA target recognition has been taken into account from several bioinformatic tools in order to design an efficient algorithm. Different miRNA prediction programmes use different algorithms to predict mRNA-miRNA target recognition interactions. These include the following categories: a) Searching for complementary sites which also considers thermodynamic calculations; b) Thermodynamic based algorithms, calculate and compose the free energy of predicted binding interactions; c) Databases based on microarray expression studies or gene expression (Bartel, 2004).

1.4.1 *Bioinformatic tools*

Computational predictions are often used to microRNA Recognition Elements (MREs) within an mRNA sequence. These algorithms look for miRNA-mRNA complementarily and give a high score according to the thermodynamics, matching miRNA seed region, conservation throughout species.

For this project we used GeneSet2MiRNA (which collects information from 11 algorithms), microrna.org, miRBase (which are based on complementary and conserved binding predictions) and Ingenuity Pathways Analysis (which searches for protein-protein and protein-miRNA interactions) to predict miRNAs. As these miRNAs are being predicted, there is no published information about their interactions. Therefore, by comparing the results from these bioinformatic tools that employ different algorithms and finding an overlap on mRNA-miRNA interactions, the aim was to increase the probability of finding a network and validate it experimentally.

Microrna.org is a bioinformatic tool based on a modification of the miRada algorithm, which focuses on almost perfect 3'UTRmRNA-5'miRNA complementary sites and analysed thermodynamically (Mazière and Enright, 2007). The microRNA library is originated from a mammalian microRNA expression based on small RNA library sequence data collected from human and rodent tissues and cell lines (Landgraf *et al.*, 2007). The genome library from *Homo sapiens* (172), *Mus musculus* (64), and *Rattus norvegiocus* (16) was downloaded from the University of California Santa Cruz (UCSC) genome browser database (Kuhn *et al.*, 2007).It scores to the potential alignment by allowing mismatches in the seed region, but seeking complementary binding in the gene from mature miRNAs and considering mRNA multitargeting. For a complementary binding, a miRNA mature and miRNA target conservation among species are needed; the conservation is filtered by PhastCons. The first stage to

determine miRNA-mRNA binding is looking for A:U and C:G interactions, it scans a file with the genome of interest and the miRNA sequences

Microrna.org scans the possible miRNA-gene interactions from two files, one containing miRNA sequences (file1) and the other containing RNA/DNA sequence (file2), in two stages. The former stage includes reading sequences from files 1 and 2 to find A:U and G:C complementary sites. The latter stage will then score the alignment by its thermodynamical role.

miRBase searches for complementary sites (ranking them from 0 to 100) in the Ensembl Genome library of 37 species. It also works with the miRanda algorithm and contains miRNA sequences gathered from 58 species and seeks for conservational pairing. It contains miRNA and miRNA* sequences; which corresponds to an miRNA mature sequence that is generally loaded into the RISC (miRNA), and the passenger strand (miRNA*) which is commonly degraded by endonucleases, but has also shown to play a role in gene expression regulation (Okamura *et al.*, 2008); ~60% of the mature miRNAs within the database are experimentally validated (Griffiths-Jones *et al.*, 2008).

Features when predicting miRNA-mRNA interactions include: perfect seed pairing, mRNA conservation at least between 2 species at the 3'UTR, and thermodynamic interactions are measured by Vienna RNA folding routines (Griffiths-Jones *et al.*, 2008).

GeneSet2MiRNA is as a web-tool that takes a list of genes (by accession number or gene name) and computes it with miRecords, a website that contains documented validated and predicted targets from 9 species. The later targets are collected from 11 algorithms; the GeneSet2MiRNA score is based on the number of commonly predicted miRNA hits (The University of Minnesota and Houston, 2008; Antonov *et al.*, 2009).

Ingenuity Pathways Analysis (IPA) is a software rather that an algorithm that provides a visual analysis of gene interactions and builds pathways of gene interactions in health and disease. Biological and molecular scientific publications from *Homo sapiens*, *Mus musculus* and *Rattus novergicus* data are the source of IPA's library; it also supports 9 other species. The building structure pathway includes information from scientific articles and the chemical and physiological behaviour of the components (Ingenuity, 2008).

miRWalk was created recently from modifying miranda algorithm, but it also compares its predictions with other well established bioinformatics tools and includes experimentally validated and documented miRNAs-mRNA binding (Dweep *et al.*, 2011).

1.4.2 *Finding miRNAs through "hands in" laboratory techniques*

Although these miRNA computational predictions give different scores according to the likeliness of an miRNA-mRNA cleavage, the rate of a false positive miRNA-mRNA interaction is between 30-50% (Yokoi and Nakajima, 2011), this might be because the partial complementarity to the mRNA is not enough for the binding to occur or because several miRNA characteristics are not taken into account for this forecast: miRNA tissue specificity, cell compartment of the mRNA target.

Several experimental approaches have been suggested, miRNA target cloning through RNA ligase-mediated amplification of cDNA end (RLM-RACE), QRT-PCR for generally recommended when identifying a few miRNAs, Labelled miRNA pull-down assay (LAMP), Photoactivatable-ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) (Jin *et al.*, 2010) luciferase reporter assays

- miRNA target cloning through RNA ligase-mediated amplification of cDNA end (RLM-RACE), a microarray based technology designed to identify miRNA-mRNA binding events in vivo by ligating the 5’ monophosphate ends at the 3’ of the cleaved mRNA end, polyA+ to an RNA oligonucleotide adapter. Once ligated, the RNA is synthesised into a singles stranded cDNA by reverse transcription and generating the cDNA second strand with an homologous RNA adapter sequence for the promoter of the bacteriophage T7 as forward primer and a gene specific reverse primers designed followed by PCR amplification, cDNA template removal, cDNA fragmentation and microarray hybridization (Franco-Zorrilla *et al.*, 2009). Because miRNA and siRNA contain ligation-competent 5’ monophosphate when the RISC bind to the small RNA, it does not distinguish between these two snRNA species (Jin *et al.*, 2010). It is generally used for identifying plant miRNA targets (Franco-Zorrilla *et al.*, 2009; Llave *et al.*, 2011)

- Luciferase reporter assays, a bioluminescence technology that clones and isolates luciferases to construct a reporter assay that would catalyse light emitting reactions by producing a photon when luciferin is catalysed into oxyluciferin (Thorne *et al.*, 2010).

- Labelled miRNA pull-down assay (LAMP) in which miRNA biogenesis is mimicked by means of inserting a digoxigenin (DIG) labelled pre-miRNA and once this interact with the endogenous DICER, the DIG-miRNA-mRNA complex can be precipitated and analysed by QRT-PCR;

- Photoactivatable-ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) an adapted cross-linking and immunoprecipitation assays which crosslinks RBN proteins facilitated by 4-thiouridine (4SU) labelled cellular transcripts to detect miRNA targets; cross-linked 4SU residues are the sequenced to recognize the T to C transitions and identify RBN binding, the downside of this technique is that 4SU labelling could

be affected by neighbour aminoacids (Hafner *et al.*, 2010), miRNA profiling through PCR platforms, microarrays,

The most commonly used experimental validation of *In silico* miRNA-mRNA interactions are the following approaches:

a) Microarray. This method screens the expression of multiple mature miRNAs based on hybridization of test RNA samples to arrays of miRNA probes (Yin *et al.*, 2008; Kong *et al.*, 2009). Several commercial microarray platforms are available and they all produce consistent functional pathways according to a study in mouse kidney where mice treated with and without aminoacids to determine by Affymetrix, Applied Biosystems, Agilent, and GE health microarray platforms the pathways found by Ingenuity Pathway Analysis from the genes differentially expressed; showed that the four platforms produced an overlap of 70-80% on functional pathways from the top 5 ranked pathways (Li *et al.*, 2009) which in reality gives limited data of the specific function of a set of genes as the listed functions include tumour, tumorogenesis, cancer, neoplasia, etc. Therefore, more specific comparisons are needed especially towards miRNA expression. At the time of carrying out the experimental work, described in this thesis, the Agilent platform offered the option to customize the chip probes and it had greater sensitivity for detection of individual member of miRNA families or clusters at it showed high correlations with QRT-PCR when comparing Agilent Human miRNA Microarray 1.0; Exiqon miRCURY Locked Nucleic Acid (LNA) microRNA Array, v9.2; and Illumina Sentrix Array Matrix 96-well microarray expression profiling (Git *et al.*, 2010).

Human genome V2 Agilent platform is a one colour microarray consisting of a short probe which hybridizes the target miRNA and contains a hairpin loop that will join the synthesized sequence to the miRNA, this is to standardize all the miRNAs Tm. Uses the miRNAs sequences from Sanger 10.1.

b) Massively parallel/next-generation sequencing (NGS). miRNAs have also been detected through next generation sequencing (Landgraf *et al.*, 2007; Sato *et al.*, 2009; Git *et al.*, 2010) where the small non-coding RNAs are measured using their mature sequence independently, groups of miRNAs that share the mature sequence, and miRNA precursors which are either sharing mature sequences or organized in clusters. (Landgraf *et al.*, 2007)

c) Northern blot analysis. This method is used to analyse and quantify small RNAs. It verifies miRNA-mRNA target stabilizing hybridization by building a methylene bridge between the ribose and the nucleotides with LNA modified probes. From test samples, RNA is electrophoresed through a denaturing polyacrylamide gel to membrane and probed using specific labels miRNA probes (Castoldi *et al.*, 2006; Hino *et al.*, 2008; Kuhn *et al.*, 2008; Várallyay *et al.*, 2008).

d) QRT-PCR. This approach involves reverse transcript followed by PCR amplification using primers specific to the miRNA to be quantified. Due to miRNA sequences being short, several approaches like polyadenilation, LNA technology and stem loop miRNA primers have been designed, to reverse transcribe the miRNAs: and technologies like LNA that needs modified probes to stabilise the ribose and ensures hybridization, and the stem loop technology that extends the miRNA using a universal primer are designed with a standardise the melting temperature of the miRNAs (T_m). PCR products could be analysed on a 2% agars gel (Raymond *et al.*, 2005; Kuhn *et al.*, 2008), or monitor the fluorescence by QRT-PCR (Ro *et al.*, 2006).

1.5 Selenium

The trace element selenium (Se) is an essential micronutrient that has been shown to have an important function in health and disease processes since the early 1930s, when an excess of Se in soils was indicated to be the cause of alkali disease in farm animals consuming high-Se plants. In 1957 Se was found to be not only a toxic nutrient, but also an essential one, protecting liver tissue in vitamin E deficient rats (Burk R. F., 1994).

1.5.1 *Selenium and health*

The impact of Se on human health was first identified in Keshan province of China, where Se deprivation caused an endemic cardiomyopathy that involves circulation disorders as well as myocarditis, and which was related to extremely low soil Se content and viral infection; incidence was considerably decreased once the crops were treated with sodium selenite (Liu *et al.*, 2002). Myxedematous cretinism (Zimmermann and Kohrle, 2002) (growth and mental retardation as well as facial maturation problems are found in this neurological disease common in Zaire) and Kashin-Beck (Moreno-Reyes *et al.*, 2003) (joint and bone disorders) are also diseases related to severe selenium deficiency (Gromer *et al.*, 2005); in the case of cretinism it is also related to iodine deficiency. More recently, Se deficiency has been associated with increased mortality risk in HIV (Fawzi *et al.*, 2005). In case of an insufficient Se intake, the distribution of Se among the tissues will migrate so as to protect brain, thyroid and reproductive system (Ferguson *et al.*, 2012), suggesting hierarchy in order to maintain the amount of Se vital for this organs. For any form of Se supplementation, beneficial effects can be observed after a month of supplementation, this is based on dietary studies which are usually from 4-6 weeks since this is enough to stabilise epithelial kinetics (Hu *et al.*, 2011).

Evidence for an influence of Se on the immune system has come from both observational and association studies. For example, an association study of serum Se and risk of mycobacteria pulmonary infection in an Indian population (Ramakrishnan *et al.*, 2012) found that low serum Se was associated with increased risk of pulmonary tuberculosis. In addition a selenomethionine supplementation trial of 12 weeks in a UK population has suggested improved immune status (Sanderson *et al.*, 2010); in this study the immune response to influenza vaccination was assessed by measuring IgG levels in plasma and cytotoxicity. Immune function as assessed by cytotoxicity was enhanced after supplementation with 50 µg Se/d; and improved immune responses to vaccines observed after supplementation with at 100 µg Se/.

Often, these studies investigate more than one micronutrient at a time. For example, a study of 25 asthmatic Taiwanese patients found low plasma Se possibly to compensate for antioxidant imbalance (Guo *et al.*, 2011). In another study elderly volunteers were fed with a 5 portions/day fruit and vegetables that increased plasma Se (among other micronutrients like retinol, carotenoids, ascorbic acid) for 12 weeks and after this intervention, one group was given was tetanus toxoid and Pneumovax while the another one was given a placebo. Natural killer cell cytotoxicity was higher in the immune challenged group, but no significant differences were found in response to the vaccine on the treated groups, except that patients eating the 5 portions/day diet reported a sense of well-being against illness (Sanderson *et al.*, 2010).

A further study where Se yeast supplementation (200 µg) reduced the admission to hospitals in American HIV patients has been reported (Hurwitz *et al.*, 2007). This effect was hypothesised as being due to Se increasing the production of activated T cells, increase tumour cytotoxicity and natural CD4 killer activity. Other studies in elderly people showed increased immune response (50% higher T lymphocyte count) against infections after a daily supplementation with 400 µg as Se yeast (Wood *et al.*, 2000). More recent publication linked the incidence of viral infection diseases with low Se content in the soil (Harthill, 2011).

Further evidence of enhanced immunity has come from the improved offspring survival and a reduction in mortality because of gastrointestinal infections among women with HIV (Siegfried *et al.*, 2012). In contrast, Se supplementation and drug consumption among HIV patients that has been reported to increases the risk of suffering from bacterial infection (Rayman, 2012).

The proposed roles of low Se intake in the aetiology of asthma in children has been proposed as being partly responsible for increased incidence of this disease in New Zealand, a country that historically has had inadequate Se intake (Thomson *et al.*, 2011). The mechanistic basis for this may the modulation of pro-inflammatory leukotriene metabolism, and thus reduced inflammatory response, by increased Se intake. Increased Se intake has also been reported to protect against recurrent wheeze from children exposed to smoke (Thomson *et al.*, 2012).

The roles of Se in humans include thyroid hormone metabolism (Schomburg and Kohrle, 2008), enhancement of immunity mediated by selenoproteins (Arthur *et al.*, 2003), male fertility (Flohe, 2007), antioxidant defence, brain function (Schweizer *et al.*, 2004), anti-proliferative effect in skin cancer (De Felice *et al.*, 2011), muscle metabolism (Brown and Arthur, 2001), and it is even considered a favourable factor of longevity (Lv *et al.*, 2011). Selenium intake recommendations are based on the amount of Se required to increase GP*x*3 activity in plasma to normal levels. Estimates of recommended Se intake based on GPx activity levels vary from 55µg/d (Burk *et al.*, 2006) to 60-75µg/d for women and men respectively (Sunde *et al.*, 2008). In contrast, a high intake of 200µg/d has been indicated to reduce risk of skin, prostate, lung, and colon cancer (Burk *et al.*, 2006) as well to improve immune system function in HIV (Fawzi *et al.*, 2005). In Australia, concentrations of Se in plasma of 95-100µg/l are considered as adequate (Huang *et al.*, 2011a)

Se is also thought to affect musculoskeletal heath. Supplementation at 5 µM of sodium selenite on BMMs cells and macrophage such as RAW 264.7 cells, had an antioxidant effect inhibiting

reactive oxygen species (ROS) generation which might reduce osteoclast differentiation so preventing bone reabsorption (Moon *et al.*, 2012). Low levels of Se are associated with the deforming musculoskeletal condition Kashin-Beck disease when plasma Se is from 5-27ng/ml (Downey *et al.*, 2009). Controversially a study of the importance of Se for bone development and that abnormal metabolism of postmenopausal Turkish women with different bone mineral densities, found little difference in plasma Se concentration, suggesting no relationship between osteoporosis and Se status (Arikan *et al.*, 2011).

In the case of thyroid function, a Se sufficient diet is required for optimal function of this gland, more specifically for the selenoproteins glutathione peroxidases and deiodinases which promote thyroxine (T4) and triiodothyronine (T3) synthesis by protecting iodine from oxidation; iodothyronine deiodinases are responsible for the catalysis of the thyroid hormone into active form; however supplementation in elderly people neither altered thyroid function on its own (Rayman *et al.*, 2008) when Se and iodine supplementation were given to the same participants (Thomson *et al.*, 2011) nor increased the conversion of T4 into T3 (Rayman *et al.*, 2008); however, Se supplementation is recommended as well as iodine to ensure an adequate thyroid function (Thomson *et al.*, 2011).

Se induces DNA damage response in MRC-5 normal lung fibroblasts (Wu *et al.*, 2011a) where ataxia-telangiectasia mutated (ATM) kinase is induced by Se so as to activate senescence response and limit tumourogenesis (Wu *et al.*, 2011a). In a study where human LNCaP were supplemented with sodium selenite and SeMet, a higher cell survival was observed after ROS exposure (de Rosa *et al.*, 2012).

1.5.2 *Food sources*

Dietary Se intake is dependent on the Se content of the soil on which food is grown (Harthill, 2011).. For example, Scottish soil is considered Se deficient being 0.34-0.48 mg/kg (Fordyce *et al.*, 2010)

when the recommended threshold by the Food Standards Agency is 0.6 mg/k; soil Se levels were was found to correlate with the Se concentration in potatoes, broccoli, beef, and milk (Fordyce *et al.*, 2010). The most common sources of Se for UK population are bread and cereals with a 0.01-30 mg/kg (26%) and meat where the Se content in muscle is 0.2 mg/kg and up to 0.6 mg/kg (26%); followed by a 0.01-0.03 mg/kg of Se from milk and dairy products (21%), fish (10%), and vegetables and fruit (7%). Eggs containing 3-25 µg/kg and up to 0.58 mg/kg on supplemented Se (4%) hens and finally other sources like Se enriched food and supplements (6%) (Fairweather-Tait *et al.*, 2011). Se content in fruits and vegetables is rather low in those that have not been supplemented, apart from garlic with a natural Se content of ~0.5 mg/kg in the forms of selenomethionine (12%) and seletane (4%); for example, potatoes contain 0.5 mg/kg in equal distribution of selenomethionine and selenite. An exceptionally high Se source are Brazilian nuts (*Bertholletia excelsa*) with 3800ng/g (Navarro-Alarcon and Cabrera-Vique, 2008). As a result of Se food content depending on soil Se levels food tables are notoriously inadequate for calculated Se intake.

In the UK, the main Se compounds contained in bread and cereals are selenomethionine (up to 85%) and as selenite or selenate (up to a 19%). Cattle, chickens and lambs are often supplemented with Se-enriched yeast and their meat is 60% of selenomethionine; while up to 30% in chicken and 50% in lamb of the total Se is in the form of selenocysteine. Cod fish and canned tuna are high Se fish (~1.8 and 5.6 mg/kg respectively) in the form of selenomethionine (29-70%) or selenite/selenate (12-45%). The most common Se-enriched fruits and vegetables are broccoli, brussel sprouts and onions (Fairweather-Tait *et al.*, 2011).

1.5.3 *Organic or inorganic Se compounds?*

Se supplementation can be implemented with organic (e.g. selenocysteine, selenomethionine, selenised-yeast) or inorganic sources (sodium selenite or selenate); all these sources are subsequently used in

the body for selenoprotein biosynthesis. Increased activity of the selenoprotein glutathione peroxidase can be achieved by inorganic Se, SeCys, or SeMet (Finley, 2006). Activity of the plasma selenoprotein glutathione peroxidase (GP*x*3) is often used as a measure to assess Se status because the activity of this enzyme is proportional to Se obtained from diet (Brown and Arthur, 2001).

The inorganic Se compounds sodium selenate and sodium selenite are both approved as supplements for animals and humans; in addition, sodium selenite is used for anticancer treatment due to its chemopreventive effect (Nafisi *et al.*, 2012). Both are efficiently absorbed in the small intestine but their different effects on liver suggest that they may be absorbed and metabolized differently; in rats supplemented with either sodium selenate or sodium selenite, liver gene microarray analysis showed that using a cut off of 2.5 fold change, 300 genes were significantly regulated by selenate whereas 43 genes were regulated by selenite (and selenoproteins GPX1, SELP, 15kDa SEP, SELW, DIO1, or GPX4 were unaffected)(Bosse *et al.*, 2009). However, the effects of organic or inorganic Se compounds differ according to the organ, species, length of exposure and concentration used. Example of this is the expression of GPX3 in Cos7 (monkey kidney cells) which were subjected to inorganic and organic selenium supplementation; up to a 145nM concentration sodium selenite increased GPX3 expression by ~50% compared to the organic compounds whereas in comparison supplementation with selenomethionine or Se-methylselenocysteine hydrochloride required 580nM to increase the expression of GPX3 to a similar extent (Ottaviano *et al.*, 2009). Human prostate cancer cells, LNCaP cells, treated with sodium selenite had a higher cell survival after the cells had been exposed to oxidative agents than the cells treated with SeMet (de Rosa *et al.*, 2012). Se in milk has been found in human intervention studies to have greater impact on plasma selenoproteins than supplementation with selenised-yeast (selenomethionine and Se-methylselenocysteine and γ-glutamyl-Se-methylselenocysteine). This could be due to milk proteins helping with the absorption or colorectal

retention of Se compounds. Se compounds in milk are of a molecular size greater than 10kDa because of their preparation (Hu *et al.*, 2011).

Se from diet comes in both organic and inorganic forms and often, human and cell models attempt to investigate the effects of these selenium compounds; however, as described in Figure 4, Se compounds are incorporated into selenoproteins. I*n vivo*, both SeMet and selenite have been shown to increase the levels of SEPP when Se intake is limited (Xia *et al.*, 2005), and *in vitro* viability of HepG2 cells is similar at 24 and 72 h after supplementation with 150 µg/L of either SeMet or selenite (Hoefig *et al.*, 2011). A human study in the US involved supplementation of both male and female volunteers with150 µg of stable isotopic tracers of selenite and selenomethionine for 10 days and then followed over 4 months where selenite estimated absorption was about 57% whereas SeMet was 97% measured by triple isotope dilution; in terms of excretion high levels of SelMet were found in urine while selenite concentration was higher in faeces; overall, Se urinary excretion was 70 µg/d (Wastney *et al.*, 2011). A study in which Danish participants were given 74.4 µg of sodium selenite had a urinary excretion of ~32 µg/d after 24 h (Gammelgaard *et al.*, 2012b). Nonetheless, the different levels of absorption in these studies could be due to a competition of Se incorporation as all the Se compounds were administrated at a rather high concentration at the same time.

1.5.4 *Selenoproteins*

Selenoproteins are a group of proteins present across species (except yeast) that contain selenocysteine (Sec) as the amino acid analogue to cysteine with Se replacing sulphur (Mariotti and Guigo, 2010) (Bellinger *et al.*, 2009). 25 selenoproteins have been identified and classified, including glutathione peroxidises, iodothyronine deiodinases, thioredoxin reductases and selenoprotein P (SELP).

These selenoproteins respond to lower Se supply in a hierarchal manner that is presumed to reflect functionality in cell survival and cell

cycle regulation (Mariotti *et al.*, 2012). It is also commonly paralleled by the effect of Se status on selenoprotein mRNA expression (Driscoll and Copeland, 2003). Figure 3 shows the minimum Se requirement for rats supplemented with sodium selenite in order to obtain an mRNA breakpoint expression of 12 selenoproteins (Sunde and Raines, 2011). As shown in Figure 3, the relatively low Se intake is needed to maximise GPx4 mRNA levels in liver is Gpx4, followed by Selk and Selt; in contrastGpx1 and Selw require a higher concentration of Se. This was also supported by data from a feeding experiment in mice which showed that in the colon *SelW*, *GPx1*, (in addition to *SelH*, and *SelM*) mRNAs levels are most sensitive to low Se intake (Kipp *et al.*, 2009). Such findings are consistent with the use of GPX1 (Pagmantidis *et al.*, 2005; Combs *et al.*, 2012) and SELW (Pagmantidis *et al.*, 2005) as biomarkers of Se status in human models. It has been suggested that the selenoprotein hierarchy could be explained through a high selenophosphate synthetase (SEPHS2) (Squires *et al.*, 2007) and/or methylation of Sec tRNA^{Sec} that facilitates the biosynthesis of GPx1, GPx3, SelR, and SelT (Reszka *et al.*, 2012).

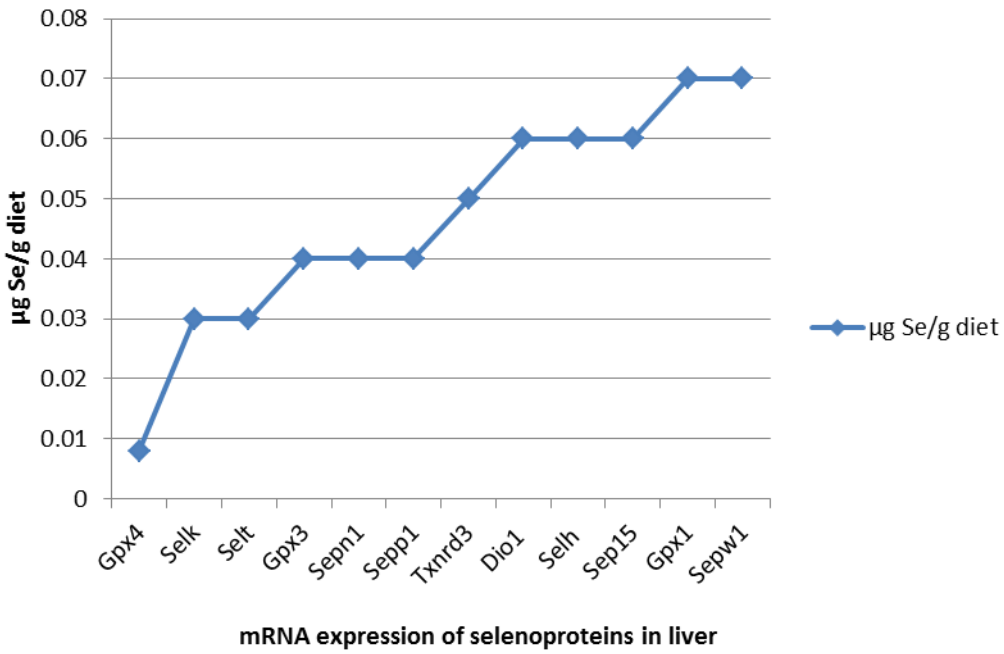


Figure 3 Minimum dietary Se necessary to reach plateau levels in liver from rodents.
This figure adapted from (Sunde and Raines, 2011).represents the amount of sodium selenite required for rats to reach plateau levels of different selenoprotein mRNAs. Values are representative of the hierarchy in the rat liver, where the selenoproteins situated on the left occupy a higher hierarchy.

The characterization of selenoproteins has not been completed (de Rosa *et al.*, 2012). However, the major functions of several families have been identified. The role of the iodothyronine deionase family is the regulation of T3 and T4 metabolism. It has been found that thioredoxin reductases have a redox function and in addition have roles in transcription factor regulation (Gromer *et al.*, 2005).

Glutathione peroxidises (GP*x*s), the most studied group, confer protection against oxidative stress at a cellular level (Méplan C., 2006; Lu and Holmgren, 2009), and have a structural role in spermatozoa (GPx4); reduced glutathione peroxidase levels lead to an increased viral virulence (Moghadaszadeh and Beggs, 2006). From which GPX1, GPX2, GPX3, and GPX4 will be addressed in this thesis.

One of the most abundant GP*x*s (Reeves and Hoffmann, 2009) highly expressed in neurons, astrocytes, nervous system, brain (Gardaneh *et al.*, 2011)and high sensitive to Se supply is glutathione peroxidase 1(GPX1) (Pagmantidis *et al.*, 2005; Reeves and Hoffmann, 2009), a homotetramer (The UniProt, 2012) of cytosolic localization and the most abundant intracellular GPX (Felice *et al.*, 2010) acting as oxidative stress protector where GPx is oxidised to reduce H_2O_2 to H_2O, maintaining the equilibrium between free radicals and superoxide dismutase enzymes (Gardaneh *et al.*, 2011). GPX1 deficiency is associated with NFκB and JNK activation, atherogenesis, and proinflamatory response (Lubos *et al.*, 2011). Similarly, increased levels of lipopolysaccharide-induced ROS on human microvascular and coronary endothelial cells, was achieved by GPX1 knockdown (Lubos *et al.*, 2010). The overexpression of GPX1with a pLV-GPX construct protected astrocytes from the toxicological effects of the parkinsonian neurotoxin 6-hydroxydopamine, and improving cell survival (Gardaneh *et al.*, 2011). ROS were induced after 24 h of 25nM of D-glucose exposure which decreased at 48 h of treatment possibly due to the endogenous increase of GPX1 activity and expression (Felice *et al.*, 2010).

Gastrointestinal glutathione peroxidase or glutathione peroxidase 2 (GPX2) is an homotetramer with cytosolic localization (The UniProt, 2012), mostly expressed in the gastrointestinal tract to confer antioxidant stress protection (Reeves and Hoffmann, 2009) and it has been suggested to regulate intestinal renewal as a Wnt signalling pathway target (Kipp *et al.*, 2012).

GPx2- Knockout (KO) mice fed with 0.086, 0.15, or 0.64 mg/kg as selenomethionine developed colitis (particularly affected at 0.086 mg Se/kg) and showed faster tumour growth with no big differences with among Se diets (Krehl *et al.*, 2012). Further work on these mice fed with the same diet described above, showed increased apoptosis in colonic crypts (Florian *et al.*, 2010) suggesting antiproliferative and antiinflammatory roles.

Glutathione peroxidase 3 (GPX3) is an homotetramer (The UniProt, 2012) extracellular enzyme produced by the kidney and thyroid, and secreted into plasma and body fluids in general (Brigelius-Flohe and Kipp, 2009). *GPX3* is usually downregulated on cancer patients (Brigelius-Flohe and Kipp, 2009) and expressed in head and neck, lung, ovary, bladder, oesophagus, melanoma human cancer cell lines, and healthy lymphocytes (Chen *et al.*, 2011a). GPX3-KO mice show increasing H_2O_2 in plasma which has a protombotic effect, leading to vascular injury (Jin *et al.*, 2011) since *GPX3* regulates the bioavailability of nitric oxide from platelets and vascular cells (Reeves and Hoffmann, 2009). GPX3 hypermethylation was associated with chemotherapy response (Chen *et al.*, 2011a). However, methylated *GPX3* promoter is more frequent in oesophageal squamous cell carcinoma and this phenomenon downregulates *GPX3* which is a proposed carcinogen effect (He *et al.*, 2011).

Glutathione peroxidase 4 (GPX4) is a monomer (The UniProt, 2012) with three cellular isoforms, a cytoplasmic, a mitochondrial (The UniProt, 2012) and one nuclear (Speckmann *et al.*, 2011). GPX4 is involved in regulation of arachidonic acid metabolism (Reeves and

Hoffmann, 2009). Cell death by lipid peroxidation is a known result of GPX4 silencing (Ueta *et al.*, 2012); and GPX4 has protective effects against Alzheimer disease symptoms in mouse brain (Speckmann *et al.*, 2011). More studies suggest that GPX4 malfunction could lead to functional changes in photoreceptors caused by lipid peroxidation and compromised cell survival due to inefficient energy supply to the primary cilia (Ueta *et al.*, 2012).

SEPHS2 is essential for selenoprotein biosynthesis (Azevedo *et al.*, 2010) (see page 39). Mutations in exon 3 of *SEPHS2* caused an R10X nonsense mutation in a 12 year old girl and this lead to high T4 and low T3, reduced levels of Se and GPX3, DIO2 mutations, no expression of SELP, developmental coordination disorders, congenital myopathy, hypotrophy, respiratory depression and delayed mental development. After 30 days of 200 µg/day Se yeast supplementation, GPX3 and T4 increased but SELP and T3 levels did not change (Azevedo *et al.*, 2010).

Recently, several selenoproteins have been found to be present in the endoplasmic reticulum (ER). Selenoprotein S (SELS) has been suggested to have a role in brain inflammation by reducing the synthesis of proinflammatory cytokines (Fradejas *et al.*, 2011) and regulating interleukin 6 release by astrocytes under ER stress (Maraldi *et al.*, 2011). Because of its ER localization, it has also been proposed to be involved in protein degradation *via* the unfolded protein response (Maraldi *et al.*, 2011). In terms of metabolic roles, SELS has also been linked to glucose metabolism after studies in participants in western Sweden with coronary disease and another group of obese patients showed an increased in *SELS* expression on adipocytes during hyperinsulinemia, leading to the suggestion that *SELS* is a biomarker of chronic diseases (Olsson *et al.*, 2011). Another endoplasmic reticulum selenoprotein with a wide highly expressed in the brain (Hwang *et al.*, 2005; Reeves *et al.*, 2010) is Selenoprotein M (SELM). It is thought to reduce toxicity from oxidative stress and it also modulates baseline calcium levels (Reeves *et al.*, 2010). 15kDa selenoprotein 15 (SEP15), another ER selenoprotein is also thought to play a role in the protein folded response (Kasaikina *et al.*,

2011a). It is expressed in secretory organs (Kasaikina *et al.*, 2011a), and has been linked to a reversed cancer phenotype in the intestinal cell lines HCT116 and HT29 (Kasaikina *et al.*, 2011a); a recent report suggests that there is increased chance of survival on cancer patients when the expression of Sep15 in plasma is high (Slattery *et al.*, 2012).

Selenoprotein H (SELH) has been suggested to be responsible for the regulation of mitochondrial respiration by upregulating the genes involved in mitochondrial biogenesis and oxidation - NRF-1, PGC-1α, and Tfam-, and decreasing NF-KB, which inhibits the transcription of PGC-1α in the nuclear fraction (Mendelev *et al.*, 2011). By inhibiting mitochondrial membrane depolarization, SELH has been proposed to protect brain cells against mitochondrial induced apoptosis (Mendelev *et al.*, 2011). In the hippocampal neuronal cell line HT22, SELH overexpression enhanced neuronal cell survival after UVB exposure by stabilizing mitochondrial membrane potential, inhibiting p53 and caspades 3 and 9 activation (Mendelev *et al.*, 2009).

Finally, SELP (Selenoprotein P), which includes 10 Sec, accounts for 50% of the Se in plasma and is responsible for extra hepatic Se balance (Burk and Hill, 2009) as an hepatic-tissue Se transporter (Li *et al.*, 2007). It appears important in Se transport to the brain (Moghadaszadeh and Beggs, 2006) and testis where Se levels are maintained by means of a receptor mediated uptake of SELP (Burk and Hill, 2009).

1.5.5 *Selenium uptake*

Determining Se uptake *in vivo* is challenging because it is difficult to control every condition in the intestine. For this reason, cellular models are preferred to explain the absorption of this essential micronutrient (Combs *et al.*, 2012).

Zeng et al compared the transport of 500 nmol/L selenite, SeMet, methylselenocysteine from the apical to the basolateral side of Caco-2 monolayers and found that SeMet and methylselenocysteine transport

were more efficient than selenite but that transport did not correlate with GPx activity, being selenite > SeMSC > L-SeMet (Zeng *et al.*, 2011).

Subsequently, an X-ray absorption spectroscopy approach showed that in A549 human lung adenocarcinoma epithelial cells treatment with organic Se compounds SelMet was incorporated into selenoproteins after 24h while incorporation was slower with Se-methylselenocysteine (Weekley *et al.*, 2011).

A recent study using,Caco-2 cells, a cell model of human intestinal absorption, investigated the uptake of different forms of Se in more detail, They found that the best absorptive flux was with SeMet followed by semethylselenocysteine and that this occurred through a transcellular transport mechanism. In addition they suggested that selenate was transported paracellularly and passively across the intestinal monolayer, and that selenite reacts with thiol groups of glutathione and cysteine reaching the intestinal layer as selenotrisulfides. (Gammelgaard *et al.*, 2012a).

1.5.6 *Selenoprotein biosynthesis*

This essential trace element in proteins is in the form of selenocysteine (Sec) metabolized from either organic or inorganic Se compounds through different pathways (Papp *et al.*, 2007; Zeng *et al.*, 2011). Although the organic form of Se, selenomethionine (SeMet) is more bioavailable, it does not correspond to the best source of selenoprotein biosynthesis (Zeng *et al.*, 2011).

Incorporation of Se as the aminoacid Sec into selenoproteins at the UGA codon was first discovered when the TGA codon was found in the Open Reading Frame (ORF) of the *GPX1* gene (Moghadaszadeh and Beggs, 2006; Papp *et al.*, 2007). Sec is coded by UGA (usually a stop codon) and incorporated during translation. This requires a special structure within the mRNA, the SECIS (selenocysteine insertion sequence), to recode the UGA. In eukaryotes the SECIS is located in the

3’UTR of the mRNA (Méplan C., 2006; Papp *et al.*, 2007; Hesketh, 2008; Mariotti and Guigo, 2010). A specific tRNA, Selenocysteyl-tRNA-Sec (Sec-tRNA^{Sec}) (Hubert *et al.*, 1996; Fletcher *et al.*, 2001), is necessary for Sec incorporation and this is derived from serine by the reactions of a SEPHS2 and selenocysteine synthase (Papp *et al.*, 2007).

Figure 4 illustrates how selenoprotein biosynthesis is proposed to occur. It starts with the incorporation of serine to tRNA^{Se} into Seryl-tRNA^{Sec} (Ser-tRNA^{Sec}) by the enzyme SeryltRNA synthetase (I) (Hubert *et al.*, 1996). Either organic (II A) or inorganic sources (II B) are reduced to selenide which is then phosphorylated by adenosine triphosphate (ATP) (Hoffmann and Berry, 2005) by SEPHS2 (III) to form selenophosphate. Selenophosphate and Ser-tRNA^{Sec} are then incorporated into selenocysteyl-tRNA^{Sec} (Sec-tRNA^{Sec}) by the action of selenocysteine synthase (IV). The SECIS-binding protein 2 (SBP2) then binds the Sec-tRNA^{Sec} (Hubert *et al.*, 1996) (V) to form a complex formed with selenocysteine-tRNA specific elongation factor (EF-Sec)(VI) (Hubert *et al.*, 1996). This complex will finally bind the SECIS and lead to Sec UGA codon recognition (VII).

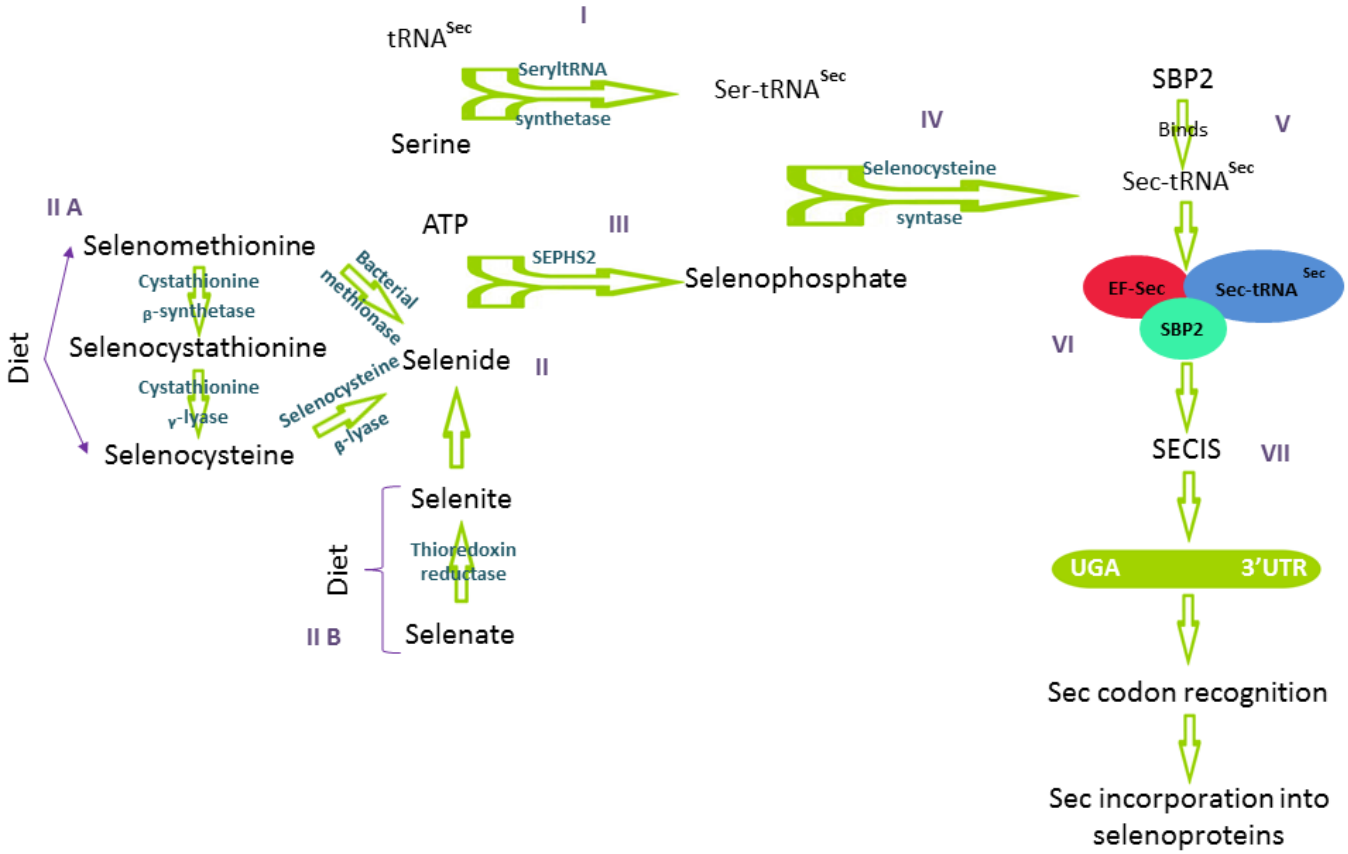


Figure 4 Selenoprotein biosynthesis.
(I)The incorporation of Se into selenoproteins requires a specific tRNA that will interact with a highly reactive reduced Se compound (selenophosphate) originated from the metabolised organic and inorganic Se compounds. Se from diet is converted into selenocysteine and metabolised into selenide via enzymatic action or through gut bacterial action (IIA); selenide can be obtained or from (IIB) oxidation of the inorganic form selenate into selenite and further selenide. (III). The metabolised selenide is then catalysed by ATP and SEPHS2 into selenophosphate, which is incorporated into Sec-tRNA^{Sec} (IV). This forms a complex with selenocysteine-specific factor (EF-Sec), SBP2 (VI), and the SECIS (VII) to allow translation of selenoproteins at the UGA codon (VIII).

1.5.7 *Downstream targets of Se*

Despite the fact that the main target of Se studies is the effects of Se supplementation/depletion on selenoproteins, a group of downstream genes has also been identified from microarray experiments when assessing the effects of Se. Monitoring changes in expression of such targets could help understanding the physiological effects of Se intake.

In a microarray study of normal and cancer prostate cells grown with an increasing concentration of Se-methylselenocysteine 16 genes - other than selenoproteins- were identified as showing altered expression (Hurst et al,2008); these included genes with functions in collagen transport, cancer and cell proliferation functions (Hurst *et al.*, 2008).

Microarray analysis of RNA from lymphocytes from healthy volunteers before and after Se supplementation with sodium selenite showed changes in expression of genes within the protein biosynthetic pathways, including 61 genes from ribosomal and mitochondrial proteins, as well as translation elongation and initiation factors (Pagmantidis *et al.*, 2008). A similar pathway was identified in the mouse colon in which the mice where fed either a Se deficient or Se-sufficient diet and colon analysed by gene microarrays (Kipp *et al.*, 2009); a large number of genes were either up (1979) or down (2016) regulated under marginal Se deficient compared with Se-sufficient conditions.

Other potential targets of selenium like apoptotic regulators (BAX, CASP3 and CASP9), tumour suppressors (PTEN) were identified by Se supplementation of the human osteosarcoma cell line U2OS (Chen *et al.*, 2012). BAX, CASP 3 were upregulated at mRNA and protein level. Sodium selenite supplementation of the human prostate epithelial cancer cell line PC3 was found to inhibit the protein expression of the prostate cancer progression regulators $TGF\beta_1$ and VEGF, as assessed by ELISA (Pei *et al.*, 2010).

1.5.8 *Selenium and colon*

Previous studies have associated increased Se intake with lower colorectal cancer risk and improved colorectal cancer survival (Peters and Takata, 2008; Meplan and Hesketh, 2012). However, the mechanistic basis for such proposed links are not known. It has been hypothesised that these effects of Se are due to oxidative stress protection mediated by Se-sensitive genes and selenoproteins (Slattery *et al.*, 2012).

SNP in the gene encoding selenoproteins thioredoxin reductase (TRX), SELX, and SELN on a case-control study of colon cancer patients and controls were related to environmental factors like smoking, aspirin and a minimum effect from Se intake, while *TRX1*, *TRX2*, and *TRX3* SNPs were associated with smoking and colorectal cancer (Slattery *et al.*, 2012). A similar correlation between SNPs and cancer was found on Korean patients with colon and rectal cancer where a SNP on one variant T allele of rs34713741 on SELS was associated with a significant risk increase in female patients; nevertheless on the same study, SNP on rs5845 or rs5859 in SELP15, the risk of rectal cancer was higher in men (Sutherland *et al.*, 2010).

Selenium, as sodium selenite, has been linked to tumour reduction and therefore proposed for chemopreventive therapy (Thirunavukkarasu *et al.*, 2008). Colon adenocarcinoma cells is a cell model commonly used in Se experimental models as scientific evidence is been found that Se efficiency is linked to gastrointestinal diseases (Peters and Takata, 2008; Connelly-Frost *et al.*, 2009; Wu *et al.*, 2010). Once Caco-2 reach confluency, their phenotype behaves as the morphology and functionality of an intestine (Piana *et al.*, 2008) and the absorption and transport of nutrients such as selenium, iron, zinc, calcium, and copper is comparable to physiological conditions due to its structure and function is similar to an enterocyte (Zeng *et al.*, 2011).

More recently, a 6 week supplementation with either 150µg Se/d of diary (83% selenomethionine 5% selenocysteine) or yeast (83% selenomethionine 5% selenocysteine and 3% selenite) Se of 23 volunteers (mean age of 64 years) with an initial plasma concentration of 103µg/l, investigated the effect of Se supplementation on either organic or inorganic forms to study the benefits of Se on colon cancer risk selenoprotein related. Plasma concentrations of Se increased, finding a higher level under Se-yeast supplementation; although women had higher plasma Se concentrations, GPx activity in plasma was higher in men. Rectal selenoprotein expression had no gender related variations. Colorectal biopsies revealed that Se diary supplementation increased GPx 1 and 2 mRNA levels. Dairy-Se supplementation was enough to maintain SELP levels at the 6-week washout period (Hu *et al.*, 2011; Huang *et al.*, 2011a).

The beneficial effects of Se have also been studied in the rat 1,2-dimethylhydrazine (DMH) induced tumour model. Treatment with Se as sodium selenite resulted in a 60% reduction of the tumour growth after 20 week of supplementation. Although some nuclear polarity was lost, reduced lipid peroxidation, reduced levels of DMH, increased levels of glutathione-S-transferase, superoxide dismutase, and glutathione were observed after Se treatment (Ghadi *et al.*, 2009). In a study in Caco-2 cells in which cells Se-methylselenocysteine sufficient and deficient media was compared, the von Hippel-Lindau tumour suppressor gene was downregulated under Se-methylselenocysteine deficient conditions. The study hypothesises that Se could provide methyl groups altering methylation pathways or by methyl group modification since liver and colon from rats has shown a proportional response in the expression of Se and DNA methyltransferase (Uthus *et al.*, 2010).

Selenomethionine, methylselenocysteine (SeMSC) and Se enriched yeast are 3-fold more effectively transported than selenite (Zeng *et al.*, 2011). However, selenite activates the AKT/β-catenin signalling inducing apoptosis of colorectal cancer cells (Luo *et al.*, 2012). Selenoproteins like SELS have been linked to colorectal cancer;

selenoproteins by reducing the impact of ROS by the glutathione peroxidases and thioredoxin reductases would protect against cancer (Meplan and Hesketh, 2012).

1.6 Aims and objectives

Se role in human health has been widely studied in different cell models, tissues and biopsies of several organisms, and has pointed out the beneficial effects of Se intake in human health with selenoproteins as key regulators of metabolic pathways. Apart from these Se-containing proteins, Se intake also influences expression of other group of genes/proteins by indirect, and largely unknown, mechanisms. The purpose of the present study was to investigate possible relationships between Se metabolism and recently discovered small RNA species (miRNA) with the ability to regulate expression of a protein or group of proteins by interacting with its/their mRNA.

To achieve this aim, the following objectives were set:

- To use a bioinformatics approach to identify miRNAs with the potential to interact with selenium sensitive genes/pathways.

- To use a custom miRNA microarray to find miRNAs expressed in the Caco-2 cell line and detect how their expression changes after selenium supplementation/depletion.

- To use transcriptomics to identify transcripts that show changes in expression in Caco-2 cells grown under Se supplemented/depleted conditions.

- To analyse the datasets from the miRNA microarray and the mRNA microarray with bioinformatic tools in order to find possible miRNA-mRNA interactions that respond to Se supply.

- To test the effect of altered miRNA expression on these potential Se-sensitive target genes.

A summary diagram of the aims of my study is shown in Figure 5.

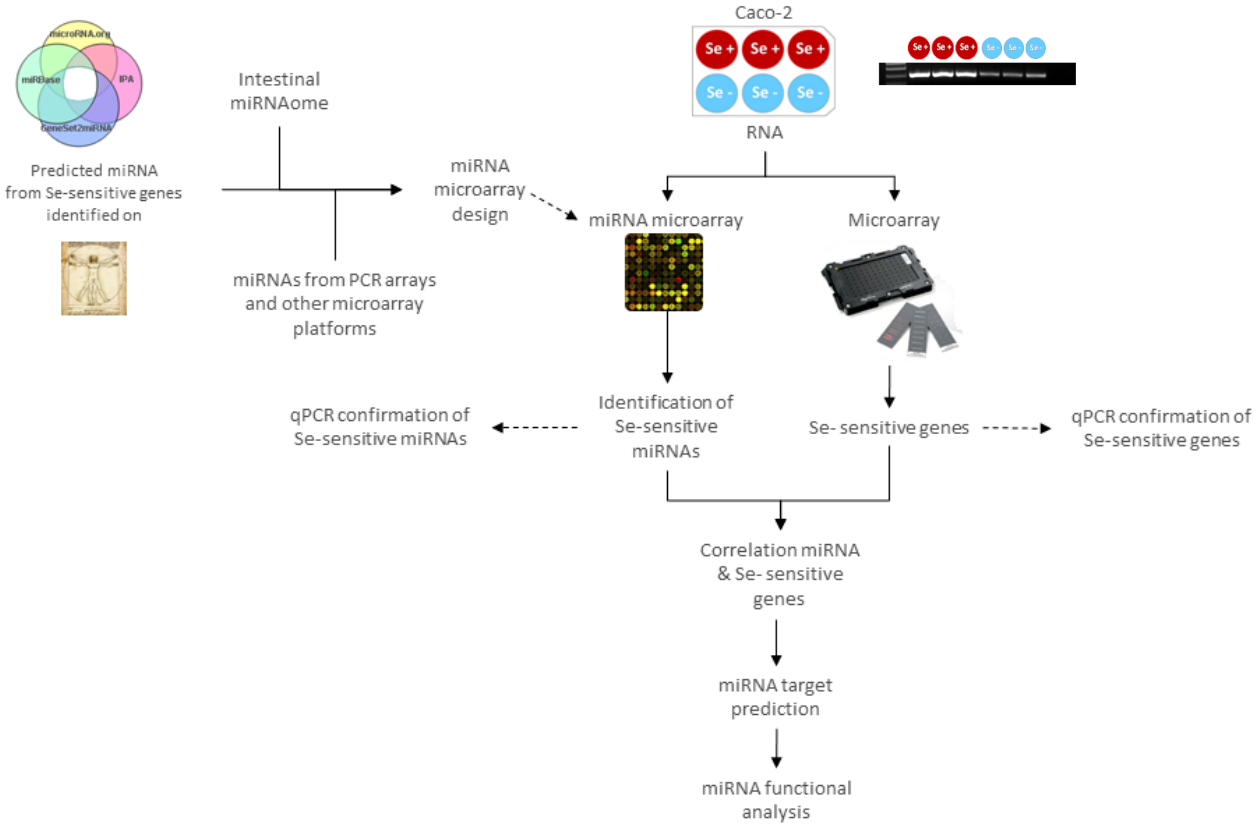


Figure 5 Aims and objectives.
In order to investigate the effect of Se on miRNAs and their possible targets in Caco-2 cells, RNA from Caco-2 cells grown in cells supplemented with 40.5nM of sodium selenite for 72 h was analysed in a total transcriptome experiment and a customised miRNA microarray chip. The chip design was based on bioinformatics predictions of miRNAs using genes within a Se-sensitive pathway identified in a human supplementation trial, miRNAs expressed in gastrointestinal tissues and cell lines, miRNAs identified on PCR arrays. Microarray results were confirmed by Real Time PCR for both miRNAs and mRNAs. The Se-sensitive miRNAs and transcripts were analysed bioinformatically in order to identify potential interactions n and provide a basis for a miRNA functional analysis.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 *Equipment*

Jouan High Speed cooling Centrifuge MR 23**,** Laborgerätebörse GmbH
NanoDrop® ND-1000 Spectrophotometer, Labtech International
Jouan B4 Centrifuge
Agilent 2100 bioanalyzer
Nikon TMS microscope
LightCycler® 480 Real-Time PCR System
Applied Biosystems 7900HT Fast Real-Time PCR

2.1.2 *Chemicals*

Trypan blue Sigma-Aldrich, Germany
DMEM - Dulbecco's Modified Eagle Medium
Fetal-calf serum
Penicillin Streptomycin
Insulin Bovine with 10mg/mL insulin in 25 mM HEPES, pH8.2 by Sigma-Aldrich, Germany
Transferrin
DMEM - Dulbecco's Modified Eagle Medium, Invitrogen®
MEM Non-essential amino acids 100x by GIBCO®
Sodium selenite Sigma-Aldrich, Germany
Invitrogen TRIzol® Reagent
TRIzol® Plus RNA Purification Kit
Chloroform, minimum 99%, Sigma-Aldrich®

Agilent RNA 6000 Nano Kit
Reverse transcriptase (Roche)
dNTP by Bioline, UK
Hyperladder IV, Bioline, UK
Hyperladder V, Bioline, UK
Ultrapure agarose, Invitrogen™
TAE buffer Tris-Acetate-EDTA, Sigma
TEB Tris-Borate-EDTA, Fisher
Diethyl Pyrocarbonate (DEPC) 97% NMR, Sigma ®.
Ethanol, Analytical reagent grade. Fisher scientific, UK
Oligo $(dT)_{15}$ primer for DNA synthesis, Roche, Germany
Primers Eurofin MWG Operon oligonucleotides, UK
QIAquick PCR Purification Kit (QIAGEN, UK).
Transcriptior RT reaction Buffer, 5x, Roche, Germany
Protector RNase Inhibitor
LightCycler 480 SYBR Green I Master**,** Roche, Germany
Opti-MEM® (Invitrogen, UK)
Lipofectamine® 2000 Transfection Reagent (Invitrogen, UK)
Anti-MiR-185 Applied Biosystems, UK
Anti-MiR negative control Applied Biosystems, UK

2.2 Methods

2.2.1 *Cell culture*

Human colon adenocarcinoma cells (Caco-2) were used as a model of Se effects on the colonic epithelium. Cells were cultured in 75cm tissue culture flasks and fed with 15mL of DMEM - Dulbecco's Modified Eagle Medium (with 4.5g/L glucose and Glutamax) with 10% Fetal-calf serum (FCS) and 1% Penicillin/streptomycin (100 units/mL and 100 µg/mL, respectively) and grown at 37°C in the presence of 5% CO_2. The cells were harvested at 80-90% of confluence.

For investigation of the effects of Se supply, $5x10^5$ - $10x10^5$ /mL cells were seeded into each well of a 6-well plate. Cells were cultured in either Se-sufficient or Se-deficient medium comprising 2mL of DMEM (with 4.5g/L glucose and Glutamax) without added FCS but containing 5µg/mL insulin, 100 µg/mL Penicillin/streptomycin, 5µg/mL transferrin, and in for Se-sufficient medium 7ng/mL sodium selenite (40.5nM). Cells were fed every 2 days and harvested after 72 h in Se sufficient/ depleted media.

To count the number of cells, cells were rinsed with 2mL of Phosphate buffered saline (PBS), and incubated with 450 µL of Trypsin at 37°C for 10 minutes to detach them. After the incubation period, a cell pellet was obtained by centrifugation at 1500 RPM from 5 min, the pellet was resuspended gently (to avoid lysis) in 10mL of DMEM, diluted 1:2 dilution in trypan blue and then counted in a haemocytometer a Nikon E 10 0.25 160 / - Ph 1 DL objective lens.

2.2.2 *Sodium selenite 7ng/mL; 40.5nM*

Sodium selenite was resuspended to a final concentration of 10µg/mL in PBS and filtered with a 0.22µm filter, Millipore express PES membrane.

2.2.3 *RNA isolation for total transcriptome analysis*

Total transcriptome isolation was performed to ensure conservation of the miRNA population in the final RNA preparation.

Medium from Caco-2 cells was removed and cells were rinsed with a volume of PBS equal to that the medium that the flask/well contained. 1mL of TRIZOL® Reagent was added per $10cm^2$ dish area. Cells were detached from the bottom with a cell scraper and cell lysate

was homogenized by pipetting up and down with an RNase-free filtered tip. Cell lysate was collected into a microcentrifuge tube and incubated at room temperature for 5 min.

For the separation of aqueous (containing RNA) and organic phases, 200µL of chloroform was added to the lysate and the mixture was shaken vigorously by hand for 15 sec. After a 3 minute incubation at room temperature, the tubes were centrifuged at 12,000 x g for 15 min at 4°C to separate the RNA (the upper phase) which was transferred into a fresh microcentrifuge tube without disturbing the interphase. Then, in accordance with best practice gene expression profiling experiments, a second separation phase was performed before ethanol precipitation (Naderi *et al.*, 2004).

Then the upper layer was thoroughly mixed with an equal volume of ethanol and RNA isolated using a Spin cartridge column. A maximum of 700 µL added to a Spin Cartridge and centrifuged at ≥12,000 x g for 1 min at room temperature. The flow through was discarded and the step was repeated with the remaining sample solution.

500 µL of washing buffer II with ethanol were added to the centre of the column and centrifuged at ≥12,000 x g for 15 sec at room temperature. Flow through was discarded and the washing step was repeated a second time.

The membrane of the column was dried by centrifuging at ≥12,000 x g for 1 min. To obtain 60 µL of purified RNA, 62 µL of 0.05% DEPC-treated water was added to the centre of the Spin Cartridge and centrifuged in a Recovery tube ≥12,000 x g for 2 min at room temperature. The RNA was aliquoted 2 or 3 aliquots of RNA were prepared and stored at -80°C.

2.2.4 *Measurement of RNA Concentration*

RNA concentration and purity were determined spectroscopically using the NanoDrop® ND-1000 Spectrophotometer. The RNA concentration was based on 260 nm absorbance. 260/280 ratios between 1.8-2 indicated low contamination with protein, phenol or other contaminants. 260/230 ratios from 1.8-2.2 indicated the nucleic acid purity. Such values for both ratios where routinely achieved.

2.2.5 *RNA integrity assessment*

From each sample, 1µL of 5-500ng/µL total RNA was loaded into a RNA Nano chip and analysed by gel filled microcapillary electrophoresis using detection through laser-induced fluorescence on an Agilent 2100 bioanalyzer.

65µL of filtered gel at room temperature was centrifuged after vortexing for 10 sec and mixed with 1µL dye. The mixture (Gel-Dye Mix) was vortexed and centrifuged in a microcentrifuge at 13,000g for 10 min at room temperature.

9µL of the Gel-Dye Mix was loaded into a well marked **G** from the RNA 6000 Nano chip, and compressed for 30sec in the chip priming station, 9µL was loaded the wells marked with ; the remaining gel was discarded. 5µL of RNA 6000 Nano marker was then loaded into each well. Subsequently, 1µL of ladder was loaded into its corresponding well, and 1µl of RNA previously denatured for 2-3 min at 70°C RNA was loaded into the remaining wells.

The RNA 6000 Nano chip was vortexed for 1 min at 2400 RMP and was the sample analysed within 5 min to avoid evaporation of the reagents.

After approximately 30min, the RNA concentration and RNA Integrity Number (RIN) were shown for each sample. The RIN is quantified from 1 to 10 and indicates the extent of RNA degradation (Schroeder *et al.*, 2006).

2.2.6 *miRNA target prediction for the protein biosynthetic pathway*

An earlier study in of human lymphocytes had suggested that the protein biosynthetic pathway was sensitive to Se supply. The 61 Se sensitive genes within the protein biosynthetic pathway (Pagmantidis *et al.*, 2008) were loaded in 3 bioinformatic tools and 1 software for miRNA prediction in order to obtain possible miRNA-mRNA interactions that would then lead us to build an interactive network.

Microrna.org was chosen because its algorithm has been used in target prediction obtained high scores and most of them experimentally validated (John *et al.*, 2004) and it offers the flexibility of not relying on the seed region completely and because scientific evidence has shown that miRNAs interactions occur beyond this region (Brodersen and Voinnet, 2009), (Grimson *et al.*, 2007) (Lytle *et al.*, 2007), (Mazière and Enright, 2007), (Stark *et al.*, 2007).

At microrna.org (Betel *et al.*, 2008), each gene symbol was loaded into the Target mRNA search tool for *Homo sapiens* (*H. sapiens*) in a window, as shown in Figure 6.

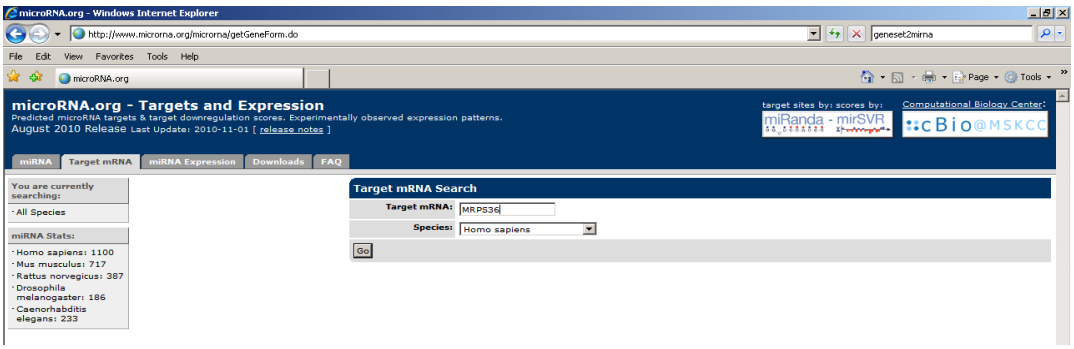


Figure 6 Example of data entry for microRNA.org.
A gene name from the protein biosynthetic pathway is typed into the “Target miRNA” box, and by selecting *H. sapiens* under “Species”, the given mRNA ID will be scanned by microRNA.org in order to localize MRE on the gene’s sequence.

A graphic list of predicted alignment was then visualized. This window contains the miRNA sequence and its mRNA alignment site, the number of miRNA targets, alignment score, PhastCons score which measures the changes in orthology across time and space, and a list of miRNA organized by the number of target sites and alignment scores hierarchy, where more target sites, a high alignment score, and high PhastCons score (higher than 0.57) are the optimal conditions for a mRNA-miRNA interaction to occur.

After microRNA.org had scanned the sequences and found miRNA-mRNA cleavage sites, the alignments and the scores alignments (conservation across species, miRNA energy required, and overall score) are displayed along the gene sequence. An example of a protein sensitive to Se supply used in this study and its MRE is showed in Figure 7.

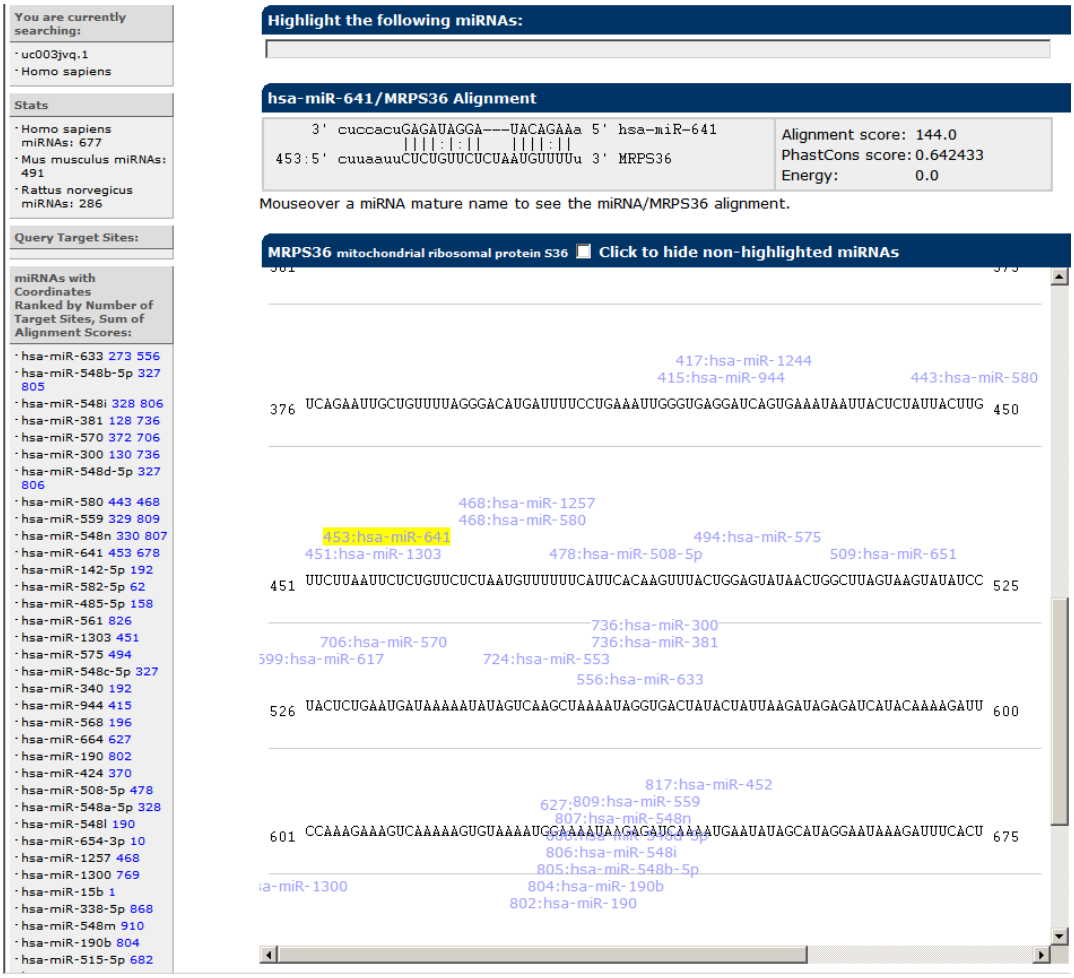


Figure 7 MRE identified by microRNA.org for Mitochondrial Ribosomal Protein S36 (MRPS36).
microRNA.org prediction of miRNAs interacting with MRPS36, a protein from the PBP. For each miRNA-mRNA alignment, a PhastCons score and the energy needed for the interaction are given. These factors set an overall score, the higher the score, the more likely for the cleavage to occur.

miRBase from the Sanger Institute, rebranded as Microcosm and managed by the European Bioinformatic Institute (EBI) during the course of this study, used the miRanda algorithm to find miRNA-mRNA complementary base pairing. Differing from microrna.org, perfect seed pairing is required. Predicted mRNA targets were obtained by accessing to http://microrna.sanger.ac.uk/cgi-bin/targets/v5/search.pl, and by entering the gene name selecting *H. sapiens* as the selected genome as shown in Figure 8.

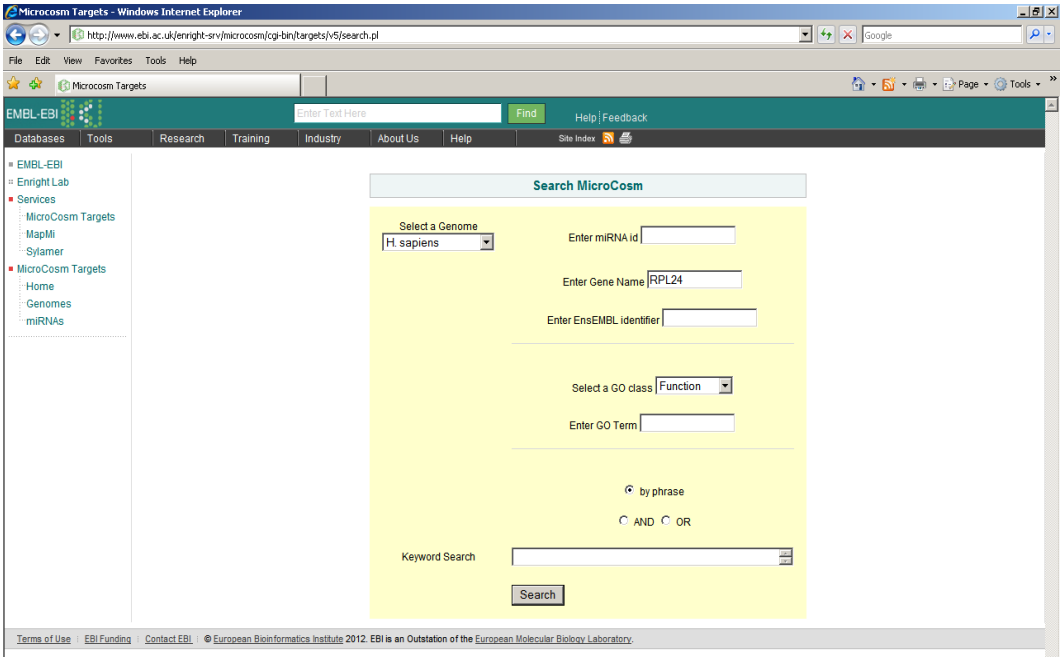


Figure 8 MicroCosm data entry screen.
A gene sensitive to Se supply in lymphocytes, the Ribosomal Protein Large 24 (RLP24), is showed as an example of data entry for miRBase/ MicroCosm using the *H. sapiens* sequence.

The displayed information now includes miRNA sequences and the mRNA binding, p-value, score, total energy required for the cleavage and the conservation across species. RPL24 was analysed by the miRBase/MicroCosm algorithm and a predicted miRNA interaction, score, orthology is represented by Figure 9.

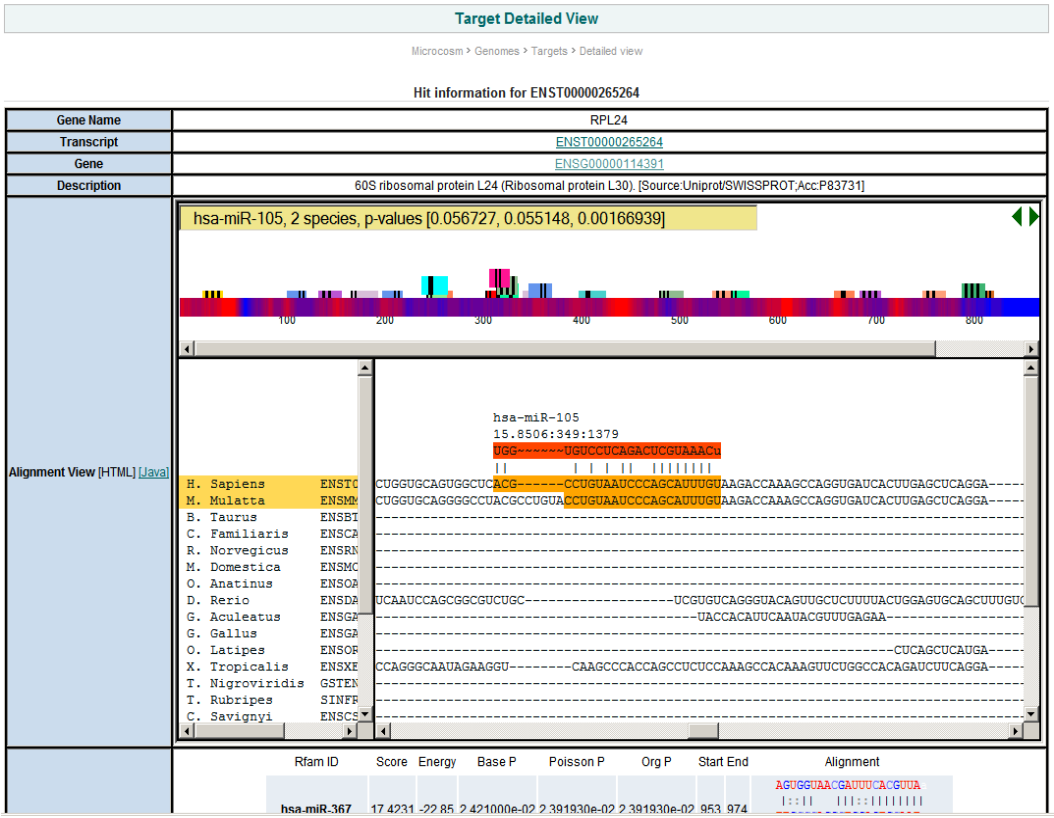
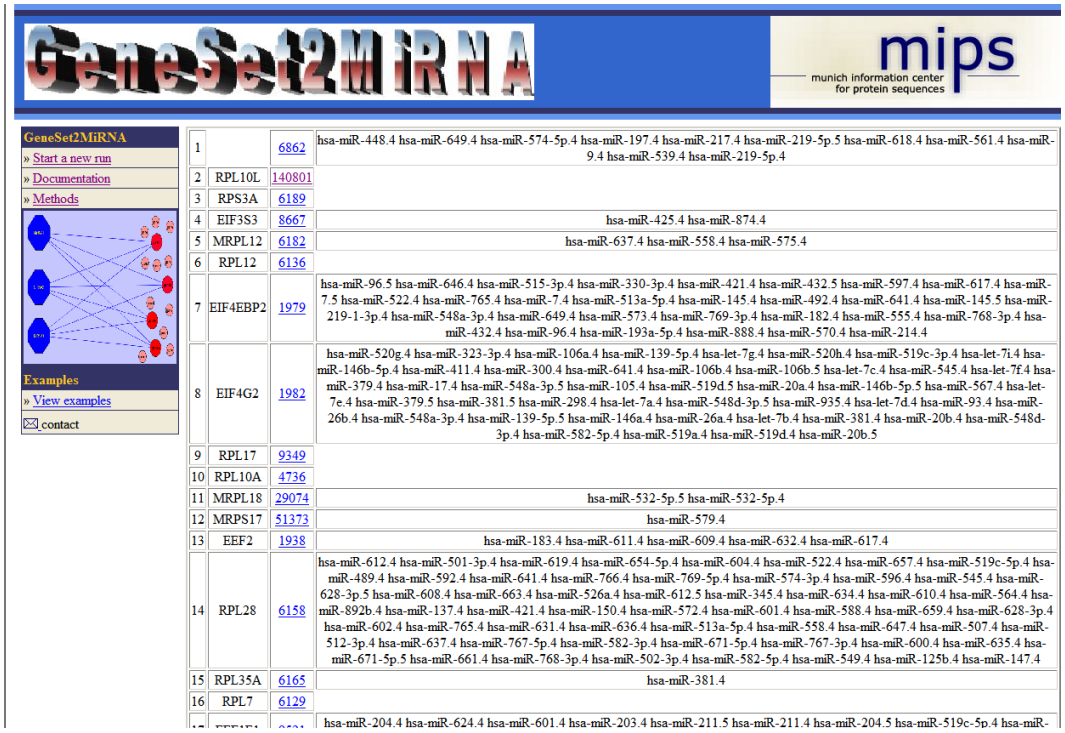


Figure 9 MicroCosm results screen.
MicroCosm MRE result page showing the alignment, conservation among species, p-value, energy and overall score of a RLP24.



1		6862	hsa-miR-448.4 hsa-miR-649.4 hsa-miR-574-5p.4 hsa-miR-197.4 hsa-miR-217.4 hsa-miR-219-5p.5 hsa-miR-618.4 hsa-miR-561.4 hsa-miR-9.4 hsa-miR-539.4 hsa-miR-219-5p.4
2	RPL10L	140801	
3	RPS3A	6189	
4	EIF3S3	8667	hsa-miR-425.4 hsa-miR-874.4
5	MRPL12	6182	hsa-miR-637.4 hsa-miR-558.4 hsa-miR-575.4
6	RPL12	6136	
7	EIF4EBP2	1979	hsa-miR-96.5 hsa-miR-646.4 hsa-miR-515-3p.4 hsa-miR-330-3p.4 hsa-miR-421.4 hsa-miR-432.5 hsa-miR-597.4 hsa-miR-617.4 hsa-miR-7.5 hsa-miR-522.4 hsa-miR-765.4 hsa-miR-7.4 hsa-miR-513a-5p.4 hsa-miR-145.4 hsa-miR-492.4 hsa-miR-641.4 hsa-miR-145.5 hsa-miR-219-1-3p.4 hsa-miR-548a-3p.4 hsa-miR-649.4 hsa-miR-573.4 hsa-miR-769-3p.4 hsa-miR-182.4 hsa-miR-555.4 hsa-miR-768-3p.4 hsa-miR-432.4 hsa-miR-96.4 hsa-miR-193a-5p.4 hsa-miR-888.4 hsa-miR-570.4 hsa-miR-214.4
8	EIF4G2	1982	hsa-miR-520g.4 hsa-miR-323-3p.4 hsa-miR-106a.4 hsa-miR-139-5p.4 hsa-let-7g.4 hsa-miR-520h.4 hsa-miR-519c-3p.4 hsa-let-7i.4 hsa-miR-146b-5p.4 hsa-miR-411.4 hsa-miR-300.4 hsa-miR-641.4 hsa-miR-106b.4 hsa-miR-106b.5 hsa-let-7c.4 hsa-miR-545.4 hsa-let-7f.4 hsa-miR-379.4 hsa-miR-17.4 hsa-miR-548a-3p.5 hsa-miR-105.4 hsa-miR-519d.5 hsa-miR-20a.4 hsa-miR-146b-5p.5 hsa-miR-567.4 hsa-let-7e.4 hsa-miR-379.5 hsa-miR-381.5 hsa-miR-298.4 hsa-let-7a.4 hsa-miR-548d-3p.5 hsa-miR-935.4 hsa-let-7d.4 hsa-miR-93.4 hsa-miR-26b.4 hsa-miR-548a-3p.4 hsa-miR-139-5p.5 hsa-miR-146a.4 hsa-miR-26a.4 hsa-let-7b.4 hsa-miR-381.4 hsa-miR-20b.4 hsa-miR-548d-3p.4 hsa-miR-582-5p.4 hsa-miR-519a.4 hsa-miR-519d.4 hsa-miR-20b.5
9	RPL17	9349	
10	RPL10A	4736	
11	MRPL18	29074	hsa-miR-532-5p.5 hsa-miR-532-5p.4
12	MRPS17	51373	hsa-miR-579.4
13	EEF2	1938	hsa-miR-183.4 hsa-miR-611.4 hsa-miR-609.4 hsa-miR-632.4 hsa-miR-617.4
14	RPL28	6158	hsa-miR-612.4 hsa-miR-501-3p.4 hsa-miR-619.4 hsa-miR-654-5p.4 hsa-miR-604.4 hsa-miR-522.4 hsa-miR-657.4 hsa-miR-519c-5p.4 hsa-miR-489.4 hsa-miR-592.4 hsa-miR-641.4 hsa-miR-766.4 hsa-miR-769-5p.4 hsa-miR-574-3p.4 hsa-miR-596.4 hsa-miR-545.4 hsa-miR-628-3p.5 hsa-miR-608.4 hsa-miR-663.4 hsa-miR-526a.4 hsa-miR-612.5 hsa-miR-345.4 hsa-miR-634.4 hsa-miR-610.4 hsa-miR-564.4 hsa-miR-892b.4 hsa-miR-137.4 hsa-miR-421.4 hsa-miR-150.4 hsa-miR-572.4 hsa-miR-601.4 hsa-miR-588.4 hsa-miR-659.4 hsa-miR-628-3p.4 hsa-miR-602.4 hsa-miR-765.4 hsa-miR-631.4 hsa-miR-636.4 hsa-miR-513a-5p.4 hsa-miR-558.4 hsa-miR-647.4 hsa-miR-507.4 hsa-miR-512-3p.4 hsa-miR-637.4 hsa-miR-767-5p.4 hsa-miR-582-3p.4 hsa-miR-671-5p.4 hsa-miR-767-3p.4 hsa-miR-600.4 hsa-miR-635.4 hsa-miR-671-5p.5 hsa-miR-661.4 hsa-miR-768-3p.4 hsa-miR-502-3p.4 hsa-miR-582-5p.4 hsa-miR-549.4 hsa-miR-125b.4 hsa-miR-147.4
15	RPL35A	6165	hsa-miR-381.4
16	RPL7	6129	
			hsa-miR-204.4 hsa-miR-624.4 hsa-miR-601.4 hsa-miR-203.4 hsa-miR-211.5 hsa-miR-211.4 hsa-miR-204.5 hsa-miR-519c-5p.4 hsa-miR-

Figure 10 Screen with the predicted miRNA from GeneSet2miRNA after the analysis of the genes within the PBP.
The miRNAs predicted by 4 or more bioinformatics tools used by GeneSet2miRNA are displayed in a chart in which a column containing miRNAs is next to the gene they have an alignment with.

GeneSet2MiRNA miRNA predictions included the miRNA-mRNA interactions that have been predicted for at least 4 other algorithms from the 9 it compares in order to minimise false positives. Another advantage is that a list of gene IDs can be analysed at the same time, so once a file in a text or excel format is uploaded into GeneSet2miRNA, a screen like the one shown in Figure 10displayed the miRNAs interacting with each gene but no further information about energy, score or position of the alignment was included.

A Core analysis with the list of genes from the PBP as a Reference Set was created in Ingenuity Pathway Analysis (IPA) Figure 23. This Network generation tool was used so pathways were created as part

of the analysis. The data sources for the analysis were the available miRNA-mRNA interactions options (miRecords, TarBase, and TargeScan Human), and selecting the experimentally observed as confidence feature through the human genome and across all tissues and cell lines.

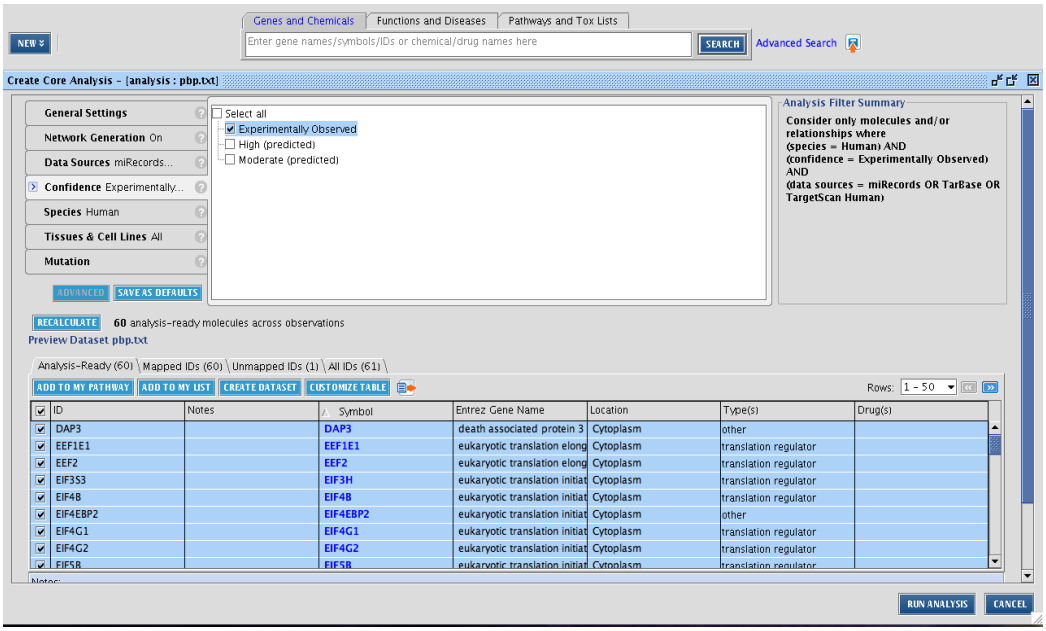


Figure 11 IPA core analysis generator screen.
The 61 genes from the PBP loaded into IPA in order to obtain miRNAs that have been shown to interact with these genes.

An image with direct and indirect protein-protein connections was created once the analysis was carried out and 9 networks were composed. Those networks were merged to produce a visual image of miRNA-mRNA interactions and direct and indirect interactions among the genes of the PBP. The resulting image was displayed as a sub-cellular level and in shown in Figure 23 in Chapter 3.

All the miRNAs from IPA were included because the miRNAs from that bioinformatic tool were already experimentally validated.

miRWalk was introduced in 2011 as a new miRNA prediction algorithm that also included a dataset of validated targets and uses eight other bioinformatic tools (Dweep *et al.*, 2011). microRNA.org started to look for mRNA-miRNA complementarity in other gene regions other than the 3'UTR since this interaction had been reported previously as also functional; for that reason, miRWalk searches for MRE on the 5 and 3' UTRs, ORF and promoter sequences for human, mouse and rat. *H. sapiens*, *M musculus* and *Rattus novergicus*.

2.2.7 *Primer design*

Primers used for RT-PCR or QRT-PCR were designed by examining the sequence obtained from NCBI and the accession number of each gene in Entrez Nucleotide database. The NCBI Blast website (NCBI) was used to assess primer design specificity I*n silico*. A 3' to 5' 20-25nt long sequence was selected bearing in mind that: the melting temperature should be between 55 and 72°C, the 3' end of the primers should have C or G, and the CG% between 45% and 55%. Assessment of CG% content of the oligos and annealing temperature of the forward and the reverse sequences so the difference is of maximum 1°C was achieved through the IDT primer design website (Integrated DNA Technologies, 2009).

2.2.8 *Primer optimization*

Annealing temperature was determined empirically by modifying a the predicted annealing temperature by the primer design websites (NCBI; Integrated DNA Technologies, 2009) by a series of by 2°C steps. The optimal annealing temperature is chosen after visualization on agarose gel electrophoresis stained with ethidium bromide after PCR with 29 cycles. The assessment of the optimal annealing temperature for *SELK*, *EPHX2*, and *MT1F* at 55°C, 57°C, 59°C, and 61°C is shown in Figure 12. After a range of temperatures have been tried and if several bands were visualised, the primers were designed again, which is the case of *SELK*. In Figure 12; on the same figure, a single band of the predicted size was visualized at 61°C for *EPHX2* and at 59°C for *MT1F*. A similar procedure is carried out to find the adequate cycle number after the ideal annealing temperature had been determined.

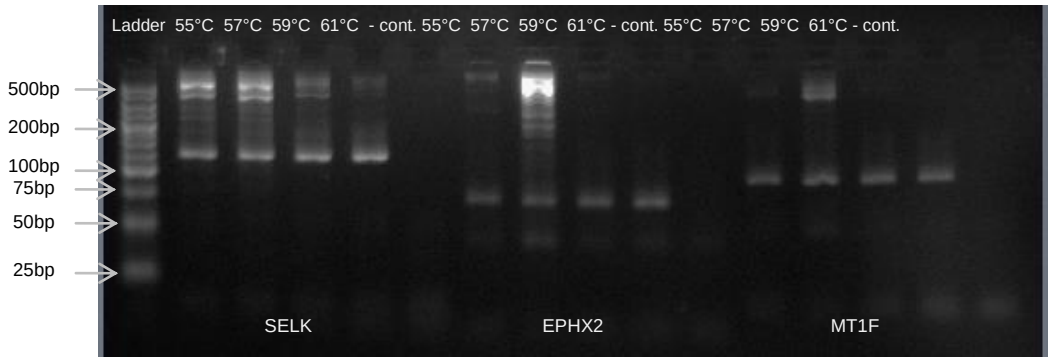


Figure 12 Annealing temperature optimization.
Shows the optimization of the annealing temperature for SELK, EPHX2 and MT1F. cDNA was PCR-amplified, at 55, 57, 59 and 61°C after 29 amplification cycles, the PCR products were electrophoresed and visualised; the annealing temperature was chosen when only one band at the predicted size was visualized.

2.2.9 *cDNA synthesis*

The RNA was synthesized from first-strand cDNA using from 10ng to 1µg of RNA template, 0.08A_{260}units of Oligo $(dT)_{15}$ primer (Roche, Germany), and up to 13µL of 0.05% DEPC water. The reaction was incubated at 65°C for 10 min and placed on ice. After incubation, the following components were added: Transcriptor RT Reaction Buffer 5x (Roche, Germany) to a final concentration of 1x, 20U of Protector RNase Inhibitor (40 U/ mL)[1] (Roche, Germany), 2 mM dNTP-Mix (dNTP; Bioline UK), and 10U of Transcriptor Reverse Transcriptase (Roche, Germany) for a 20 µL reaction. The mixture was vortexed, incubated for 30 min at 55°C, then heated at 85°C for 5 min, and placed on ice.

2.2.10 *Polymerase Chain Reaction (PCR)*

In order to assess the expression of the genes of interest a semi quantitative PCR was performed with a 20µL reaction containing the following: 0.5µL of the first strand cDNA, 3nM of $MgCl_2$ (Bioline, UK), 1.2 µL of MgCl2 solution, 2 µL 10x NH4 buffer, 3µL 100mM 1:10 dNTP Mix, 1µL of each forward and reverse primers (© Eurofins MWG Operon) in a 10pmol/µL concentration (SELW, GPX1, GAPDH, SELK, GPX3, PIR, SEPHS2, ACBT, and MT1F) and 0.25µL BIOTAQ™ DNA polymerase. The reaction was amplified under the following conditions:

Denaturing stage	95°C- 3 min		
Amplification stage			
Denaturing	94°C - 1 min		
Annealing	*55°C - 60°C	1 min	26-33 cycles
Elongation	72°C – 2 min		
Final elongation	72°C – 12 min		

* Annealing temperature for each primer as well as the number of cycles are shown in Table 1.

Primer name	Accession number	Sequence	Length (bp)	Amplicon length (bp)	Cycles	*Annealing temperature
SELW_for	NM_003009.2	GTTTATTGTGGCGCTTGAGGC	21	352	26	60°C
SELW_rev		GAACACAGGGAAAGACCAGG	21			60°C
GPX1_for	NM_000581	TTGGGCATCAGGAGAACGCCA	21	407	26	56°C
GPX1_rev		CTGCAACTGCCAAGCAGCCG	20			56°C
GAPDH_for	NM_002046.3	TGAAGGTCGGAGTCAACGGATTTG	23	112	26	55°C
GAPDH_rev		CATGTAAACCATGTAGTTGAGGTC	25			55°C

Table 1 Primers and conditions used for RT-PCR.
Gene name, accession number, sequence, primer length, amplicon size, number of cycles programmed in the thermal cycler and annealing temperature used for RT-PCR are shown for the three genes analysed.

2.2.11 *Agarose gel electrophoresis*

The DNA products of the PCR reaction were then separated by agarose gel electrophoresis. The gel was prepared by dissolving 1-3% agarose (Ultrapure agarose, Invitrogen^{TM}) in 1xTAE buffer (Tris base 0.04M, Acetate 0.02M, EDTA 0.001M, pH at 25°C 8.25-8.55; Sigma, UK) or 1x TEB (Tris-Borate-EDTA, Fisher, UK) buffer at high temperature and then adding and mixing 1microlitre of ethidium bromide before gel solidification. 8 µL of PCR product mixed with 2 µL of 5x loading dye (50 nM Tris-HCl pH8.5 nM EDTA, 20% glycerol and 0.1% bromophenol blue) were loaded into the gel. 5 µL of DNA marker (Hyperladder IV; Bioline, UK) were also loaded to indicate the size of the amplicon.

Electrophoresis was carried out at 70 V for 40 minutes and the RNA bands were visualized under UV light.

To assess the ideal cycle number for the different pair of primers PCRs were carried out by testing increasing the cycle number for each pair of primers. PCR products were electrophoresed. Figure 13 shows GAPDH as an example of cycle number assessment in which once visualized in a trans illuminater, the appropriate condition was chosen according to the brightest intensity of the band whilst in the linear amplification phase before reaching saturation.

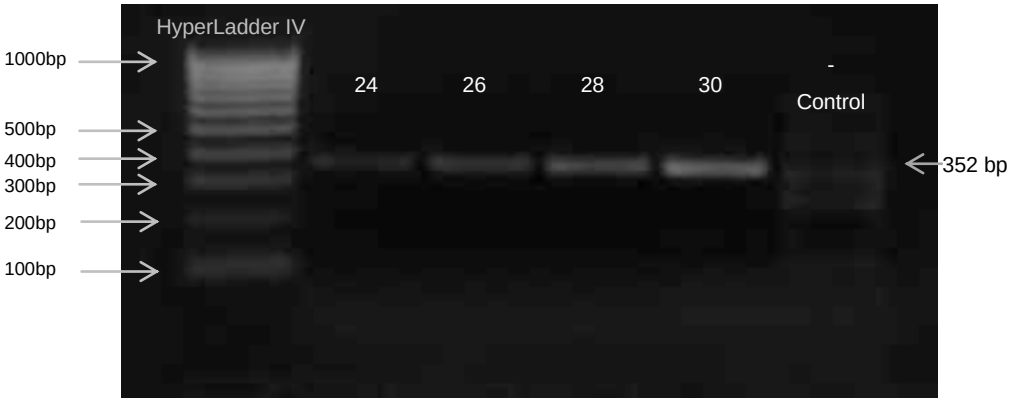


Figure 13 Experimental assessment for cycle number determination in RT-PCR for *SELW*.
For each experiment and sets of primers, an assessment of the number of cycles was performed in order to determine the right conditions (i.e. linear phase before saturation) to quantify the expression of a gene by RT-PCR.

2.2.12 *Polyacrylamide Gel Electrophoresis (PAGE)*

To prepare a 6% polyacrylamide gel, 240 g of urea, 50mL of 10x Tris/Borate/EDTA buffer (TEB), 75mL of 40% acrylamide and the volume of de-ionized H_2O required for a final volume of 500mL was warmed at 37°C; once urea was dissolved, 250 µL of 10% ammonium persulphate and 25 µL of tetramethylethylenediamine were added and immediately loaded into the appropriate glass plates inserted into the electrophoresis apparatus and filled with 0.5X TBE buffer solution.

The samples were denatured at 95°C, loaded into the 6% polyacrylamide gel and electrophoresed at 30-40mA 200V for 30 min. The gel was stained soaking the gel in a 0.1mg/mL solution of ethidium bromide in 1x TBE. The gel was washed for 5 min in 1x TBE and visualized with a UV transilluminometer.

2.2.13 *miRNA microarray design*

The human genome V2 Agilent platform is a one colour microarray consisting of short probes which hybridize to target miRNAs and which contain a hairpin loop that will join the synthesized sequence to the miRNA. The platform uses the miRNAs sequences from Sanger 10.1. In the work described in this thesis an array was customized using Agilent's eArray Design Tool. This allowed experimental confirmation of predicted miRNAs.

The design of the customised slide was 8x15K where 737 distinct *H. sapiens* miRNAs were included, distributed according to the following:

- 194 miRNAs from a review of the literature -referred to in this thesis as the intestinal microRNAome- (Bandres *et al.*, 2006; Cummins *et al.*, 2006; Han *et al.*, 2007; Ogier-Denis *et al.*, 2007; Monzo *et al.*, 2008; Peng *et al.*, 2009; Sarver *et al.*, 2009).

- 292 bioinformatically obtained miRNAs described in section 2.2.6 by miRNA target prediction (Chapter 3 of this thesis).

- 414 miRNAs were taken at random from Illumina human v2 chip platform miRNA array and PCR arrays.

- 10 *Mus musculus* miRNA as negative control.

miRNA sequence, ID and either literature review, bioinformatic or platform origin for each miRNA are shown in Appendix A.

Figure 14 illustrates the distribution of the 754 miRNA IDs listed in Appendix A which were loaded into eArray, Agilent’s tool for designing miRNA microarray chips. Twenty features per miRNA were chosen which were distributed at random across the chips.

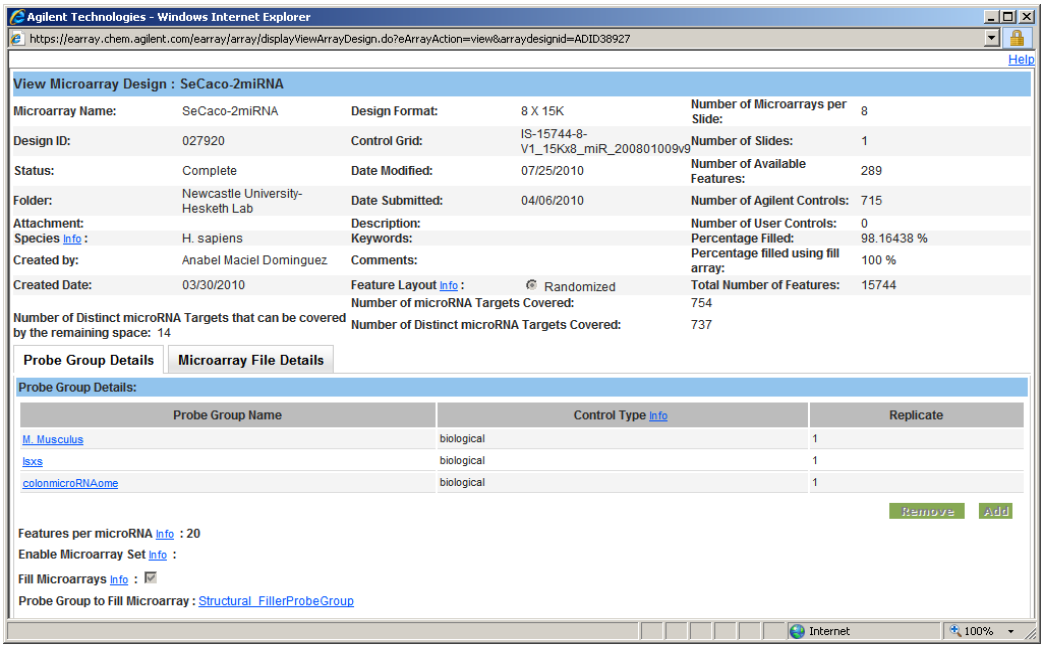


Figure 14 eArray miRNA microarray chip design summary window.
Details of the miRNA microarray chip design including the chip format, total of miRNA featured, control types, distribution of layout, number of microarrays per slide, and name of Microarray.

2.2.14 *miRNA microarray hybridization*

The miRNA expression profiling was assessed by microarray technology and the hybridization was carried out as a service provided through Warwick University.

Briefly, the expression of 737 miRNAs (predicted, characterized or experimentally validated in intestine tissue/cell lines/ biopsies) was investigated in RNA samples obtained as described in section 4.1. 100ng/ul of total RNA was labelled with Cyanine 3- pCp ligated to the 3' end. Spike ins to assess the labelling and hybridization efficiencies were included at both the labelling step and the hybridisation stage. Spike in solutions were prepared using the 3rd serial 100:200dilution of Labelling Spike-In and the 3rd serial 100:200 dilution of Hyb Spike-In.

Two µL Calf Intestine Alkaline Phosphatase (CIP) Master Mix prepared with 0.4µL 10X Calf Intestinal Phosphatase Buffer, 1.1 µL Labelling Spike-In and 0.5 µL, Calf Intestinal Phosphatase was added to each sample. Samples were dephosphorylated by incubation for 30 min at 37˚C. Samples were denatured with 2.8µL of 100% Dimethyl Sulfoxide (DMSO) and incubated for 5-10min at 100˚C. A ligation was carried out using 1µL/reaction 10X T4 RNA Ligase Buffer at 37˚C to dissolve precipitate, Cyanine3-pCp (3 µL/reaction), and T4 RNA Ligase (0.5µL/reaction) to prepare Ligation Master Mix for T4 RNA Ligase and subsequently an incubation for 2 h at 16˚C.

The labelled miRNA was purified using separation 2 mL microcentrifuge tubes with columns to desalt samples. These columns contain a buffer they were separated from the microcentrifuge tube to let the buffer drain for 2 min and then the buffer discarded. After centrifugation of the column at 1000 *x* g for 2 min; after which, the column was placed back into the microcentrifuge tube 38.7 µL of RNase-Free water was added to the labelled sample for a total 50 µL volume that were added to the previously drained gel containing spin column**.** Samples were was eluted by centrifugation at 1000 *x* g for 4 min. Samples were translucent and slightly pink after elution and were kept on ice until they were vacuumed dry at 50˚C for about an hour and the resulting pellets resuspended in 17 µL of RNase Free Water. 4.5 µL of 10X GE Blocking Agent and 22.5 µL of 2X Hi-RMP hybridization Buffer were then added to each sample to give a final volume of 45 µL that was mixed gently and then added to the array slide and incubated at 100˚C for 5 min and transferred to an ice water bath. The array slide was then placed into the Gasket Slide, and the 60-mer oligo microarray slide into the Chamber Base so 490 µL of hybridization solution and samples were dispensed into the array slide. The 2 slides were assembled into a chamber that was incubated in an oven at 55˚C and continuously mixed at 20 RMP for 20 h. After incubation, slides were washed twice with Wash solution 1 and separated carefully so the Gasket slide was submerged into a wash staining dish (containing Wash Solution 1 and 2mL of 10% Triton C-102).

The Oligo microarray slide was submerged into a second wash staining dish and washed with rewarmed (37˚C) Wash solution 2 for 5 min.

The array was scanned using an Agilent G2565CA scanner. Data were extracted using Agilent's Feature Extraction, two text files containing probe and gene level data from all features as well as array statistics and parameters and a list of expressed miRNAs was generated. In spike controls within the microarray pointed the obtained signals as biological significant rather than the result from processing issues. Because of an error message at the moment of the data extraction for one of the Se-2 sample array (a Se depleted media treated sample), possibly due to a leak, this sample was discarded.

2.2.15 *miRNA Quantitative Real Time Polymerase Chain Reaction (QRT-PCR)*

QRT-PCR for miRNA was performed using the TaqMan® MicroRNA assay (Applied Biosystems).

TaqMan® Universal PCR Master Mix, No AmpErase® UNG in combination with TaqMan miRNA assays was used for the reverse transcription of the miRNAs and further PCR amplification were carried out and analysed on an Applied Biosystems HT7900 apparatus. The miRNA assay uses a FAM labelled probe and there will be ROX within the Mastermix as a passive reference dye which might interfere with the fluorescent signal from cDNA amplification.

TaqMan® MicroRNA Assays incorporate a target-specific stem–loop reverse transcription primer to address a fundamental problem in miRNA quantitation: the short length of mature miRNAs (~22 nucleotides) prohibits conventional design of a random-primed RT step followed by a specific real-time assay. The stem–loop structure provides specificity for only the mature miRNA target and forms an RT primer/mature miRNA chimera that extends the 3' end of the miRNA. The resulting longer RT

product presents a template amenable to standard real-time PCR using TaqMan® Assays.

2.2.16 cDNA synthesis of miRNAs

Due to its short length, several approaches for the reverse transcription of RNA into its complementary DNA have been developed (Raymond *et al.*, 2005; Ro *et al.*, 2006; Sharbati-Tehrani *et al.*, 2008; Benes and Castoldi, 2010). The method utilized in this project was the reverse transcription of RNA with specific miRNA primers annealed to a prime reverse transcription stem-loop which due to its design will differentiate from mature to miRNA precursors (Benes and Castoldi, 2010).

Each miRNA was reverse transcribed with 0.15 µL 100mM dNTPs (with dTTP), 1µL MultiScribe™ Reverse Transcriptase, 50 U/µL, 1.5µL 10 X Reverse Transcription Buffer, 0.19 µL of 20U/µRNase Inhibitor, and 4.16µL of nuclease-free water for a volume of 7 µL which was mixed and placed on ice while 3 µL of 5 X RT primer and 5 µL of 10ng RNA were added to the mix. After a 5 min incubation period on ice, the 15 µL final volume reaction was transferred to a thermal cycler and reverse transcription carried out using the following conditions:

16°C	30 min
42°C	30 min
85°C	5 min
4°C	Hold

The reaction was placed on ice after cDNA synthesis and kept at -20°C.

2.2.17 *miRNAs QRT-PCR*

After cDNA synthesis for the specific miRNA of interest, the assessment of the expression by Real Time PCR was carried out by adding 1 µL of specific miRNA primers corresponding to the miRNA used in the first step of cDNA synthesis, 7.76 µL of 0.05% DEPC water, 1.33 µL of RT product (from cDNA synthesis), 10 µL of 2 X TaqMan® Universal PCR Master Mix No AmpErase® UNG to a final volume of 20µL per well. A clear 96 well-plate was centrifuged at 1,500g for 2 min to bring all reagents to the bottom and eliminate bubbles that could interfere with the correct capture of the fluorescent signal.

Using the Applied Biosystems HT7900, FAM was selected as non-fluorescent dye detection system and the Default fast run was used with the following conditions:

Enzyme activation	95°C	10 min	
Denature	95°C	15 sec	45 cycles
Anneal	60°C	60 sec	

Ct values were extracted with a threshold of 2.0 and a baseline of 2-3 cycles before the start of the amplification curve.

2.2.18 *Reference miRNAs*

A common problem in miRNA RT-PCR analysis is the data normalization. There are several approaches described in scientific publications; one approach is the development of an algorithm using the data from over 400 miRNA RT-PCRs (Mestdagh *et al.*, 2009) and another is to carry out data normalization using a miRNA reference such as let-7a (Bandres *et al.*, 2006). In this study the reference miRNA approach was used.

2.2.19 *mRNA microarray hybridization*

Caco2 cell total mRNA expression profiling was carried by using hybridization of total RNA samples to Illumina HumanRef-8 v.3 Expression BeadChips. This was carried out commercially by Service XS B.V.

Aliquots of the RNA samples previously analysed by miRNA microarray were sent to Service XS B.V. for gene microarray analysis. According to Service XS standard operation procedures, quality control was determined by Nanodrop ND100 spectrophotometer and Agilent Bioanalyzer.

Each RNA sample was reverse transcribed by converting the mRNA into a single stranded cDNA with a T7 promoter sequence using T7 Oligo (dT) Primer, 9 µL Reverse Transcriptase Master Mix and an incubation period of 2 h at 42˚C, after which, the plate was placed on ice to prepare the synthesis of a second strand; for this purpose, 80 µL Second Strand Master Mix was added to each sample (containing Nuclease-free water, Second Strand Buffer, dNTP Mix, DNA Polymerase, and RNase H) and incubated at 16˚C for 2 h in a thermal cycler previously cooled down to 16˚C and with the heat lid off. The resulting cDNA was placed on ice and purified with 180 µL cDNA Pure (magnetic beads in binding buffer), transferred to an UBottom Plate and shaken for 2 min so the cDNA within the samples bind to the magnetic beads. The UBottom Plate was then transferred to a magnetic stand to collect the magnetic beads and the supernatant was discarded by careful aspiration. The Plate was removed from the magnetic stand for the washing step in which 150 µL of cDNA Wash Buffer was added, shaking the plate at moderate speed for 1 min and the wash step repeated one more time. The magnetic beads were captured and the supernatant was discarded carefully by aspiration; the plate was then shaken vigorously for 2 mins to evaporate residual ethanol. cDNA was eluted with 20 µL of 55˚C Nuclease-free Water. 17.5 µL of the purified reaction was transferred to a PCR plate for further in vitro transcription of Biotin-labelled cRNA by adding 7.5 µL of IVT Master

Mix (containing Biotin-NTP Mix and an Enzyme Mix) to each sample and the reaction was incubated for 14 h at 37˚C. cRNA was purified with 70 µL of cRNA Binding Mix and 95 µL of ethanol, RNA Binding Beads were removed and cRNA was washed twice with 100 µL cRNA Wash Solution. Finally, cRNA was eluted with 40-100 µL 55˚C cRNA Elution Buffer.

After RNA Cy3-labeled and biotinylation into cRNA, 750ng cDNA was hybridized an overnight to Illumina HumanRef-8 v.3 Expression BeadChip containing the complementary gene-specific sequences according to Whole-Genome Gene Expression Direct Hybridization Assay protocol. A Beadchip image was produced by iScan that scans at two wavelengths simultaneously creating two images per array. after the BeadChip scan by the Illumina iScan array scanner, fluorescence signal intensity from the laser excited fluor of the hybridized single stranded products of the beads was measured for each bead location. Finally, the signals were converted into a TIFF file for transcript quantitation.

From the high resolution images, raw signals were normalized with no background adjustment using the Illuminas's Genomestudio v.2010.1 software. Internal controls (spiked-in controls and sample dependant controls) complied with specifications, gene signals were strong enough and all highly expressed genes were present and therefore signal generation was considered "passed" for analysis. Data exported from Illumina Bead Studio in a GeneSpring compatible format was loaded into GeneSpring GX 11 for analysis and analysed using the Illumina.SingleColor.HumanRef-8_V3_0_R1_11282963_A technology annotation package.

2.2.20 *mRNA QRT-PCR*

Primers were designed using with a 45-55% CG content, forward and anti-sense sequences should have no more than 2°C Tm difference; and an amplicon size of less than 150 bp to ensure the closest efficiency to 2.0 (see section 2.2.7) Gene name and accession numbers, primer

sequences, amplicon length, and annealing temperatures used for QRT-PCR are summarized in Table 2.

Primer name	Accession Number	Sequence	Primer length (bp)	Amplicon length (bp)	Start	Stop	Annealing temperature
GPX1_for	NM_000581.2 NM_201397.1	ACGAGGGAGGAACACCTGATCTTA	24	102	780	803	57°C
GPX1_rev		TTAGTGGGGAAACTCGCCTTGG	22		881	860	
GPX2_for	Nm_002083.2	AGCTCTGGGCCTTCACAGAATGAT	24	117	871	895	57°C
GPX2_rev		TCCACACCTGCCCTTTATTGGTC	24		960	984	
GPX3_for	NM_002084.3	CAGCCTGGCACAAATGGATGC	21	163	1018	1038	57°C
GPX3_rev		GTACACATTCCCAGAAAGACAC	22		1180	1159	
SELK_for	NM_021237.3	GACAACCGGACATGCGCATTC	21	116	479	499	63°C
SELK_rev		TGTGAGGACACAGAGCAGTCCT	22		594	573	
SEPHS2_for	NM_012248.2	ACGTCTCTCCCTCTAAACCCCA	22	111	2103	2124	57°C
SEPHS2_rev		TCACCAGGAATCTGCCGCAA	20		2213	2194	
ACBT_for	NM_001101.3	TGTTACAGGAAGTCCCTTGCCATC	24	90	1481	1504	60°C
ACBT_rev		CTCCCCTGTGTGGACTTGGG	20		1570	1551	
GAPDH_for	NM_002046.3	TGAAGGTCGGAGTCAACGGATTTG	24	128	113	136	57°C
GAPDH_rev		CATGTAAACCATGTAGTTGAGGTC	24		240	217	
MT1F_for	NM_005949.3	ACCTTTCTCCCAGATGTAAACAGAGAGA	28	82	313	340	60°C
MT1F_rev		GAATGTAGCAAATGGGTCAAGGTGG	25		394	370	

Primer name	Accession Number	Sequence	Primer length (bp)	Amplicon length (bp)	Start	Stop	Annealing temperature
PIR_for	NM_001018109.2	CGGAAGAGCAGGTCTTGATGTG	22	108	V1_1332 V2_1366	V1_1353 V2_1387	60°C
PIR_rev	NM_003662.3	GGAATTCTTTAGAAGCACCGGCT	23		V1_1439 V2_1473	V1_1417 V2_1451	

Table 2 Sequences, amplicon length, and conditions for primers used for QRT-PCR.

Optimal amplification temperatures were determined empirically by performing a semi quantitative PCR and visualized on a PAGE as described in section 2.2.8.

2.2.21 *Reference genes in Caco-2 cell line*

A relative quantification analysis was performed to compare the different targets normalized to an endogenous reference gene after the concentration from the samples had been determined according to the crossing point of the sample in the standard curve.

Beta actin (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), beta-2-microglobulin (*B2M*), hypoxanthine phosphoribosyl-transferase I (*HPRT1*), succinate dehydrogenase complex, subunit A (*SDHA*), and tyrosine 3-monooxygenase/tryptophane 5-monooxygenase activation protein, zeta polypeptide (*YWHAZ*) are six of the most common reference genes used when normalizing genomic data from Caco-2 cells (Piana *et al.*, 2008). The best expression stability was found in *ACTB*, *GAPDH*, *HPRT1*, the stability decreased on *YWAHZ*, *B2M*, and *SDHA,* hence of GAPDH and ACBT were used in this study in order to normalize the expression of the genes assessed by RT-PCR and QRT-PCR.

2.2.22 *PCR product purification*

In order to produce a standard curve RNA was reverse transcribed into cDNA, a series of ten 10-fold dilutions of the concentrated DNA carried out and then the diluted DNA amplified by PCR with the specific set of primers intended to be used for the standard curves. The PCR product was the purified with QIAquick PCR Purification Kit (QIAGEN, UK). After PCR amplification, 1 volume of PCR product was mixed with 5 volumes of Buffer PB and placed into a spin column and centrifuged at 13,000 RPM for 30-60 sec. After the flow-through was discarded, the

column was washed with 750 µL Buffer PE and centrifuged at 13,000 RPM for 30-60 sec. The columns were dried at 13,000 RPM. PCR product was then eluted with 50µL of Buffer EB applied to the centre of the membrane and collected in a microcentrifuge tube by a 1 min centrifugation at 13,000 RPM. Purified DNA was then stored at -20°C and used within 3 months. PCR of the purified products was used to produce a standard curve was produced for each gene of interest.

2.2.23 *cDNA amplification*

1:40 dilution of cDNA and 1:40 of each serial dilution was pipetted into a 96 well plate plus forward and reverse primers 1 pmol each, 2x LightCycler 480 SYBR Green I Master and 0.05% DEPC water to a final volume of 20µL.

QRT-PCR for mRNA assessment was run in a Light Cycler 480 Roche with the following programme:

Pre-Incubation	95°C	5min		
Amplification	Denature	95°C	10sec	45 cycles
	Annealing	*°C	15sec	
	Elongation	72°C	5sec	
Melting curve	95°C	5sec		
	65°C	1min		
	95°C	continuous		

And a final cooling step at 40°C for 30sec

* Refer to Table 2 for annealing temperature.

The quantification was programmed at the end of every amplification cycle.

2.2.24 *QRT-PCR quality control*

Since SYBR Green I will bind to any double stranded structures included primer dimers, the accurate detection of the gene of interest needs to be assed. For that purpose, the Light Cycler 480 software performs a melting curve analysis which will generate a peak according to the detected amplified products produced by the fluorescent dye at 65°C-95°C. The melting temperature of cDNA is monitored at the end of each cycle by raising the temperature by a fraction of a degree while the fluorescence is monitored; because SYBR Green does binds to any double stranded structure, in case there is presence of primer dimers or contamination, these different strands will denature and produce a different peak. As shown in Appendix B, the melting peaks for each gene of interest as well as the house keeping genes used to normalized the data showed a similar melting point represented by a single peak across all samples and standard curves. Standard curves showing the efficiency, error and slope are also presented on Appendix B as means of quality control for QRT-PCR.

2.2.25 *miRNA silencing*

miRNA silencing was performed with an antisense technology. A perfect complementary oligonucleotide (anti-miR) can bind to the endogenous miRNA to prevent RISC formation (Figure 15) and this will silence miR function and lead to down-regulation of its expression.

Figure 15 Anti-miR-185 sequence.
Perfect complementary cleavage of Anti-miR-185 oligonucleotide sequence to the mature miR-185.

Anti-miR-185 and a negative control were purchased from Applied Biosystems, UK and transiently transfected using cationic lipid-mediated gene delivery which reacts with the negatively charged phosphate groups of nucleic acids to then be easily adhered to the membrane (Karra and Dahm, 2010). This approach has been used previously to achieve anti-miR delivery and miRNA silencing.

Prior to transfection, $3.5x10^5$ Caco-2 cells were seeded and grown for 24 h. Caco-2 cells were then treated with a transfection media containing either Anti-miR-185 or Negative control. The concentrations tested were 30, 50 and 70nM for Anti-miR and Negative control. For each 6 well plate, 5 µL of Lipofectamine® 2000 Transfection Reagent (Invitrogen, UK) was added to 250 µL of Opti-MEM® (Invitrogen, UK) and incubated for 5 min ("A"); at the same time, 30, 50, and/or 70 nM ("B") of Anti-miR-185 and negative control ("C") with corresponding concentrations was added to 250 µL of Opti-MEM® and incubated for 5 min. After incubation, "A" and "B"; and "A" and "C" were mixed together, incubated for 20 min, and transfected into the cells. Media was changed 24 h after transfection.

2.2.26 *hsa-miR-185 knockdown*

Gene silencing using anti-miR often requires empirical assessment of the culture conditions required to achieve optimal knock-down of miRNA expression, and such assessment includes modifying the following conditions (Song *et al.*, 2009; Tsang *et al.*, 2010; Wang *et al.*, 2010a): oligonucleotide concentration, cell density at the time of transfection, and length of exposure. To determine the optimal conditions for miR-185 knock-down, a screen of cell density, anti-miR concentration and exposure time was carried out. The experimental set-up for this screen is illustrated in Figure 16, Figure 17, and Figure 18; The period of anti-miR treatment was varied but the total period that the cells were grown in Se-supplemented transfection medium (DMEM with 4.5 g/L glucose and Glutamax, Insulin 5 µg/mL, transferrin 5 µg/mL and Sodium

selenite 7ng/mL; 40nM) was kept constant at 72h since this period of supplementation has been used throughout the experiments described in Chapters 4 and 5.

As shown in Figure 16, to achieve anti-miR treatment for 24 h, cells were grown in 2mL of Se-supplemented transfection media for 48 h prior to transfection with the anti-miR. No penicillin was used in the medium since cell membrane permeability is increased during transfection which compromises cell survival. To achieve anti-miR treatment for 48h, cells were seeded in Se-supplemented transfection media 24 h before transfection (as shown on Figure 17). Finally, to achieve anti-miR treatment for 72 h cells were transferred into Se-supplemented transfection media and transfected at the same time, 24h after seeding (see Figure 18). In all cases the medium was changed 24 h after transfection and every 24h subsequently.

Expression of anti-miR knockdown was assessed by QRT-PCR using Taqman assays. For miRNA QRT-PCR, absolute quantification of Ct values were normalised to hsa-let-7a and RNU6B. However, miRNA-185 knockdown was found to affect the expression of hsa-let-7a but not RNU6B, and therefore in the knockdown experiments miR-185 expression was normalised to expression of RNU6B.

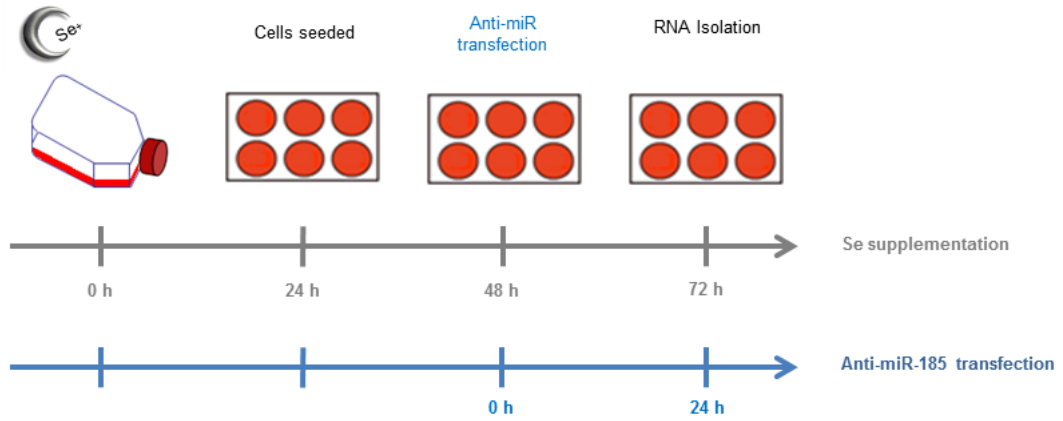


Figure 16 Experimental design of miR-185 silencing for 24 h.
The experiment was designed in order to maintain the cells under 72h of Se supplementation. Therefore, for miRNA inhibition at 24 h, cells were Se supplemented 24hrs prior seeding, transfected with 70nM of either Anti-miR-185 or negative control; RNA was isolated 24 h after transfection.

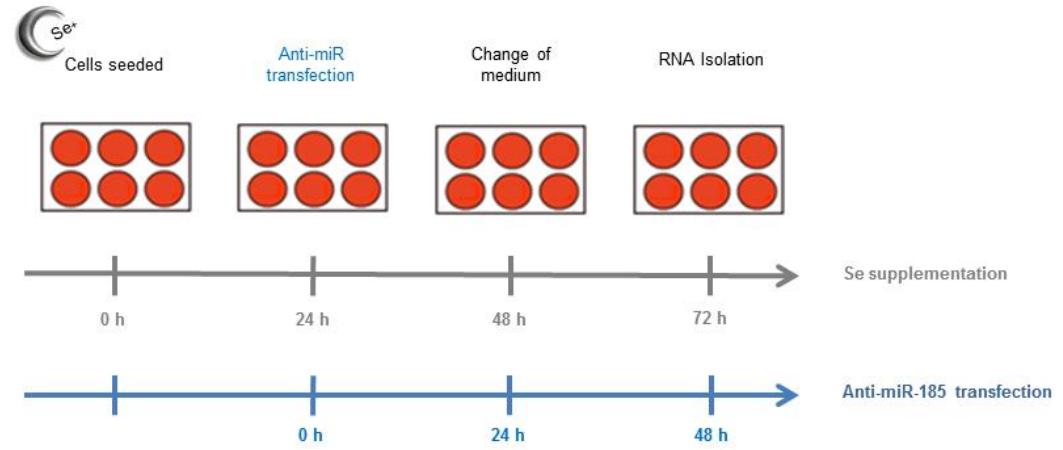


Figure 17 Anti-miR transfection optimization model at 48 h.
48h of miRNA silencing experiment, cells were seeded and Se supplemented; transfected 24 h later for RNA to be isolated 48 h after oligonucleotide treatment. Cells were fed with Se supplemented medium 24h after transfection.

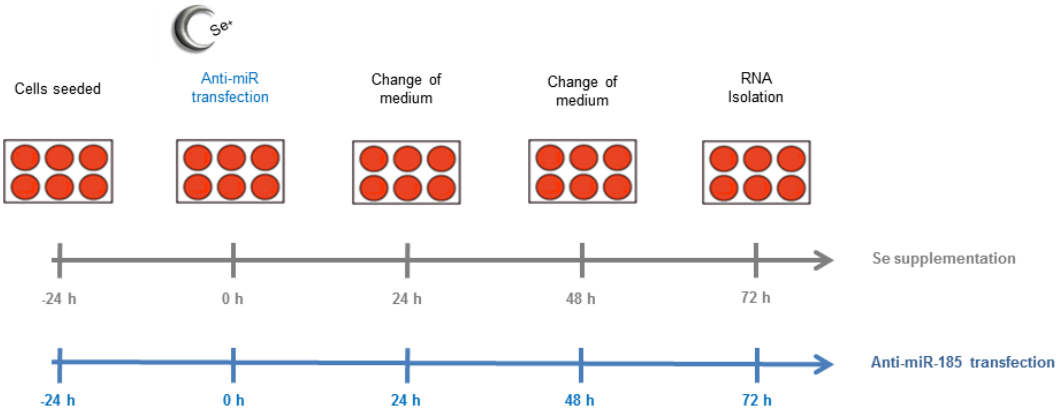


Figure 18 Anti-miR transfection optimization model at 72 h.
For the 72 h Anti-miR treatment, cells were Se supplemented and transfected 24 h after being planted; Se supplemented medium was replaced every 24 h after transfection and RNA was extracted 72 h after transfection.

As shown in Figure 19, treatment with the anti-miR for 24 h led to lower mean miR-185 expression at the 3 tested concentrations, 30, 50 and 70nM. Unexpectedly, the highest knockdown (~80%) was achieved with an anti-miR concentration of 30nM. The pattern of knockdown after 48h was different (Figure 20) with a similar knockdown of approximately 40% at anti-miR concentrations of 50 and 70nM. For anti-miR treatment for 72 h, only a concentration of 70nM experiment was used. As shown in Figure 21 with an anti-miR concentration of 70nM the extent of miR-185 knockdown fell as the length of treatment time increased from 24h to 72h. Overall, the data in Figure 19-Figure 21 suggest that the extent of knockdown fell as the length of incubation with the miR increased and treatment with 70nM ant-miRNA produced consistent knockdown.

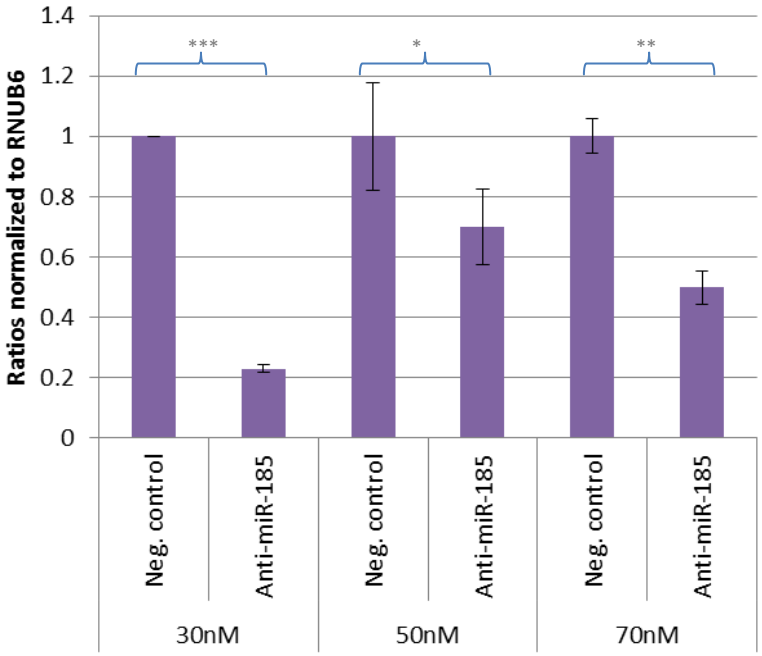


Figure 19 miR-185 silencing at 24h.
48h h after seeding, Caco-2 cells transfected with 30nM, 50nM, and 70nM of anti-miR-185 and after a further 24 h RNA was extracted 24h and the expression of miR-185 was assessed by QRT-PCR. miR-185 expression was and normalized to RNU6B levels and the data expressed relative to the expression in cells treated with the negative control anti-miR. Data shown are means =/- s.e.m, n=9.

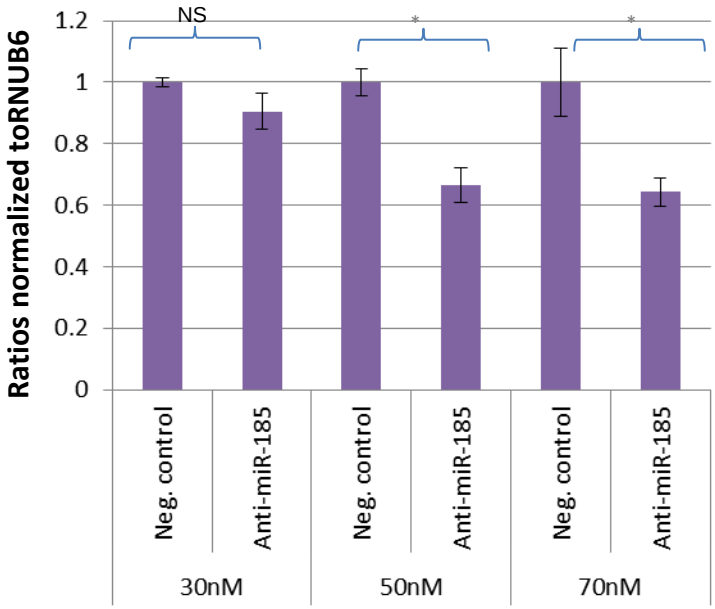


Figure 20 miR-185 knockdown at 48h.
miR-185 expression in which cells were Se-supplemented for 72h while 48h of anti-sense transfection to achieve miR-185 knockdown. RNA was extracted at 72h and the expression of miR-185 was assessed by QRT-PCR. miR-185 expression was and normalized to RNU6B levels and the data expressed relative to the expression in cells treated with the negative control anti-miR. Data shown are means =/- s.e.m, n=9.

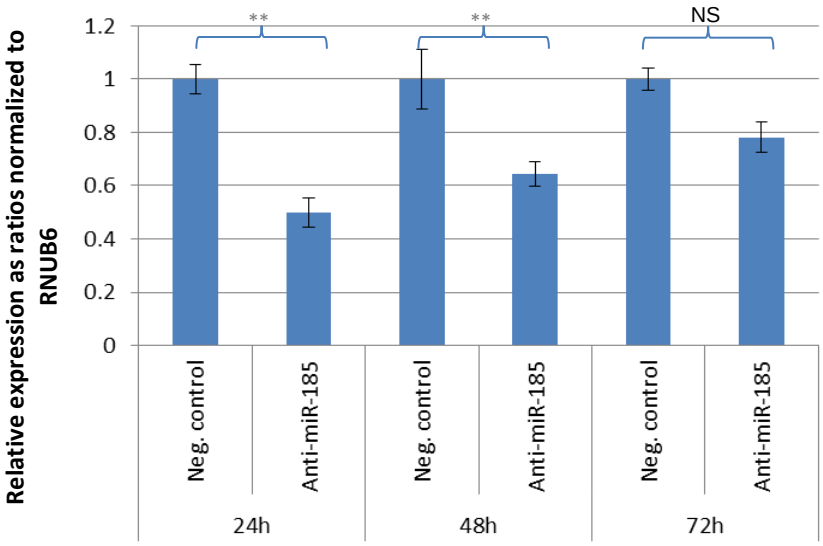


Figure 21 Expression of miR-185 after 24, 36, and 72 h of Anti-miR-185 exposure.
Caco-2 cells were treated with 70nM of either Anti-miR-185 or negative control, RNA was isolated, reverse transcribed and QRT-PCR revealed a knockdown of miR-185 normalized to RNU6B n=9.

Chapter 3 miRNA target prediction and miRNA microarray chip design

3.1 Introduction

Previous work has found changes in Se intake *in vivo* to alter expression of a number of genes e.g.(Kipp *et al.*, 2009). The work presented in this thesis was proceeded by a transcriptomic analysis in 16 healthy human volunteers who took a daily 100µg supplement of sodium selenite over 6 weeks followed by hybridization of lymphocyte RNA to DNA microarrays from 32 samples representing before and after supplementation (Pagmantidis *et al.*, 2008). The data generated from this assay was analysed by Metacore, and this software generated a biological network of 61 differentially expressed genes and the most significant network was a Protein Biosynthetic Pathway (PBP).

The transcriptomic data showed a rather small increase in expression of some genes and these were then analysed by Metacore, a pathway analysis tool, and integrated them into a protein coding network. This group of proteins included cytoplasmic ribosomal proteins of subunits 40S and 60S, mitochondrial ribosomal proteins and initiation and elongation factors. Pathway analysis pointed to ribosomal proteins and a PBP as being responsive to Se supplementation (Pagmantidis *et al.*, 2008). Another study in mouse colon investigated the mRNA expression of selenoproteins and other genes (Kipp *et al.*, 2009). For this study, GenMAPP was used for pathway analysis, finding translation factors, ribosomal protein, mRNA processing/binding, and regulation of elongation factors as well as mTOR signalling pathway to be ranked the highest. Although a different Se compound, species and tissue of interest were used for this second *in vivo* study, a protein biosynthetic pathway was also identified as affected by Se.

Although these transcriptome-level studies showed that a change in dietary Se alters expression of a range of genes, there has been no study to date of whether Se supply alters expression of miRNAs. It is possible that regulation of pathways, such as the PBP that are affected by alters Se supply may involve miRNAs. The nature of the miRNA mechanisms of action, a miRNA or miRNA cluster could regulate a group of proteins.

Many miRNAs are computationally predicted but not all are validated as being expressed in any particular cell line or tissue. Indeed, finding functional miRNAs and their targets represents a challenging task due their multiple complementarity and tissue specificity. Although many miRNA genes have been identified in the human genome search for the expression of the corresponding miRNA in a particular cell line is often the first step in understanding their role. The profile of miRNA expression in Caco-2 cells is not known. Because Caco-2 was the selected experimental model in which the investigation of Se regulated gene expression through miRNAs, it was necessary to investigate the miRNA profile of this particular cell line. The approach was to design a miRNA chip including miRNAs likely to be expressed in gastrointestinal cell lines or tissues and/or of a particular relevance of Se-regulated gene expression rather than attempt to use a genome wide miRNA array. This chapter describes the process by which a customised miRNA microarray chip was designed using a combination of 4 bioinformatic tools, a review of the relevant literature to identify miRNAs which are known to be expressed in biopsies and colonic cell lines, as well as the miRNAs for which sequences have been validated and which therefore have been previously included on PCR arrays or microarray platforms (Figure 22).

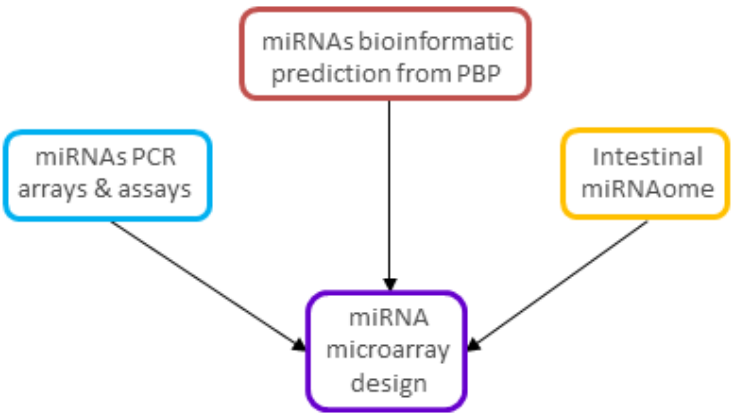


Figure 22 miRNA microarray chip design.
A customised miRNA microarray chip was designed using the most abundantly expressed miRNAs used on PCR arrays and microarrays, miRNAs predicted to target the genes from the protein biosynthetic pathway (PBP) and miRNAs expressed in intestinal tissue or cell lines identified from previous research.

For this project I used GeneSet2MiRNA (which collects information from 11 bioinformatic tools), microrna.org, MicroCosm (which are based on complementary and conserved binding predictions) and Ingenuity Pathways Analysis (which searches for protein-protein and protein-miRNA interactions) to predict miRNAs. As these miRNAs are only predicted, there is no published information about their interactions. Since metazoan miRNA cleavage does not need perfect Watson-Crick base pairing; hundreds or even thousands of targets are often the outcome of bioinformatic predictions (Jackson and Standart, 2007; Brodersen and Voinnet, 2009; Schnall-Levin *et al.*, 2011; Subramanyam and Blelloch, 2011). Therefore, by comparing the results from these tools that employ different algorithms and finding an overlap for mRNA-miRNA interactions, the aim was to increase the likelihood of finding miRNA targets to validate experimentally.

3.2 Bioinformatics tools

The design of the customized miRNA array chip was based on prediction of appropriate miRNAs that could regulate Se biological pathways using a bioinformatics tools, most of them based on the miRanda algorithm, which is designed by the Computational Biology Center of Memorial Sloan-Kettering Cancer Center. This algorithm has been used to generate many bioinformatics tools since the algorithm was

made freely available and it is flexible so that parameters can be adjusted according to requirements. Target prediction is based on complementarity between the target sequence and the mature miRNA, energy needed for the miRNA-mRNA duplex formation and orthology, focusing mainly in the 3'UTR (John *et al.*, 2004).

3.2.1 *MicroCosm Targets*

MicroCosm Targets is a bioinformatics tool developed by the Enright Lab (Grocock *et al.*). miRNA and genomic sequences are obtained from miRBase and EnsEMBL respectively. microCosm contains mature miRNA and miRNA* sequences gathered from 58 species; ~60% of the mature miRNAs within microCosm are experimentally validated (Griffiths-Jones *et al.*, 2008). Seed region complementarity, interactions within the 3'UTR, orthology alignment with at least two species and thermodynamic interactions are measured by Vienna RNA folding routines (Griffiths-Jones *et al.*, 2008) and are the criteria of miRNA-target prediction.

3.2.2 *microrna.org*

This bioinformatics tool is based on a modification of the miRanda algorithm, and focuses on almost perfect 3'UTRmRNA-5'miRNA complementary sites, which are analysed thermodynamically (Mazière and Enright, 2007). The microRNA library is based on small RNA sequence data collected from human and rodent tissues and cell lines (Landgraf *et al.*, 2007). The genome library from *Homo sapiens* (172), *Mus musculus* (64), and *Rattus norvegiocus* (16) was downloaded from the University of California Santa Cruz (UCSC) genome browser (Kuhn *et al.*, 2007).

microrna.org gives a score to the potential alignment by allowing mismatches in the seed region, but seeks complementary binding in the gene for mature miRNAs and it will give different scores to multiple

alignments of the same miRNA on the same sequence. For a miRNA to be considered in the library it must be conserved among species; the orthology is filtered by PhastCons, a conservation score that measures the probability of genome changes across time and space using Phylogenetic Hidden Markov Models (Betel *et al.*, 2008) with a cut off of 0.57. It scans the possible miRNA-gene interactions from two files, one containing miRNA sequences (file1) and the other containing RNA/DNA sequence (file2), in two stages. The first stage includes reading sequences from files 1 and 2 to find A:U and G:C complementary sites (alignment score). The second stage will then score the alignment by its thermodynamical parameters. More recently, seed site pairing, orthology, and thermodinamical stability of the miRNA-mRNA interaction were called mirSVR.

3.2.3 *GeneSet2MiRNA*

At the time of the analysis, this web-tool predicted miRNAs from a list of genes (entered by accession number or gene name) and looked for miRNA-target interactions with miRecords, an algorithm that contains documented validated and predicted targets from 9 species; these predicted targets are collected from 11 algorithms of miRNA target prediction. The GeneSet2MiRNA score is based on the number of commonly predicted miRNA hits. Unlike the previous bioinformatic tools, a complete set of genes can be uploaded into GeneSet2MiRNA, and selecting Homo sapiens as organism, the list of mapped genes with their correspondent predicted miRNAs or miRNA clusters is displayed. However, no orthology or alignment details are displayed. The score given corresponds to the number of algorithms that the miRNAs were predicted by at least 4 of these tools.

3.2.4 *Ingenuity Pathways Analysis (IPA)*

IPA is a biological pathway analysis software that provides a visual analysis of gene interactions building networks in health and disease. Biological and molecular scientific publications on *Homo sapiens*, *Mus musculus* and *Rattus novergicus*; it also supports 9 other species. IPA pathways include information from scientific articles and the chemical and physiological behaviour of other interacting components as well as miRNAs (Ingenuity, 2008) which represented a bioinformatics tool with experimentally validated.

A core analysis was created using direct and indirect relationships using the experimentally observed confidence option, species-human, and all cells and tissues, all data sources available to IPA.

3.3 Intestinal microRNA

Due to their role in different cell processes, it is not surprising that miRNAs would have tissue specificity. Since the discovery of miRNAs in 1992 and the discovery of their role in cancer 10 years later (Nugent *et al.*, 2011), there have been several reports of studies of miRNA expression using high throughput approaches. At the time of the chip design, only IPA considered miRNA tissue specificity. To include miRNAs thought to be expressed in intestinal cells, a number of publications involving the presence of endogenous miRNAs across cell lines, tissues and biopsies from healthy or metastatic intestine and even embryonic intestinal cells were also considered in the array design.

The first attempt to investigate the miRNA expression in colon was a study in which miRNA serial analysis of gene expression (miRAGE) based on cloning and a serial analysis gene expression (SAGE) was used to ligate 18- 26-nucleotide-long isolated RNA molecules by specialized linkers which were then cleaved to allow binding to streptavidin-coated

magnetic beads so that cloning and sequencing of tags enabled miRNA sequence comparison using colorectal cancer tissues. The novel miRNAs were then confirmed as Dicer dependant processing in HCT116 (Cummins *et al.*, 2006). QRT-PCR with TaqMan MicroRNA Assay kit (Applied Biosystems) on RNA from 16 colorectal cancer cell lines (using the human normal colon cell line CCD-18Co as means of comparison) and normal compared to non-tumorous colorectal tissues form USA identified 13 differentially expressed miRNAs. Their data correlated to miRNAs already identified by SAGE (Bandres *et al.*, 2006).

Later, epigenetic mechanisms of miRNA action were investigated with a focus on tissue specificity and using HCT116 and DKO cells with microarray Combimatrix and further confirmation by miRNA detection kit and QRT-PCR primers from Ambion (Han *et al.*, 2007). Ogier-Denis and Fasseu (Ogier-Denis *et al.*, 2007) reviewed the relations between intestinal pathologies and miRNAs. miRNAs were investigated in cancer cell lines as tumour suppressors or oncogenes, and in colorectal inflammatory diseases where they increase the mRNA instability of pro-inflammatory cytokines, or enhancing resistance to bacteria.

Sarver *et al* (Sarver *et al.*, 2009) compared the expression of miRNA in tumour samples against normal colon tissues from biopsies from an American population. RNA samples were analysed by Illumina miRNA BeadArray, a PCR based system which polyadenylates RNA to then reverse transcribe it to its 5'-end with a universal biotinylated oligo-dT primer sequence.

miRNAs from stem cells were considered from Monzo and Navarro's work who investigated if miRNA expression would be present in embryonic, normal and cancer tissue and assessed the relationship between the function and morphology of stem and cancer cells by analysing embryos and normal-tumour paired colorectal human samples by QRT-PCR TaqMan MicroRNA Assay. (Monzo *et al.*, 2008).

Finally, a study comparing mRNA and miRNA expression profiling efficiency to classify cancer types (Peng *et al.*, 2009) was also included into the chip design.

3.4 miRNAs from PCR arrays and assays

Since the discovery of miRNAs, several approaches have been developed for the identification of these new small RNA species. However some of these contain probes that are mis-assigned (Sales *et al.*, 2010) which leads to incorrect annotations in public databases. Often, snoRNA, snRNA and tRNA sequences are included in miRBase (Langenberger *et al.*, 2011); examples of such an incorrect annotation are miR-1274a and 1274b that share identical stretches with tRNA such as tRNA (Lys5) and tRNA (Lys3); (Schopman *et al.*, 2010).

In order to minimise the inclusion of mis-assigned miRNAs, I decided to include miRNAs from miRNA PCR arrays and microarrays because these sequences represent the most abundantly expressed and previously characterized miRNAs used by Illumina and Invitrogen.

3.5 Results

The protein biosynthetic pathway was loaded into MicroCosm Targets, selecting the genome *H. sapiens*. An interactive map with the alignment throughout the species was displayed as well as the score that represented the complementarity (the closer to 100 the more Watson-Crick base pair bound), energy required for cleavage and the sequence location of the miRNA. For microrna.org, the 61 Se-sensitive proteins were uploaded into the website as target mRNA, with search options *H. sapiens* as species. Alignment details, orthology and score were displayed. miRNAs that were predicted to target the genes of the PBP were extracted from MicroCosm, microrna.org and GeneSet2miRNA. A large number of entities were obtained from the bioinformatics tools, so the miRNAs were selected by taking the miRNAs that had the most number of interactions

with the same mRNA sequence. Examples of the readout of the analyses are shown in Figure 23 and Figure 24, and the summary of these data presented in Appendix A.

miRNA from IPA were obtained from a network image in which miRNA-mRNA direct interactions were obtained using the grow tool. The general settings used are previously described in Chapter 2 section 2.2.6 *miRNA target prediction for the protein biosynthetic pathway*; which briefly comprises: growing out all molecules from experimentally observed direct interactions in all human tissues and cells. The analysis as subcellular layout is shown in Figure 23 where direct and indirect protein-protein interactions as well as the direct miRNA interactions with these proteins are visualised. Direct interactions are represented by a continuous arrow whereas the indirect interactions have a dotted arrow. GeneSet2miRNA predicted 225 miRNAs, MicroCosm 83, microRNA.org 31, IPA 43; because several miRNAs were predicted by these tools more than once, the results from the various bioinformatic tools described above were merged to give a final number of 292 miRNAs.

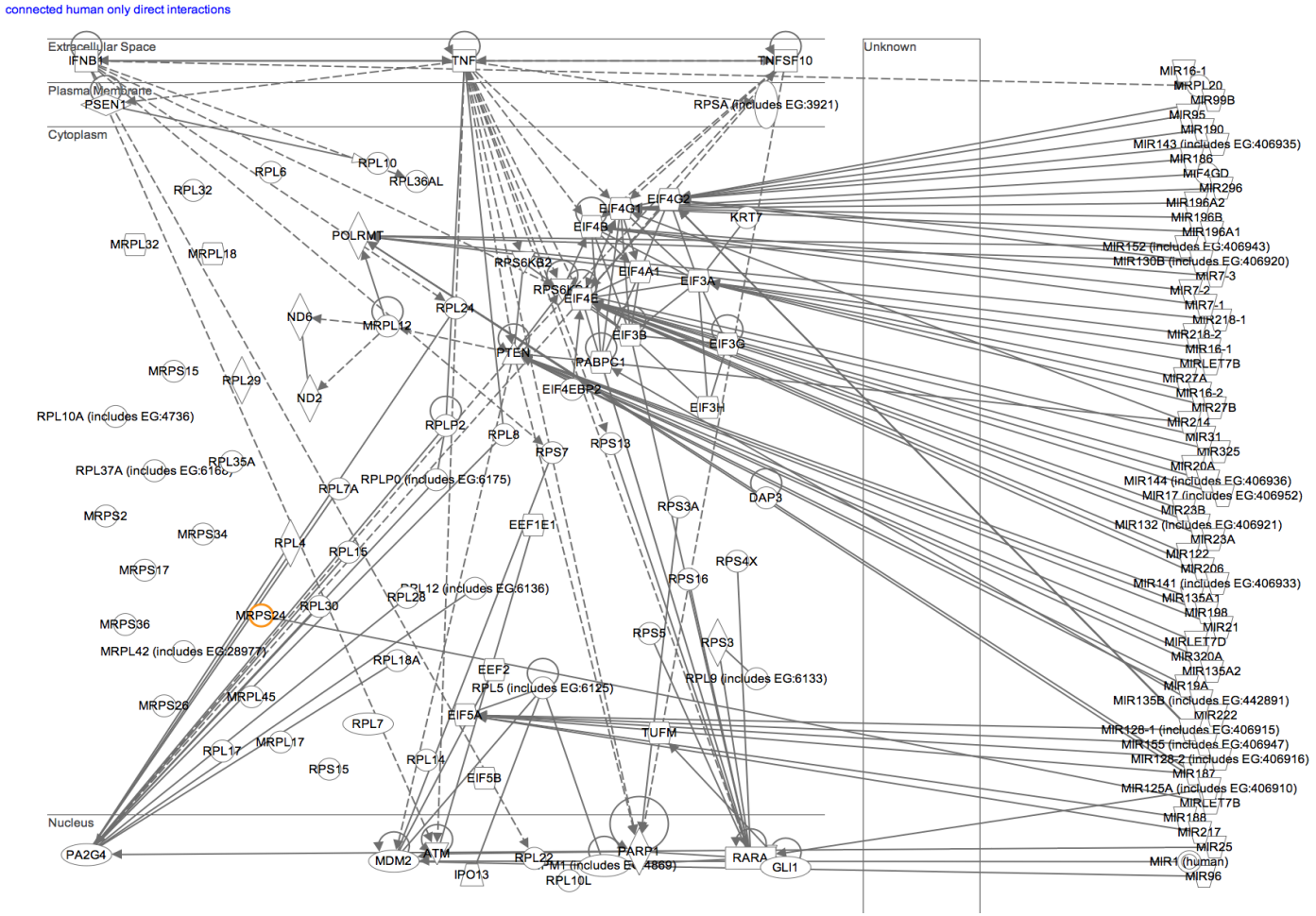


Figure 23 miRNA identification by IPA
Merged networks from a core analysis of the protein biosynthetic pathway showing experimentally validated direct miRNA interactions and protein-protein interactions as well as cellular protein localization.

Since IPA offers experimentally validated miRNAs and GeneSet2miRNA had already a filter of at least 4 bioinformatic tools for miRNA to be considered all the miRNAs from IPA and GeneSet2miRNA were included in the design. The fact that a miRNA targeted the same gene sequence a minimum of 5 times was the inclusion factor for miRNAs predicted from Microcosm and microrna.org. Lastly, mature miRNAs as well as mRNAs* were included.

In summary, and as shown in Figure 24, 61 genes were uploaded into each bioinformatics tool to predict miRNAs that may target them. microRNA.org predicted 1845, 1583 emerged form from MicroCosm, 446 from GeneSet2miRNA and 54 from IPA. miRNAs predicted from microRNA.org and MicroCosm were filtered according to the number of interactions each miRNA had with the same mRNA sequence (for 5 times or more) to give a total of 31 miRNAs from microRNA.org and 83 from Microcosm. miRNAs from IPA and GeneSet2miRNA were not filtered in this instance because IPA miRNA-mRNA interactions and GeneSet2miRNA miRNAs were both previously validated.

Because some miRNAs had more than one target and the cleavage with the same mRNA sequence occurred more than 5 times, some miRNAs were listed more than once after the first filter. So the resulting miRNA list was to eliminate the repeats from the 4 lists generates on the previous step which the numbers had been reduced to 18 miRNAs from microRNa.org, 52 miRNAs from MicroCosm, 43 from IPA, and 228 from GeneSet2miRNA. Finally, these miRNAs overlapped and when this overlap was removed there was a total of 292 miRNAs that had passed the filters from the bioinformatics tools and the software; these were included on the miRNA microarray chip.

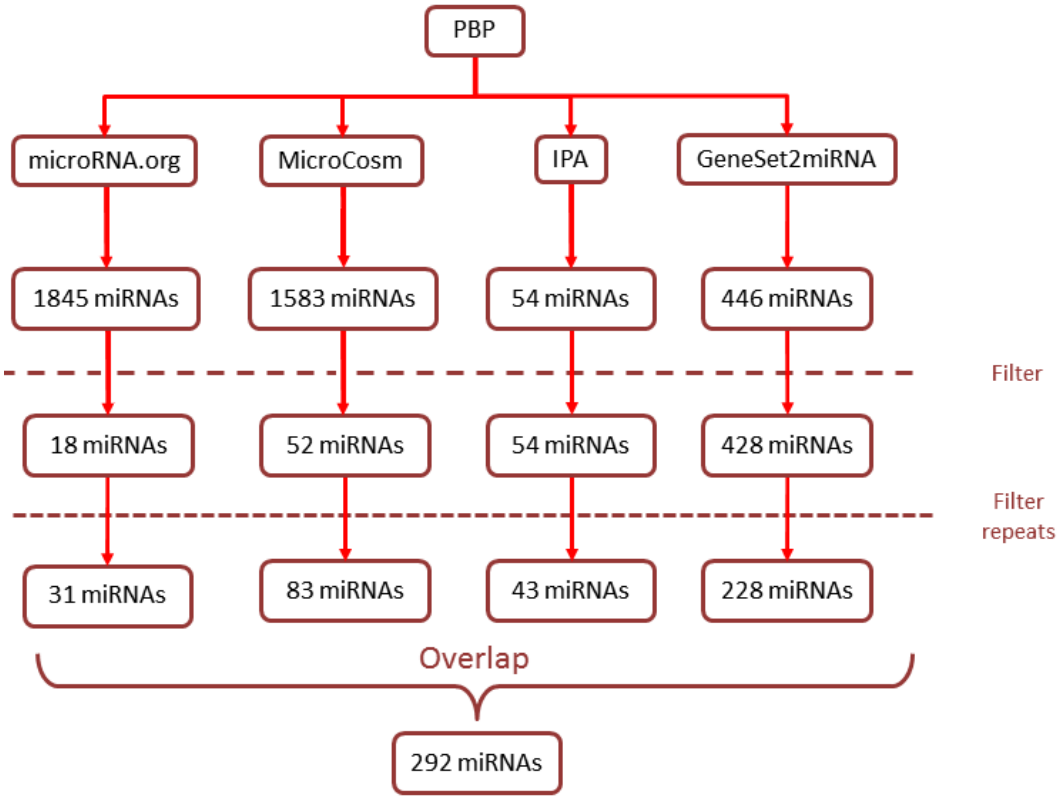


Figure 24 Bioinformatically predicted miRNAs.
61 genes from the protein biosynthetic pathway (PBP) were used to predict miRNA-mRNA interactions using 4 bioinfomartic tools: microRNA.org, Microcosm, GeneSet2miRNA, and IPA. miRNAs extracted from each bioinfomartic tool were filtered according to the number of times the same miRNA interacted with the same mRNA sequence at least 5 times (except for IPA from which all the miRNAs were included. On the second filter, miRNAs that had several targets within the PBP were counted as 1.

As shown on Figure 25, the number of miRNAs included on the chip that had been extracted from literature review are: 152 miRNAs from miRAGE technology (Cummins *et al.*, 2006); real time PCR that lead to the identification of 13 miRNA differentially expressed in cancer tissues (Bandres *et al.*, 2006), 21 miRNAs from the effects of methylation and miRNA expression research using real time PCR and microarray (Han *et al.*, 2007), finally both microarray and real time PCR studies that identified 101 miRNAs expressed in colon from embryonic and cancer tissues (Monzo *et al.*, 2008). In addition, 17 miRNAs were included from a (Ogier-Denis *et al.*, 2007) a review of functional miRNAs in intestine inflammatory processes (Ogier-Denis, Fasseu et al. 2007).27 miRNAs were included from an investigation of the relationship between epigenetics, miRNAs and their effects on tumour development (Sarver *et al.*, 2009), and 23 miRNAs

found associated with oncogenes (Peng *et al.*, 2009); both these studies had used a microarray approach.

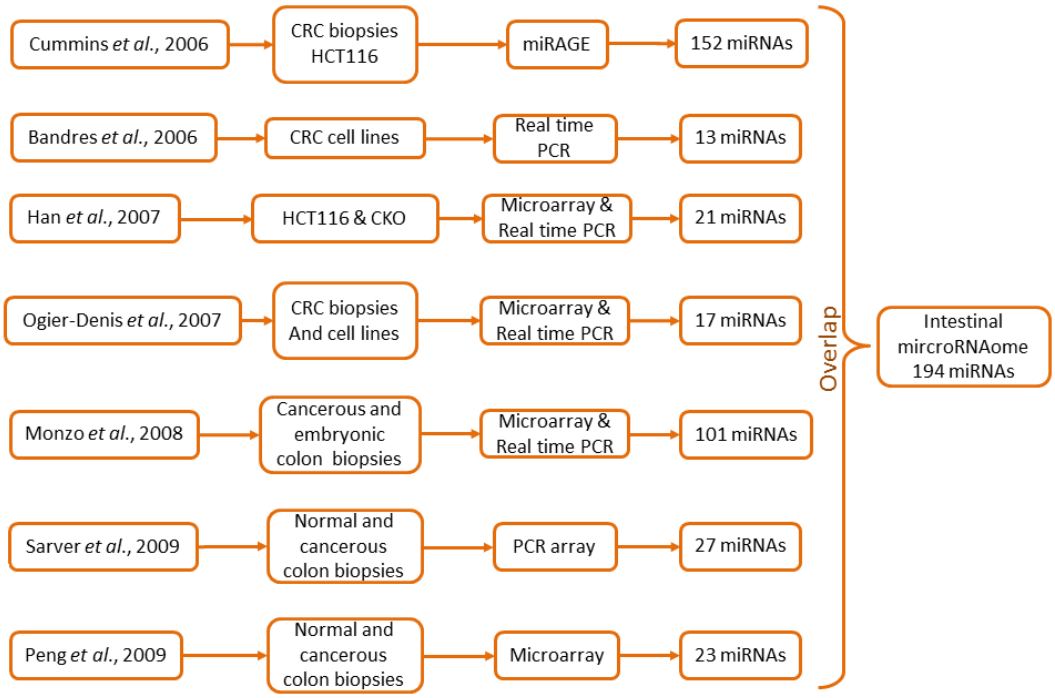


Figure 25 miRNAs from literature review.
A flow chart of the publications available up to the moment of the chip design and the technologies such as real time PCR, microarray, PCR array and miRAGE, used identify the expression of new miRNAs in colorectal cancer (CRC) samples, normal colon tissue, embryonic colon tissues, and colon cancer cell lines. These 7 lists were combined and an overlap of miRNAs from the 7 sources reduced the number to a total of 194 miRNAs which constituted the intestinal miRNAome and were included on the Agilent microRNA microarray chip.

Tools to identify miRNAs have been developed increasingly since 2002. To identify miRNA expression in 2009, microarrays and PCR platforms contained the best-characterized and experimentally validated miRNAs. A number of miRNAs from the Illumina human v2 chip platform and commercially available PCR arrays were selected and included in the chip design. These corresponded to the miRNA sequences which were best characterized and some of which were experimentally validated. In addition, 414 miRNAs were taken randomly from Illumina human v2 chip platform miRNA array and PCR arrays from.

Finally, as internal negative control 10 miRNA from *Mus musculus (M. musculus)* were also included in the design, these were not expected

to be expressed in Caco-2 cells and so were used to monitor false positives.

Note from Venn diagram on Figure 26, that there is an overlap in the miRNAs identified from the different sources, where 81 miRNAs are common to those on the intestinal miRNAome and those originated by bioinformatic analysis (53 miRNAs from commercially available platforms, the Intestinal miRNAome, and bioinformatically predicted; and 28 from Bioinformatical predictions and the intestinal miRNAome. These two sets of miRNAs could be likely to be expressed on Caco-2 cells and may have a link with the functions of Se.

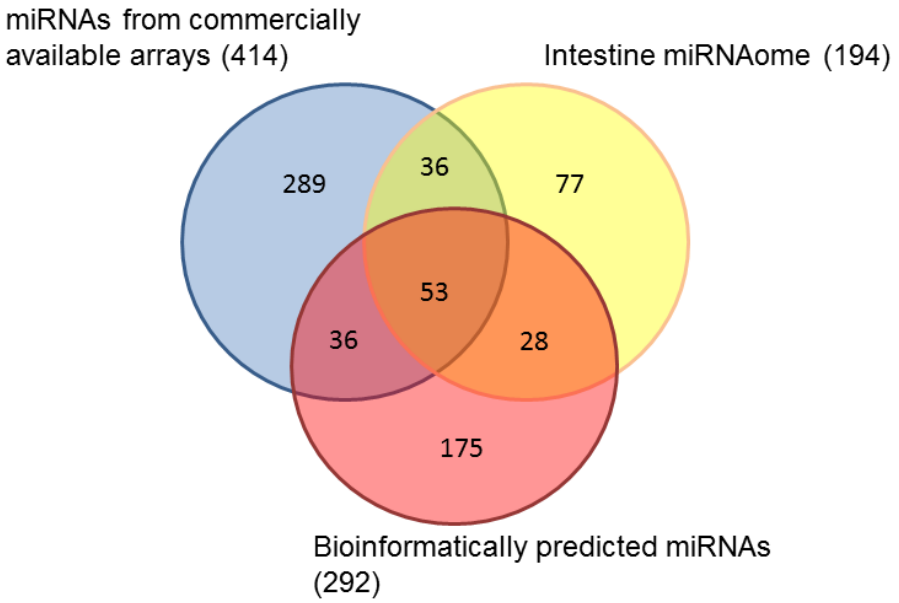


Figure 26 Venn diagram illustrating the miRNAs within the miRNA microarray design.
miRNAs were obtained from research articles in which the expression of miRNAs was identified across several intestine biopsies, tissues, and cell lines (Intestinal miRNAome); Bioinformatic predictions on Se downstream targets (Bioinformatically predicted miRNAs); finally, the best characterized mature miRNA sequences in 2010 (miRNAs from commercially available arrays).

A comprehensive list of predicted miRNAs through microrna.org, Ingenuity Pathway Analysis, GenSet2miRNA and MicroCosm, as well as the miRNAs expressed in the intestine, those obtained from commercial assays; and finally, the *M. muluscus* miRNAs used as negative control are in Appendix 1.

From a more detailed analysis of the group of 81 miRNAs common to the bioinformatically and literature review obtained, the le-7 family, miR-200 and its isoforms and the miRNAs miR-141 and miR-16 were present. Let-7 family is a widely expressed miRNA family across several organisms that is involved in various cell processes, acts as a tumour suppressor and even miRNA processing.

3.6 Discussion

Due to its flexibility to target mRNA through an imperfect complementary base pair match, miRNAs can have thousands of mRNA targets. In order to narrow down the possible mRNA-miRNAs interactions, bioinformatics tools are often used as the first step in miRNA research (Huang *et al.*, 2011b). Algorithms give the researcher the flexibility to choose the parameters of the search, as 5'UTR, 3'UTR or ORF search, and they rank the likelihood of mRNA-miRNAs interaction through a series of thermodynamics and orthology. However, mature miRNAs sequences are not necessarily conserved across species and these miRNA target prediction tools do not consider miRNA tissue specificity; finally, false positives on predictions are also common.

In order to address some of these issues, in this study we decided to scan mRNA sequences to find miRNAs that would target them with several bioinformatics tools were used with either by using the 3'UTR or the mRNA as a starting point for miRNA predictions, including all the experimentally validated miRNAs from IPA and narrowing down the miRNAs that will target the same gene several times. GeneSet2miRNA (bioinformatic tools that uses 11 other tools as filters and to support its predictions) has the largest output over the other miRNA prediction tools, and since it already takes into account at least 4 of its filters, all the predictions were included. All these measurements were taken to minimise the amount of data generated bioinformatically as well as reduce false positives.

Tissue specificity is not considered as a predicting factor because miRNA roles in cellular processes are not well understood and the expression of these genes is slightly different across cell lines which is why the expression of miRNAs in a given tissue cell line cannot be predicted easily. Thus, for this study, experimentally validated endogenous intestine miRNAs, regardless of any changes in expression, were part of the miRNA array. This way, miRNAs lost during miRNA prediction due to

lack of orthology are not being strong candidate for mRNA targeting but could be expressed in Caco-2 cells were investigated.

Because the aim of the study was to investigate the role of Se on Caco-2 cells, miRNA predictions have not been the focus of the miRNA microarray chip design, but rather the focus was miRNAs for which their mature sequence from microarray or PCR array experiments have been defined, as well as miRNAs expressed on intestinal samples. A detailed list of the miRNA chip design as well as the origin of those miRNAs (bioinformatical, literature searched, or mature miRNA characterized sequences) is presented in Appendix A.

Interestingly, a total of 81 miRNAs were found to be bioinformatically predicted and part of the intestinal microRNAome. The miRNA microarray would confirm whether those miRNAs are present on the experimental model Caco-2 cell line. Some of these miRNAs have been studied previously: for example, previous work on colorectal microdissections from German patients suggested that miR-143 and miR-145 (among others) might have a role on metastasis progression (Kahlert *et al.*, 2011); miR-200 is associated with epithelial cancer progression and cancer stem-like cells proliferation; miR-200 family is a tumour suppressor miRNA (Bao *et al.*, 2011). Homologous of miR-200 (Dykxhoorn, 2010), miR-200a, miR200b, miR-200c, and miR-141 (except miR-429), are miRNAs from the cluster 17-92, an oncogene promoter family (Tsuchida *et al.*, 2011) were bioinformatically generated and part of the literature review lists included in the miRNA microarray chip design. Colorectal microdissections from German patients suggested that the miRNAs miR-19b, miR-194, let-7b, miR-1275, miR-143, miR-145, and miR-638 might have a role on metastasis progression (Kahlert *et al.*, 2011).

Chapter 4 Experimental validation of predicted miRNAs and the effects on Se supplementation on Caco-2 cells

4.1 Introduction

Prediction of miRNA-mRNA interactions using a range of algorithms, databases (Kuhn *et al.*, 2008; Huang *et al.*, 2011b) or previously published data (Nugent *et al.*, 2011) often is the initial approach to studies on miRNAs, and was the approach also adopted for this work as described in Chapter 3. Using the list of possible miRNAs expressed in the intestine generated in Chapter 3, a high-throughput microarray-based assay to examine the expression of miRNAs in Caco-2 cells was carried out. An additional aim was to determine the effect of Se supply on their expression. There was little experimental data available to assess whether or not predicted through the bioinformatic analysis miRNAs were expressed in gut epithelial cells, and in particular the Caco-2 cell line. This chapter presents the results of microarray hybridization of 100ng/µl of total RNA from Caco-2 cells to detect expression of the predicted small RNA species, and by using RNA from Caco-2 cells cultured under both replete and deficient Se, to determine the effect of Se on their expression.

Since the discovery of miRNAs almost 20 years ago (Lee *et al.*, 1993) several techniques have been developed to assess miRNA expression; among them northern blot, QRT- PCR, Next Generation Sequencing (NGS), and miRNA microarray profiling. According to experimental design, some of these approaches are more suitable than others.

Northern blot is a widely used technique for miRNA detection; it provides high sensitivity but requires high input of RNA sample (Ro *et al.*, 2006; Sharbati-Tehrani *et al.*, 2008). The approach also requires individual

detection arrays, so they can be very costly, particularly because Lock Nucleic Acid (LNA) technology is often used.

An alternative is to use miRNA QRT-PCR, which involves amplification using specific probes. Both northern blot and miRNA QRT-PCR are highly sensitive and they are designed for individual miRNA investigation rather than investigation of groups of multiple miRNA simultaneously.

miRNA PCR arrays are based on QRT-PCR technology and will assess the expression of multiple miRNAs at the same time. miRNA probes are produced in 96-well plates or 384-well format; miRNA in the test sample are converted into cDNA by reverse transcription, which can be achieved by poly A tailing or LNA. In addition to the plate, a thermal cycling signal is generated in wells containing primers for a miRNA derived in the sample. This approach was not appropriate for the present study because the miRNAs had not been experimentally validated and available PCR arrays contained miRNA sequences only for characterized miRNA.

Assessing several miRNAs in a single assay, both to monitor the expression of miRNA and to discover novel miRNAs can be achieved through microarray technology or NGS. Like the alternative techniques already mentioned, these approaches also have their challenges and limitations. Examples of limitations of NGS are sequencing errors (Git *et al.*, 2010), high requirement for concentration of RNA than for microarrays, and poorer sensitivity than the microarrays. On the other hand, sequencing can be more reliable when detecting miRNA families that may vary by a single nucleotide in a short sequence moreover, NGS has been found to have better reproducibility (Willenbrock *et al.*, 2009) and a high ability to discover novel miRNAs because, it does not rely on use of predetermined probes (Buermans *et al.*, 2010). To achieved stable melting temperature, microarray platforms have been developed with LNA bases, the use of pre-miRNA sequences (which only gives an understanding of the miRNA processing rather than monitoring their expression) or the addition of a stem loop (Git *et al.*, 2010).

miRNA microarrays can be analysed using different platforms. It was reported that the Ambion platform was better in distinguishing signals from the background that the Agilent platform; Agilent however, included mature specific miRNA sequences and it is among the platforms with the highest scores in terms of accuracy (Git *et al.*, 2010). For analysis of two tissues containing either a small miRNA population (human liver) and a large miRNA population (human prostate), the Agilent platform showed the best reproducibility compared with QRT-PCR assays and with other microarray platforms. The reproducibility of the data was attributed to the technology used for the Agilent platform dynamic hybridization (Sato *et al.*, 2009).

As illustrated in Figure 27, this chapter describes the investigation of the effect of Se supply on miRNAs via miRNA profiling by means of miRNA microarray and the results were validated using QRT-PCR. Agilent was chosen as the microarray platform because as noted above, comparison platforms revealed it to have the highest sensitivity (Sato *et al.*, 2009) and because it offered flexibility for array design through eArray, allowing specific miRNAs of interest to be included on the chip.

For a group of miRNAs identified by this microarray analysis the confirmed as Se-regulated by QRT-PCR; bioinformatic analysis was carried out to identify their mRNA targets.

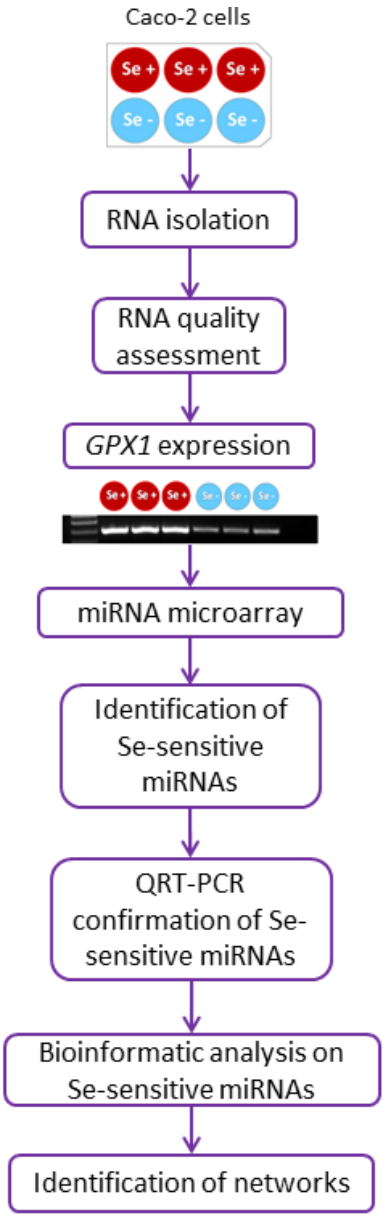


Figure 27 Chapter 4 workflow.
Caco-2 cells were treated with 40.5 nM sodium selenite and RNA was isolated after 72 h. RNA passed the quality assessment with RIN values >9.1, 260/280 ratio >1.8 and 260/230 ratio >2.0. The expression of selenoproteins is highly dependent on the Se available to the cells and the cells' responsiveness to sodium selenite was checked by measurement of expression of *GPX1* or *SELW*. After RNA passed quality assessment, and difference in Se status confirmed by *GPX1* and *SELW* expression samples were sent for miRNA microarray analysis where Se-sensitive miRNAs were identified. These miRNAs were analysed with a bioinformatics tools to predict the biological processes that the Se-sensitive miRNAs might be interacting with using miRWalk and IPA.

4.1 Results

miRNA microarray analysis of RNA from Se supplemented/deprived Caco-2 cells requires good quality RNA in order to achieve reliable expression data for these experiments, so the quality of RNA was analysed by gel electrophoresis, RNA Integrity as measurement on a Bioanalyzer, and nucleic acid purity assessed by absorbance using a Nanodrop.

As described in materials and methods, Caco-2 cells were seeded into 6-well plates. Briefly, $5x10^5$ cells were seeded into each well; 24 h later, cells were fed with either Se sufficient or deficient media, medium was changed every 2 days, and after 72h, total RNA was isolated. Either 3 or 4 biological replicas per treatment per experiment were included.

4.1.1 *Quality control of RNA used for microarray analysis*

RNA was extracted using the PureLink RNA kit QIAGEN, UK according to manufacturer's instructions to achieve total transcriptome extraction, including a high yield of miRNAs. This method uses a spin column technology based on RNA binding properties in the presence of high salt concentrations (Rump *et al.*, 2010). In order to check that the total transcriptome RNA isolated contained small RNA species such as miRNA, RNA was denatured and subjected to electrophoresis on a 6% acrylamide gel. The ethidium-bromide stained RNA fragments were visualized using an UV transilluminator. Specific miRNAs cannot be identified without specific probes therefore small RNA populations were assessed by visualization of the tRNA and the 5S and 3.5S fractions.

As shown in Figure 28, results from the electrophoresis of the RNA samples showed that all eight samples prepared for microarray analysis exhibited three clear bands, one at 175nt and two of ~300nt; these correspond to the expected sizes of 5S and 3.5S rRNA. The 4 lanes

of the left show the RNA samples from cells treated with Se whereas the 4 lanes on the right are samples from cells grown without Se. Bands corresponding to tRNA (100nt), and a small RNA population (75nt) were also observed in all eight samples. The presence of these bands implies the presence of miRNAs.

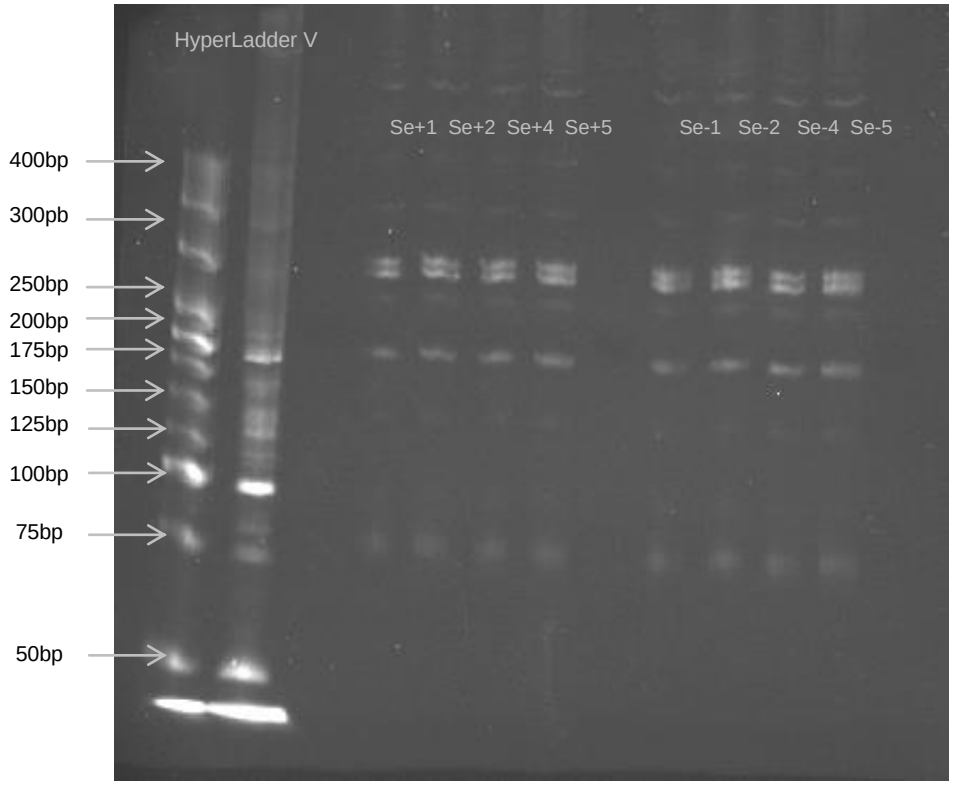


Figure 28 polyacrylamide gel of small RNA species.
RNA samples from Se depleted (right lanes) Se supplemented (left lanes) Caco-2 cells resolved by electrophoresis through a 6% polyacrylamide get. The presence of the 5S and 3.5S rRNA and tRNA population revealed the presence of low molecular weight RNA species.

RNA is susceptible to RNases that could digest it into short fragments, so reducing its integrity and potentially interfering with downstream applications by producing a difference in gene expression proportional to the level of degradation. In order to assess the RNA integrity, samples were analysed with Agilent 2100 Bioanalyzer which separated the RNA by microcapillary electrophoresis according to their molecular weight. The 28S and 18S ribosomal RNA can be visualized using fluorescent detection.

As shown Figure 29A, two major bands corresponding to 28S and 18S ribosomal RNA were observed.

Figure 29B shows a graphical representation of the 18S and 28S rRNA peaks, from which early stages of RNA degradation can be monitored. The early stages of RNA degradation can be detected when the 28S peak is lower than the 18S rRNA peak. The RNA Integrity Number (RIN) is calculated from the peak of absorbance values and is a measure of RNA integrity (Schroeder *et al.*, 2006). The analysis of these samples showed no evidence of the 28S peak being lower than the 18S rRNA and it was concluded that integrity was not compromised. The lowest RIN value measured was 9.1. Microarray quality control requirements have stated a minimum RIN of 7 (Kiewe *et al.*, 2009).

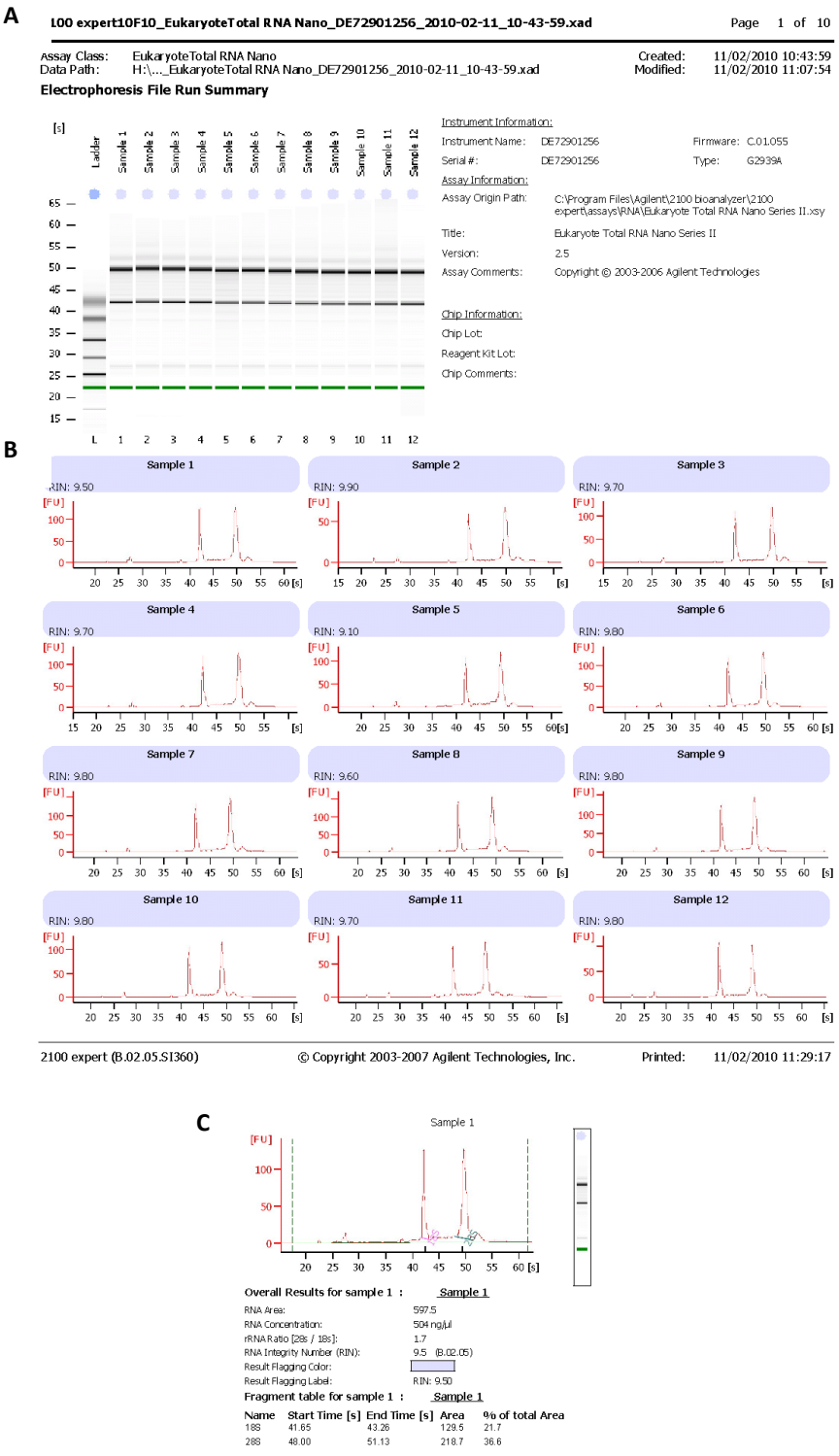


Figure 29 Assessment of RNA integrity using the Agilent Bioanalyzer. RIN values of RNA isolated on the 10 of February 2010 sent for miRNA microarray analysis and further transcriptomic analysis. A) Computer generated gel electrophoresis image; B). RIN is calculated based on the 18S and 28S ribosomal ratio with an Agilent designed algorithm; C) A 28S peak lower than the 18S peak indicates degradation of RNA.

4.1.2 *Assessment of cell response to Se supply*

RNA isolated from either Se-supplemented or Se-depleted Caco-2 cells was reverse transcribed and subjected to RT- PCR, as described in Chapter 2, to assess the level of expression of selenoprotein mRNAs known to be affected by Se depletion (Pagmantidis *et al.*, 2005). Amplification was carried out with primers specific for *GPX1* and *SELW*, selenoproteins low in the selenoprotein hierarchy and therefore expected to decrease dramatically under marginal levels of Se, and the housekeeping gene *GAPDH*. Separation of the PCR products by agarose gel electrophoresis showed single bands of 407, 352, and 128 bp respectively, sizes consistent with the expected products. As shown in Figure 30 and Figure 31, *SELW* and *GPX1* mRNA levels were lower in cells grown in Se-depleted medium whereas the reference gene *GAPDH* was unaffected.

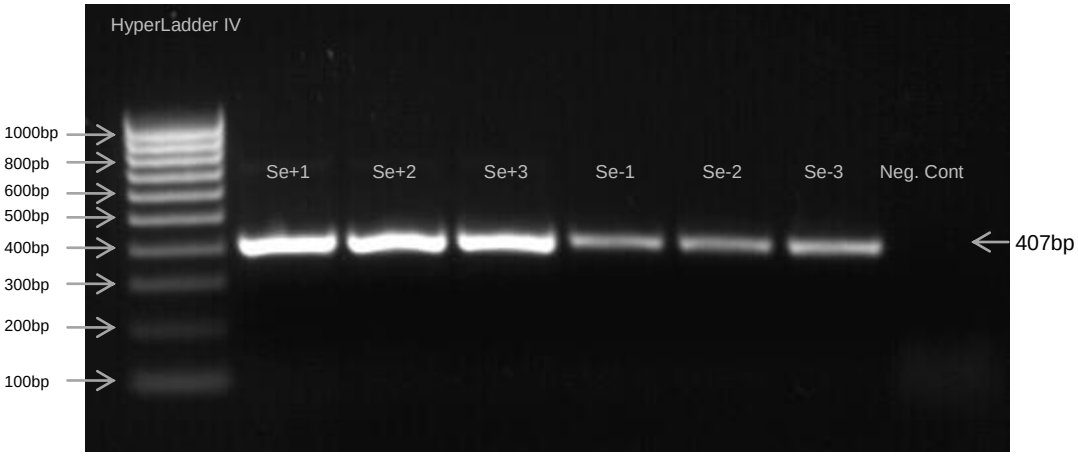


Figure 30 Semi-quantitative PCR for *GPX1* to assess Caco-2 cell response to Se supplementation/depletion.

Caco-2 cells were treated with/without 7ng of sodium selenite for 72 h after which total RNA was isolated, reverse transcribed and amplified over 28 cycles at an annealing temperature of 60˚C with specific primers to detect *GPX1*. DNA fragments were stained with ethidium bromide after electrophoresis through a 2% agarose gel at 70V for 45 min.

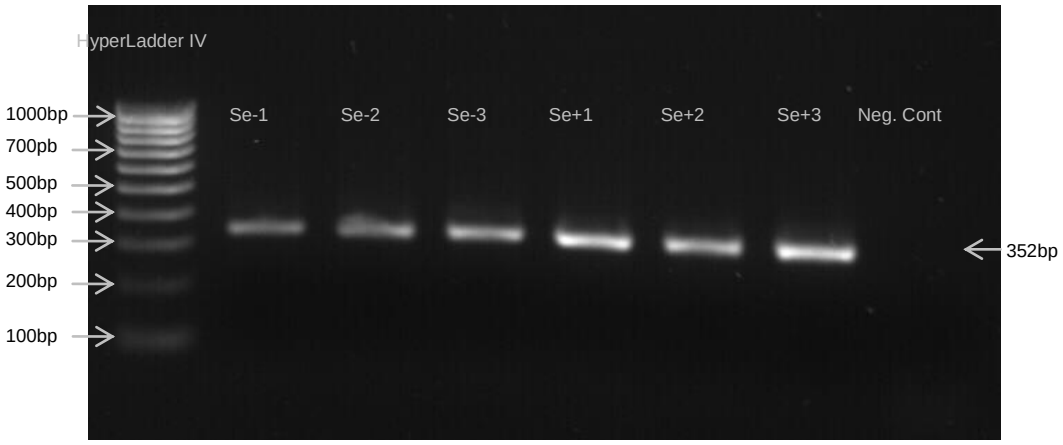


Figure 31 Semi-quantitative PCR for *SELW* to assess Caco-2 cell response to Se supplementation/depletion.

Caco-2 cells were treated with/without 7ng of sodium selenite for 72 h after which total RNA was isolated, reverse transcribed and amplified over 29 cycles at an annealing temperature of 60˚C with specific primers to detect *SELW*. DNA fragments were separated by electrophoresis in an ethidium bromide-stained 2% agarose gel at 70V for 45 min.

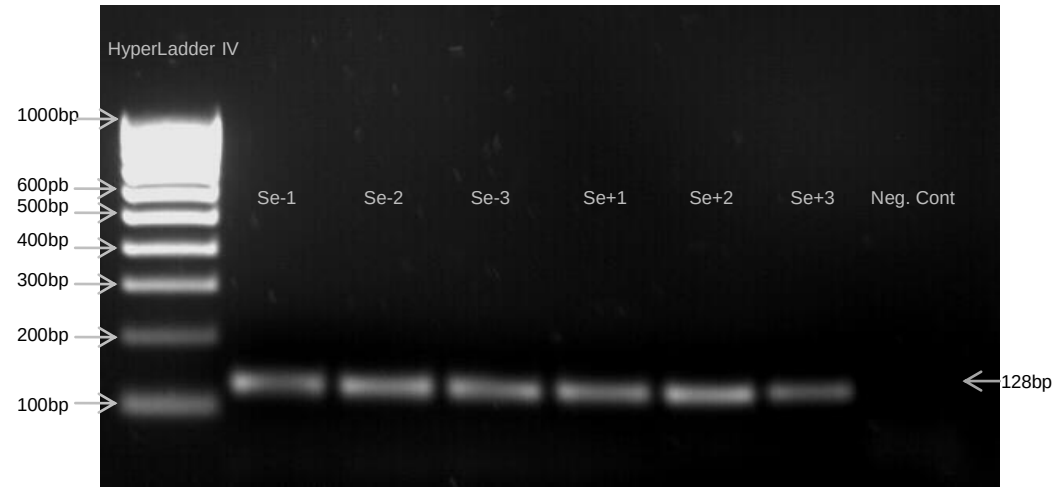


Figure 32 Semi-quantitative PCR for *GAPDH* in Caco-2 cells during Se supplementation/depletion.

Caco-2 cells were treated with/without 7ng of sodium selenite for 72 h after which, total RNA was isolated, reverse transcribed and amplified over 29 cycles at an annealing temperature of 55˚C with specific primers to detect *GAPDH* for normalization of *SELW* and *GPX1*. DNA fragments were stained with ethidium bromide after electrophoresis through a 2% agarose gel at 70V for 45 min.

SELW and *GPX1* expression was quantified by band densitometry, and normalized to that of *GAPDH* as a reference gene. As shown in Figure 33 there was a statistically significant (n=9, P<0.005) reduction of approximately 50% in both *SELW* and *GPX1* expression in cells grown in Se-depleted medium compared with cells grown in Se-supplemented medium.

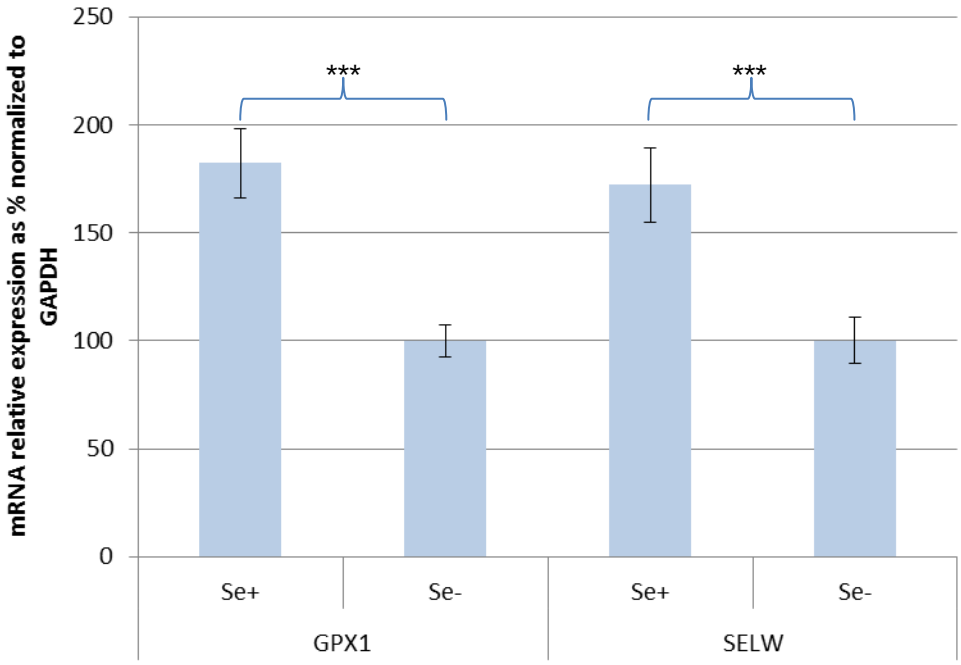


Figure 33 Assessment of Se status of Caco-2 cells at mRNA level by semi-quantitative PCR.
Gene specific bands generated by semi-quantitative PCR were quantified by densitometry of corresponding bands and data was normalized to *GAPDH*. Knockdown to 49% of basal *GPX1* level and 46% of basal *SELW* level.

4.1.3 *miRNA microarray design and analysis*

The human genome V2 Agilent platform is a one-colour microarray consisting of a short probe that hybridizes to the target miRNA and contains a hairpin loop that will join the synthesized sequence to the miRNA. In the present work, a platform of customized miRNAs was designed to investigate the global changes in miRNA expression under different Se concentrations. Human genome V2 Agilent miRNA microarray was hybridization to miRNA sequences from Sanger miRBase10.1, the

miRNA database with sequences and annotation, performed by Warwick University who also performed a second round of RNA quality control of the samples by measuring 260/280 ratios, determining RIN using a Eukaryote Total RNA Nanochip and by measuring the presence of small RNA species using a Small RNA detected on Bioanalyzer.

The array was customized using Agilent's eArray Design Tool so that predicted miRNAs could be confirmed. eArray, Agilent's microarray chip design tool, was used to introduce miRNAs of interest using the miRNA's name as identifier and an accession number was assigned to each miRNA. 20 features were selected per miRNA i.e. each miRNA was represented on the array by a total of 20 on the chip.

The design of the customised slide was 8x15K where 737 distinct *H. sapiens* miRNAs and 10 negative control miRNAs were included. The miRNA accession numbers are stated in Appendix A. Using bioinformatics analysis, which predicted miRNAs that could be interacting with the PBP, miRNAs from literature review, and miRNAs with probes already designed were included in the miRNA microarray design. The miRNAs included were the following and summarized on Figure 34:

- 194 miRNAs identified as present in colon tissue or colon cell lines (Bandres *et al.*, 2006; Cummins *et al.*, 2006; Han *et al.*, 2007; Ogier-Denis *et al.*, 2007; Monzo *et al.*, 2008; Peng *et al.*, 2009; Sarver *et al.*, 2009).

- A total of 292 miRNAs bioinformatically identified by microrna.org, Microcosm, GeneSet2MiRNA, and IPA (miRNAs predicted to target the genes within the PBP, as described in Chapter 3).

- 414 miRNAs taken from the Illumina human v2 chip platform miRNA array and PCR arrays which mature sequences were best characterized, and some even experimentally validated.

- 10 miRNAs from *Mus musculus* as an internal negative control, not expected to be expressed in Caco-2 cells and on the array to monitor false positives.

miRNA data files were loaded into GeneSpring GX 11 using the Agilent.miRNA.27924-06-Apr-2010-00-00 technology annotations. In accordance with the Agilent guided workflow expression values <1 were threshold to 1.0 and expression values were transformed to log base 2. Normalisation was performed using a 75th percentile shift and no baseline transformations were applied to the data. Control probes were removed from the dataset, prior to applying flag based filtering to remove probes where 'absent' flags were present across all the arrays. As a result, 139 miRNAs were identified as being expressed in Caco-2 cells and these are listed in Table 3; for each miRNA ID, the average of the four samples treated with Se sufficient media is represented by the column intensity Se+; and the average of the fold change of the 3 samples treated with depleted media is represented by the intensity Se-. Standard Error of the Mean (SEM) for intensity Se+ and intensity Se- are on the left of its corresponding column.

miRNA ID	Intensity Se+	SEM	Intensity Se-	SEM
hsa-let-7a	9.022078	0.048636	8.820162	0.153127
hsa-let-7b	5.570349	0.13509	5.241132	0.145011
hsa-let-7c	4.15079	0.069962	3.908818	0.156409
hsa-let-7d	5.438476	0.069183	5.043078	0.183156
hsa-let-7e	7.447854	0.061173	7.22813	0.08258
hsa-let-7f	8.443757	0.051569	8.183253	0.150259
hsa-let-7g	6.899875	0.047792	6.548454	0.146985
hsa-let-7i	4.435931	0.054651	2.917846	1.459711
hsa-miR-1	1.539187	0.889478	ND	0
hsa-miR-103	8.740459	0.056656	8.150362	0.284686
hsa-miR-106b	6.87748	0.06488	6.114391	0.142393
hsa-miR-107	8.947761	0.035605	8.316946	0.248256
hsa-miR-10a	6.717397	0.019686	5.922079	0.192433
hsa-miR-10b	5.256574	0.036311	3.961272	0.191949
hsa-miR-122	7.705111	0.0753	7.200214	0.256444
hsa-miR-1225-3p	0.761272	0.761272	ND	0
hsa-miR-1246	8.404661	0.045084	8.007711	0.100357
hsa-miR-125a-5p	4.9337	0.211012	4.583426	0.384583
hsa-miR-126	0.807424	0.807424	ND	0
hsa-miR-1260	6.241816	0.204559	5.623343	0.072961
hsa-miR-1268	5.168417	0.107281	4.466561	0.437194
hsa-miR-1274a	3.947706	0.104513	3.395757	0.106161
hsa-miR-1275	6.546358	0.067031	6.164011	0.105459
hsa-miR-128	4.845628	0.155514	3.772382	0.283495
hsa-miR-1290	4.961497	0.088476	4.45236	0.089803
hsa-miR-1308	7.489403	0.135294	6.495912	0.075958
hsa-miR-130a	7.271505	0.024923	6.463021	0.253988
hsa-miR-130b	7.686863	0.029542	7.052551	0.197461
hsa-miR-133b	4.153336	0.123501	1.211295	1.211295
hsa-miR-135a*	0.513313	0.513313	ND	0
hsa-miR-139-3p	1.596573	0.922267	ND	0
hsa-miR-140-3p	3.274347	0.080247	1.04089	1.04089
hsa-miR-141	2.656021	0.889822	1.06891	1.06891
hsa-miR-146b-5p	2.429291	0.8112	ND	0

miRNA ID	Intensity Se+	SEM	Intensity Se-	SEM
hsa-miR-148a	6.309481	0.088048	5.64309	0.116346
hsa-miR-148b	4.296767	0.082223	3.888899	0.208164
hsa-miR-149	4.214053	0.098441	2.479133	1.24343
hsa-miR-151-3p	4.82627	0.092285	4.15908	0.228345
hsa-miR-151-5p	7.646506	0.025719	7.187045	0.147258
hsa-miR-155	1.586798	0.916139	ND	0
hsa-miR-15a	6.497092	0.131338	5.896285	0.07258
hsa-miR-15b	9.523644	0.055495	9.227596	0.193577
hsa-miR-16	9.031421	0.047499	8.401617	0.211289
hsa-miR-16-2*	2.302734	0.017126	0.72783	0.72783
hsa-miR-17	8.7281	0.029747	7.973865	0.254897
hsa-miR-181a	3.662432	0.065618	1.228198	1.228198
hsa-miR-181b	0.766178	0.766178	ND	0
hsa-miR-181d	0.759912	0.759912	ND	0
hsa-miR-182	1.607439	0.538577	0.738159	0.738159
hsa-miR-1826	3.805342	0.119984	2.253012	1.127791
hsa-miR-183	5.463388	0.040492	5.293923	0.169285
hsa-miR-185	3.249075	0.047917	ND	0
hsa-miR-18a	5.46938	0.051774	4.166729	0.335517
hsa-miR-18b	3.892383	0.079531	2.945387	0.214296
hsa-miR-192	8.81948	0.073175	8.362597	0.165423
hsa-miR-192*	3.739343	0.072182	1.161362	1.161362
hsa-miR-193a-5p	3.267643	0.128876	1.085885	1.085885
hsa-miR-193b	5.722821	0.116303	5.086825	0.246936
hsa-miR-194	7.951715	0.073849	7.11759	0.21224
hsa-miR-196a	5.874406	0.039025	5.675287	0.210436
hsa-miR-197	1.487843	0.859075	ND	0
hsa-miR-19a	4.036184	0.224467	2.521994	1.261043
hsa-miR-19b	6.729723	0.116472	5.937374	0.209224
hsa-miR-200a	4.823103	0.125039	4.142203	0.169459
hsa-miR-200b	8.784505	0.027983	8.311356	0.150306
hsa-miR-200b*	2.433924	0.057725	0.770812	0.770812
hsa-miR-200c	9.509499	0.035174	9.105114	0.194941
hsa-miR-203	3.312888	0.148644	ND	0
hsa-miR-205	7.623943	0.070904	6.851522	0.349415

miRNA ID	Intensity Se+	SEM	Intensity Se-	SEM
hsa-miR-20a	8.515029	0.058233	7.855546	0.206871
hsa-miR-20b	5.779429	0.08682	4.955113	0.256972
hsa-miR-21	8.687183	0.17167	8.33308	0.146673
hsa-miR-210	5.429186	0.072436	4.313433	0.414999
hsa-miR-215	7.415388	0.103265	6.707484	0.124641
hsa-miR-22	7.041272	0.103703	5.99773	0.09331
hsa-miR-221	7.320775	0.047578	6.834014	0.207322
hsa-miR-224	6.21015	0.066881	5.718695	0.174656
hsa-miR-23a	8.76048	0.035599	8.254145	0.156926
hsa-miR-23b	8.887954	0.045638	8.427639	0.180123
hsa-miR-24	7.706882	0.015514	6.938614	0.208507
hsa-miR-25	8.730531	0.035503	8.267186	0.200301
hsa-miR-26a	6.553495	0.078728	5.604189	0.342955
hsa-miR-26b	6.903256	0.078364	6.392556	0.153007
hsa-miR-27a	5.136983	0.178805	4.290397	0.15054
hsa-miR-27b	7.367695	0.075607	6.752449	0.175431
hsa-miR-28-5p	5.326918	0.040254	4.192954	0.071334
hsa-miR-29a	6.434781	0.086205	5.962504	0.246721
hsa-miR-29c	2.315162	0.772385	ND	0
hsa-miR-30b	6.428092	0.10009	5.50396	0.153747
hsa-miR-30c	3.719069	0.0498	2.389714	1.197941
hsa-miR-30d	5.130328	0.035981	3.762747	0.499089
hsa-miR-31	4.573383	0.067303	3.784522	0.253689
hsa-miR-320a	7.162013	0.068845	6.686425	0.2042
hsa-miR-320b	8.736092	0.083738	8.282818	0.160009
hsa-miR-320d	9.186129	0.07521	8.800443	0.16449
hsa-miR-324-3p	4.25236	0.292101	3.462472	0.196601
hsa-miR-331-3p	6.223843	0.064975	5.621734	0.208552
hsa-miR-338-3p	5.707143	0.081372	5.171111	0.248368
hsa-miR-34a	5.688154	0.059739	5.300952	0.234408
hsa-miR-361-5p	6.328471	0.008713	5.725645	0.174457
hsa-miR-362-5p	5.122375	0.06553	4.560436	0.168691
hsa-miR-365	7.648694	0.033593	6.920944	0.199908
hsa-miR-371-3p	10.44318	0.013523	9.91175	0.156585
hsa-miR-371-5p	8.533873	0.071373	7.629844	0.210436

miRNA ID	Intensity Se+	SEM	Intensity Se-	SEM
hsa-miR-372	10.70369	0.064067	9.980973	0.186918
hsa-miR-373	9.382831	0.025974	8.733498	0.358062
hsa-miR-373*	4.371428	0.036219	3.900571	0.043516
hsa-miR-374b	2.313436	0.772538	ND	0
hsa-miR-375	3.706311	0.115586	2.25997	1.134564
hsa-miR-378	3.629943	0.043706	1.14892	1.14892
hsa-miR-378*	1.496086	0.863842	ND	0
hsa-miR-423-5p	5.341558	0.133256	4.69064	0.179493
hsa-miR-425	5.725467	0.031887	4.726552	0.390949
hsa-miR-429	4.201887	0.354806	ND	0
hsa-miR-455-3p	6.987688	0.058694	6.451077	0.176121
hsa-miR-483-3p	7.462019	0.105371	6.712069	0.209494
hsa-miR-483-5p	6.673096	0.064219	6.299194	0.167861
hsa-miR-494	8.721159	0.082309	8.044923	0.081757
hsa-miR-500*	4.20676	0.108002	3.495719	0.174715
hsa-miR-501-3p	0.514599	0.514599	ND	0
hsa-miR-501-5p	3.012151	0.09318	1.947929	0.978837
hsa-miR-502-3p	3.283707	0.129955	1.034	1.034
hsa-miR-517a	3.537956	0.045974	1.07354	1.07354
hsa-miR-517c	3.686805	0.05355	1.145054	1.145054
hsa-miR-518e*	3.698566	0.182194	0.844874	0.844874
hsa-miR-522	6.290417	0.038312	5.761151	0.190301
hsa-miR-532-5p	5.081926	0.04608	4.099361	0.188114
hsa-miR-572	3.61681	0.09918	2.441185	1.220714
hsa-miR-574-3p	1.647611	0.951357	ND	0
hsa-miR-574-5p	4.164115	0.296571	1.176549	1.176549
hsa-miR-623	0.499574	0.499574	ND	0
hsa-miR-625	3.371898	0.066472	ND	0
hsa-miR-629	0.55128	0.55128	ND	0
hsa-miR-638	6.01501	0.078613	5.688327	0.203988
hsa-miR-652	3.864843	0.059972	1.191884	1.191884
hsa-miR-660	4.083408	0.093736	2.250578	1.128048
hsa-miR-663	3.857931	0.10928	2.36189	1.180957
hsa-miR-671-5p	4.464573	0.071565	1.065811	1.065811
hsa-miR-7	4.440113	0.075812	4.108745	0.077011

miRNA ID	Intensity Se+	SEM	Intensity Se-	SEM
hsa-miR-720	10.40175	0.091499	9.843947	0.198153
hsa-miR-92a	9.791742	0.047591	9.312716	0.224071
hsa-miR-93	8.199486	0.070915	7.72554	0.299353
hsa-miR-940	5.483562	0.114589	5.330604	0.108044
hsa-miR-96	5.229107	0.140879	4.821802	0.116776
hsa-miR-99b	4.160071	0.081881	3.686793	0.217586

Table 3 miRNAs expressed in Caco-2 cells
RNA from Caco-2 cells was examined by hybridization to human genome V2 Agilent microarray and analysed by GeneSpring GX11. 145 positive signals were identified. intensity Se+ column shows the fold change of a particular miRNA under Se supplementation; while intensity Se- represents the intensity response of a miRNA to Se depleted conditions.

In Chapter 3, the use of bioinformatics tools IPA, MicroCosm, microRNA.org, and GeneSet2mRNA to predict miRNAs that targeted the genes in the PBP is described. The resulting miRNAs were included in the microarray chip design. The ratio of miRNAs predicted over the miRNAs expressed on the microarray was: IPA ratio 0.3, MicroCosm ratio 0.26, GeneSet2miRNA -ratio 0.25 and microrna.org -ratio 0.16.

194 miRNAs included on the chip design were extracted from literature as a group of miRNAs expressed in intestinal tissues, biopsies, and cell lines and I refer to them here as the intestinal microRNAome. From this intestinal microRNAome, 71 miRNAs signals were detected in Caco-2 cells; this number corresponds to about 37% of the intestinal miRNAome, relatively low number of miRNAs that would be expected to be expressed in this particular cell line; however, many isoforms of the miRNAs from the intestinal miRNAome were expressed. On the other hand, 74 miRNAs that were not part of that miRNAome were detected on our array. Some miRNAs detected under Se-supplementation conditions were not detected under Se deficient conditions. Perhaps rather surprisingly, the converse was not observed. This observation suggests that Caco-2 cells do not include any pathways for increased expression of miRNA that are activated by Se depletion but pathways are activated by the presence of Se or repressed by Se deficiency.

Differential miRNA expression analysis was performed using t-tests in GeneSpring (Benjamini-Hochberg False Discovery Rate (FDR) multiple testing correction applied) where probes had a corrected p-value of <0.05. Thirteen miRNAs passed the flag filter and were considered as differentially expressed, showing an upregulation in the presence of Se. The greatest changes in expression were shown by miR-185, miR-625, miR-203, and miR-429, where no signal was detected with Se-depleted RNA samples, and a FC 3.85, 3.37, 3.31, and 4.20 respectively for Se supplemented RNA samples. In addition to these dramatic changes, the difference in expression for the remainder of Se sensitive miRNAs was ~20% (Table 4).

miRNA ID	Intensity Se-	Intensity Se+	p-value
hsa-miR-185	ND	3.2490754	3.06E-08
hsa-miR-625	ND	3.3718982	1.30E-07
hsa-miR-203	ND	3.3128884	7.77E-06
hsa-miR-28-5p	4.1929545	5.326918	2.50E-05
hsa-miR-429	ND	4.201887	1.70E-04
hsa-miR-373*	3.9005706	4.3714275	3.98E-04
hsa-miR-10b	3.9612722	5.2565746	5.58E-04
hsa-miR-22	5.9977303	7.041272	8.18E-04
hsa-miR-532-5p	4.099361	5.0819263	0.00199156
hsa-miR-1308	6.495912	7.489403	0.00220046
hsa-miR-494	8.044923	8.721159	0.00235208
hsa-miR-106b	6.1143913	6.8774805	0.00298004
hsa-miR-30b	5.50396	6.428092	0.00323074

Table 4 Microarray expression of differentially expressed miRNAs.
Expression of miRNAs from RNA from Caco-2 cells after 72h of either Se supplementation or depletion analysed by miRNA microarray using the Human genome V2 Agilent platform. Data were analysed by student t-test and Benjamini-Hochberg multiple testing correction.

In summary, Figure 34 shows that from the 737 miRNAs obtained from the literature review, bioinformatic predictions or software search, and other expression platforms, 139 miRNAs were expressed in Caco-2 cells (Table 3), from which 13 were differentially expressed (Table 4).

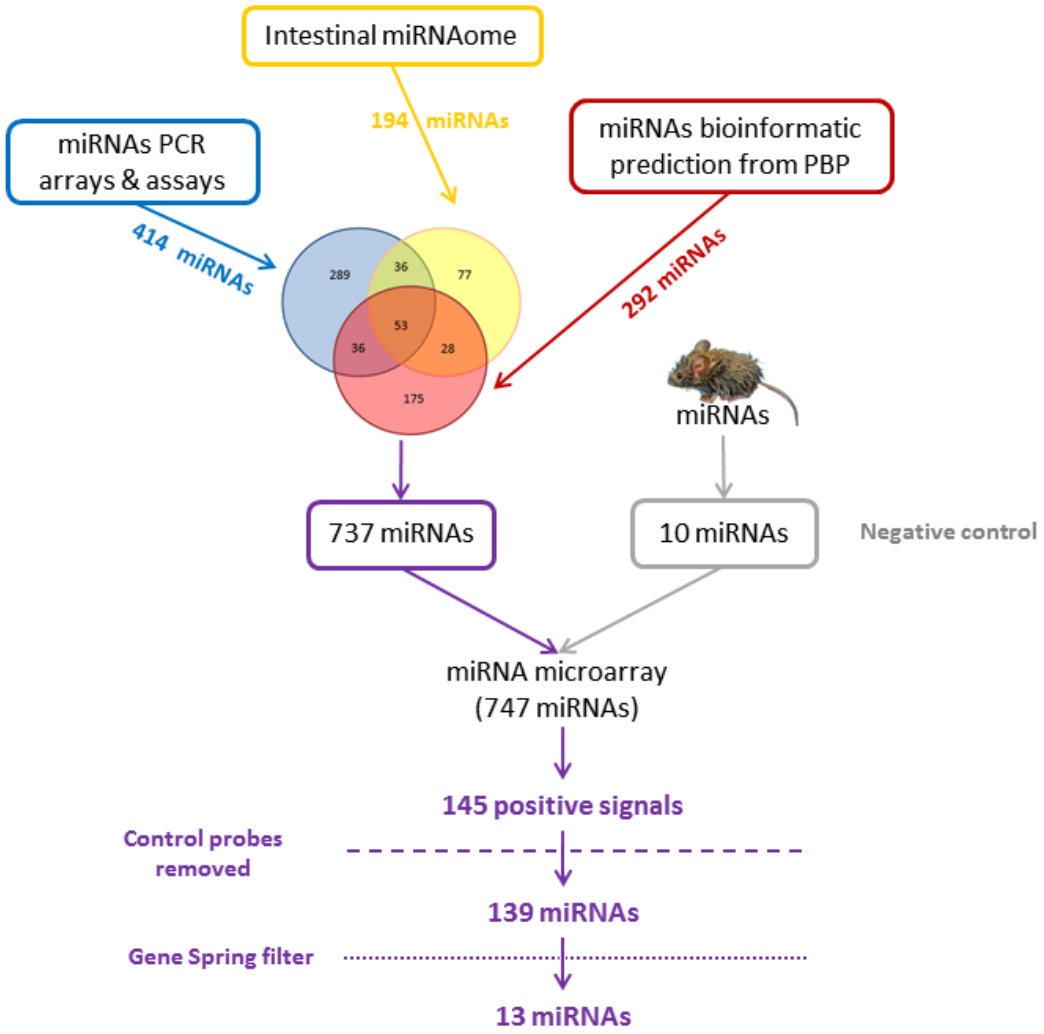


Figure 34 miRNA microarray chip design summary and miRNA expression. 737 miRNAs (purple box) from predicted from 3 bioinformatic tools and a software using the PBP as target genes (red box); miRNAs expressed in intestinal tissues and cell lines(yellow box); and miRNAs from miRNA PCR assays and microarrays (blue box) were used to investigate the effects of Se on miRNA expression. 10 *M. musculus* miRNAs (grey box) were also included on the miRNA chip design as internal negative control. The 145 positive signals were detected on the miRNA microarray were control probes and those described on the yellow, blue and red boxes of the same diagram, none of the *M. musculus* miRNAs were expressed. Table 3 shows the 139 out of 737 miRNAs that were expressed on Caco-2 cells. Finally, these 139 miRNA were flag filtered by Gene Spring detecting 13 differentially expressed miRNAs presented in Table 4.

During the course of this study, hsa-miR-1308 was erased from the Sanger database because it had been identified as a tRNA fragment rather than a miRNA, and Agilent discovered a missing G between the

hybridization sequence and the hairpin loop of the miRNA probes, which could have affected the intensity of the signal.

4.1.4 *Further experimental confirmation of miRNA microarray data.*

In order to validate the results from the microarray experiment, a second technique was used to confirm the miRNAs susceptible to Se supply/depletion performed with biological replicates. QRT-PCR is considered as the gold standard in gene expression quantitation (Git *et al.*, 2010), so was the approach used for validation.

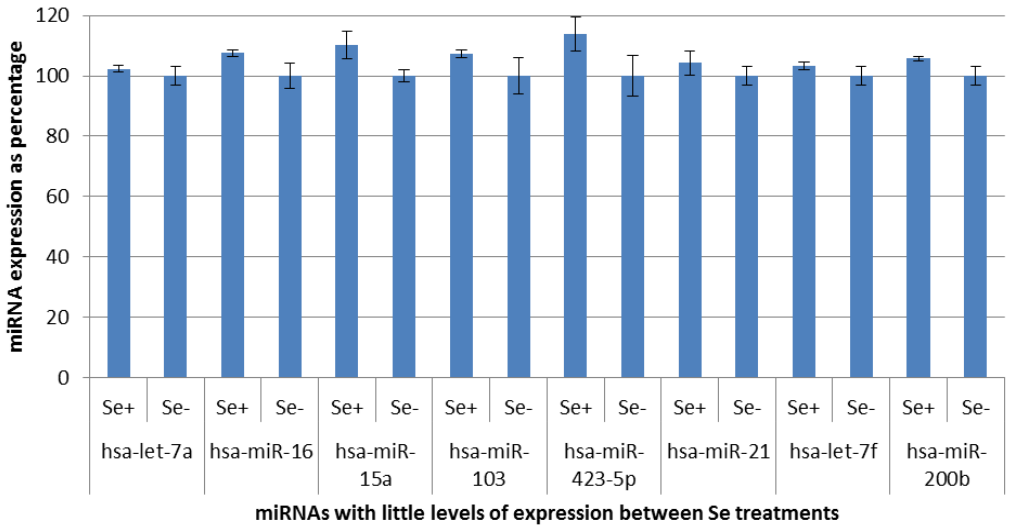


Figure 35 miRNAs with no significant change when comparing Caco-2 cells grown under Se depleted /supplemented conditions
Flag filter analysis (GeneSpring GX) was used for the identification of control miRNAs in RNA samples (n=4) isolated from Caco-2 cells grown for 72h under either Se-depleted conditions or with the addition of 7ng/ mL Se as sodium selenite and hybridized to a V2 Agilent 8x15k miRNA microarray.

In order to amplify miRNA by PCR, poly-A tailing, hairpin loop extension and LNA are some of the methods developed to reverse transcribe and amplify miRNAs. Among scientific work involving colonic epithelium tissues/cells/biopsies published in 2010, the Taqman PCR assay was the most frequent used method for miRNA QRT-PCR (Sato *et al.*, 2009) and was adopted for the current study. For the present analysis, miRNAs were prioritised according to changes in the microarray data such as the smallest p-value (<0.001), the biggest change in expression (Se supplemented/Se depleted), and scientific evidence of being present in epithelium or colon tissues. On the basis of low p-value, and large changes in expression between cells grown under Se sufficient or deficient medium, hsa-miR-185, hsa-miR-28-5p, hsa-miR-10b, and hsa-miR-106b were chosen for validation by q-miRNA-PCR. In order to normalise the data, two control miRNAs were also chosen for analysis: RNU6B and hsa-let-7a. Of a panel of miRNAs commonly used as reference (hsa-let-7a, hsa-miR-16, hsa-miR-15a, hsa-miR-103, hsa-miR-423-5p, hsa-miR-21, hsa-let-7f, and hsa-miR-200b), hsa-let-7a was chosen as it showed the smallest change in expression between the Se treatments (Figure 35) and it has also been reported as a reference/endogenous miRNA in colon tissues (Bandres *et al.*, 2006); finally, scientific evidence has suggested that the let-7 family is involved in miRNA processing (Jakymiw *et al.*, 2010) and has been reported as a stable miRNA in colorectal cancer (Chang *et al.*, 2010). RNU6B is a small nuclear RNA widely used to normalize q-miRNA-PCR data for gastric tissues, biopsies and cell lines (Tchernitsa *et al.*, 2010a; Wang *et al.*, 2010b) because it has fewer associations with physiopathological factors than the rest of the snoRNAs commonly used as housekeeping miRNAs and therefore the most stable (Gee *et al.*, 2011).

cDNA is synthesized in the Taqman assays by tailing the mature miRNA with a common sequence and using a universal primer as a template to reverse transcribe it. This universal primer is a hairpin loop designed with a single strand sequence on its 5’ end that will bind to the 3’ end of the miRNA followed by a double stranded segment with the

universal primer binding sequence contained in the loop as shown in Figure 36.

Dual-labelled hydrolysis probes that will produce a fluorescent signal corresponding to the number of amplified DNA copies are used for QRT-PCR detection.

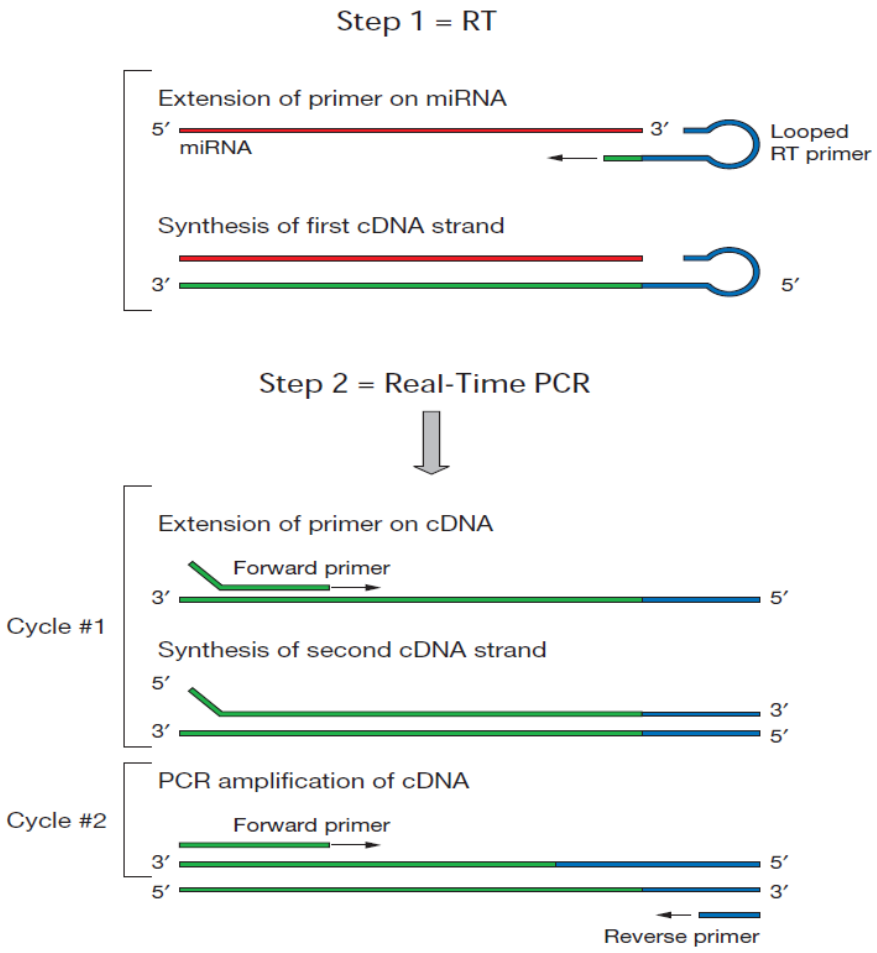


Figure 36 cDNA synthesis and QRT-PCR of miRNAs using the Taqman Assays.
Step 1 shows cDNA synthesis using a looped primer. The 3'end is complementary to the mature miRNAs sequence to extend the ~22nt sequence, where as the 5' end encodes a universal primer in step two, PRC products are amplified from cDNA synthesised from step 1 and the universal primer.

The first step of the miRNA reverse transcription is specific to each miRNA sequence to be amplified. This means that normalization will not be as accurate as using the same cDNA to amplify test and reference

genes, as for conventional QRT-PCR sample generated from a single RNA sample.

Master mixes contained: 0.075mM dDNTP mix with dTTP; 3.8 U of RNase inhibitor, 3µL of RT primers; and 5 µL of 10ng RNA concentration for a 15 µL final cDNA volume. To control such likely variation, the same sample was reverse transcribed with 8 different primers at the same time. A separate master mix, each containing the relevant miRNA primer was prepared for each of the 8 arrays.

Real-time amplification was carried out as described in Chapter 2, and Cycle threshold (Ct) values were extracted from the Applied Biosystems 7900HT Fast Real-Time PCR System using the FAM Non Fluorescent as the dye detection system. For data extraction, a 2.0 threshold and a baseline set 2-3 cycles before the amplification curve were used. Relative quantification was calculated by ΔΔ Ct formula using the Ct values of the miRNAs of interest minus the miRNAs of reference.

After normalization against both reference miRNAs (hsa-let-7a and RNU6B), the relative expression of all the studied miRNAs appeared increased in cells grown in Se supplemented medium compared with cells grown in Se-depleted medium. This finding is consistent with the microarray data. However, absolute differences in levels of expression were different when comparing Se treatments with the two different techniques. For example; miR-185, miR-429, and miR-625 were not detected in cells grown in Se depleted medium when the RNA was analysed by the microarray chip, but produced detectable signals after 28-30 cycles of QRT-PCR.

Figure 37 shows the relative expression of the miRNAs as measured by QRT-PCR when normalized to RNU6B. Compared to Se depleted conditions, RNA from cells grown in Se-supplemented medium showed a 67% increase in miR-625 expression, a 55% increase of miR-10b, 52% in miR-106b, 48% in miR-185, and 36% in miR-149.

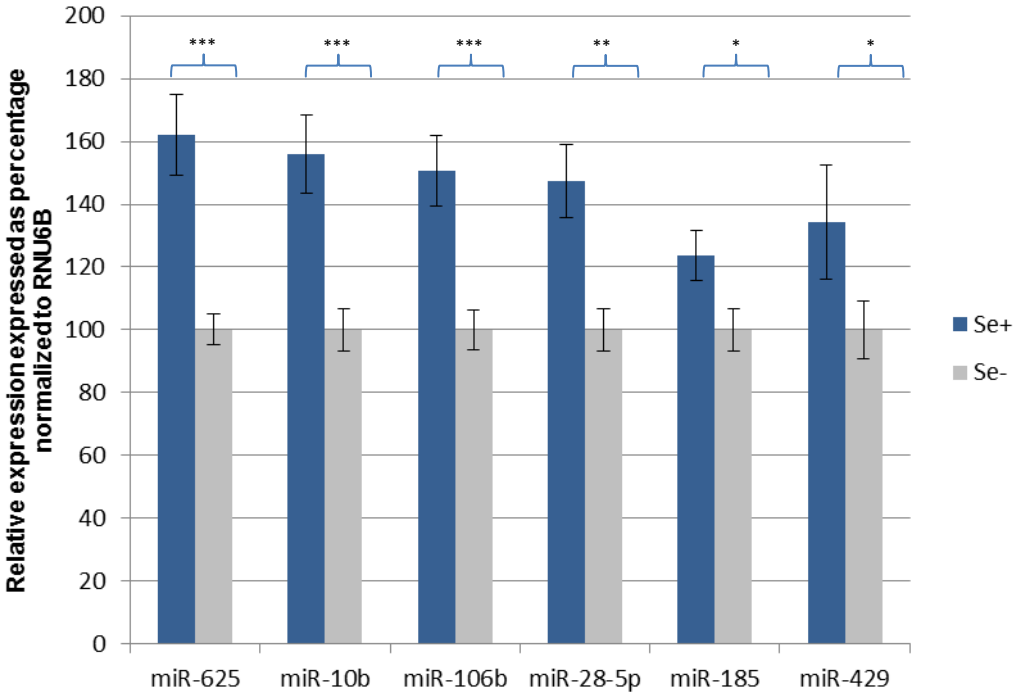


Figure 37 Relative expression of miRNAs measured by QRT-PCR and normalized to RNU6B for confirmation of selected Se sensitivity of miRNAs.
Samples sent for the analysis by miRNA microarray analysis plus two biological replicas were reverse transcribed using a specific stem loop primer and amplified by hydrolysis probes. Absolute expression was normalized to the snoRNA RNU6B and relative expression was expressed as a percentage. According to the microarray data, the expression of miR-625, miR10b, miR-106b, miR-28-5p, miR-185, and miR-429 were up-regulated in the presence of Se, which was confirmed by QRT-PCR.

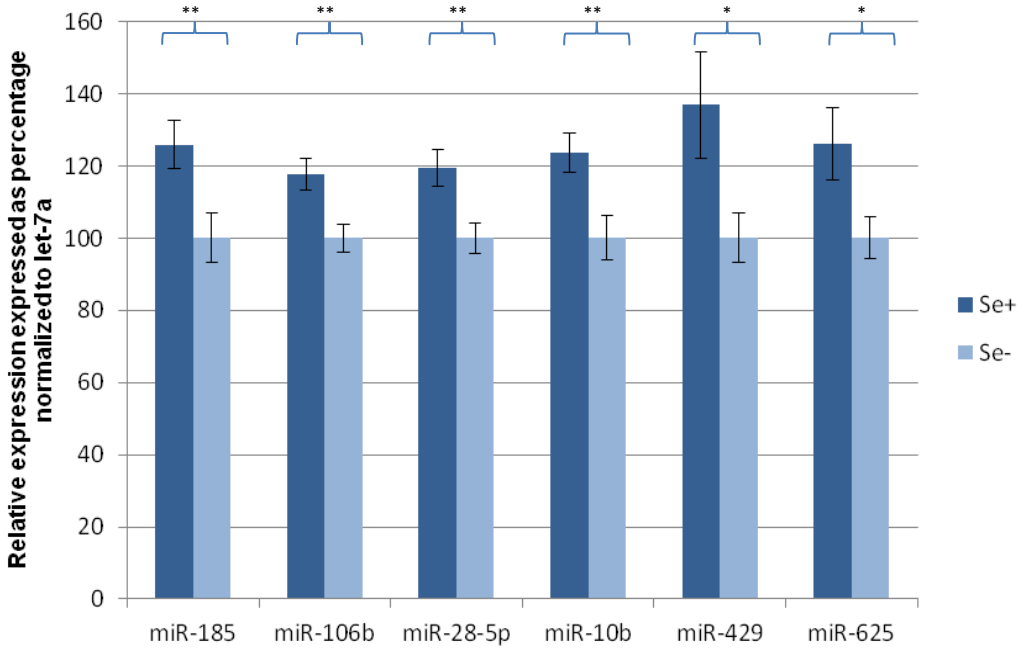


Figure 38 Relative expression of miRNAs measured by QRT-PCR and normalized to let-7a for confirmation of selected Se sensitivity of miRNAs detected.
Samples sent for the analysis by miRNA microarray analysis plus two biological replicas were reverse transcribed using a specific stem loop primer and amplified by hydrolysis probes. Absolute expression was normalized to let-7a and relative expression was expressed as a percentage. According to the microarray data, the expression of miR-625, miR10b, miR-106b, miR-28-5p, miR-185, and miR-429 were up-regulated in the presence of Se, which was confirmed by QRT-PCR.

Relative quantification was also calculated using let-7a as a second reference. As shown in Figure 38, normalization of miRNA expression relative to let-7a as the reference indicated an up-regulation of all the miRNAs I response to Se supplementation, in agreement with both the analysis of the expression data normalized to RNU6B and the microarray results. However, the level of upregulation in the presence of Se was slightly different from that after normalization with RNU6B, with a 28% increase in miR-185, 18% increase in miR-106b, 26% increase in miR-10b, 37% increase in miR-149, and 27% increase in miR-625.

4.1.5 *Bioinformatic analysis to predict effects of Se through miRNAs.*

After the expression of the Se-sensitive miRNAs was confirmed by QRT-PCR, a bioinformatics analysis was carried out to predict the targets of these 13 miRNAs.

As shown schematically in Figure 39, miRWalk was used to predict miRNA-mRNA interactions in 3 different ways: targeted predicted by 5 bioinformatic tools, predicting using experimentally validated miRNA-3'UTR interactions and predicting using interactions of the 13 miRNAs on 5' and 3'UTR sequences.

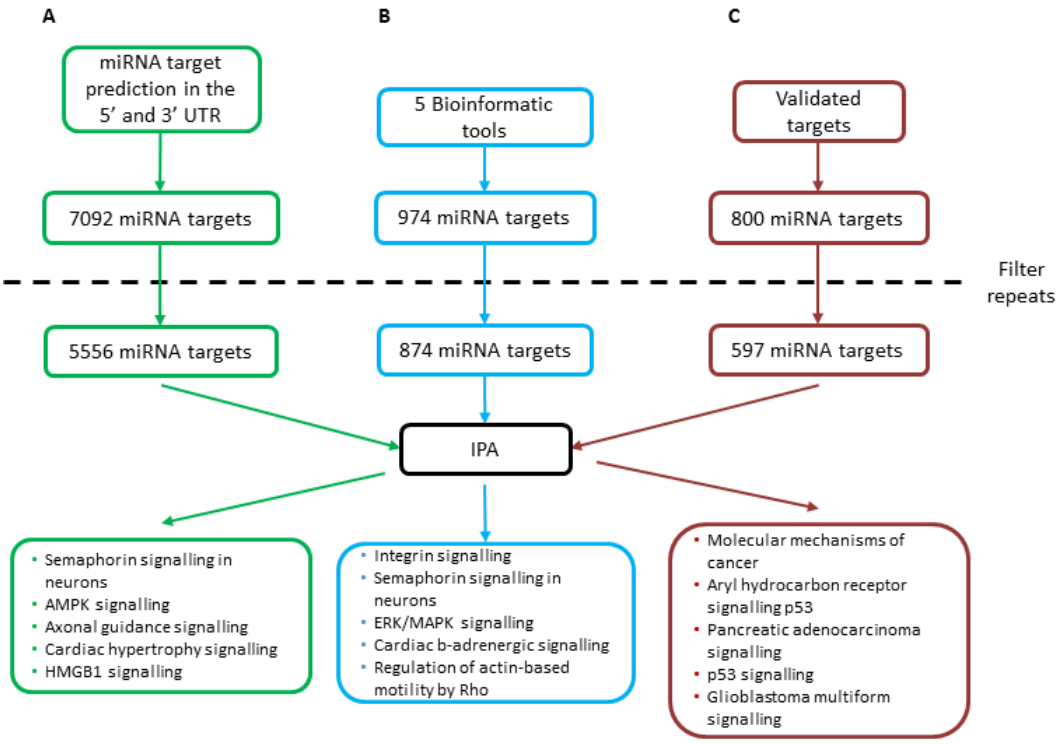


Figure 39 Se-sensitive miRNA target prediction using miRWalk.
miR-10b, miR-22, miR-28-5p, miR-30b, miR-106b, miR-185, miR-203, miR-373*, miR-429, miR-494, miR-532-5p, miR-625, and miR-1308 targets were predicted by miRWalk in three different ways: A) miRNA target prediction in selected areas (5' and 3'UTR) B) miRNA target prediction using 5 bioinformatic tools, and C) miRNA-mRNA experimentally validated interactions from the literature. Step A predicted 7092, B 874, and 597 from C, after repeated records were extracted from each list, each one of the 3 lists of miRNA targets were analysed independently and pathway analysis produced the data shown in Figure 41, Figure 42, and Figure 40 respectively.

Approximately 7,000 genes were found to be predicted targets of the 13 Se sensitive miRNAs identified in in this study, because of the promiscuity in miRNA-mRNA target pairing; typically a single miRNA could regulate 100 possible targets.

These genes were identified using a new algorithm for miRNA-gene interaction prediction developed in 2011, miRWalk. miRWalk, is similar to GeneSet2miRNA in the sense that a group of genes can be uploaded for miRNA prediction and 8 already established algorithms are used to identify possible targets. In the case of miRWalk, 4 of the 8 independent bioinformatic tools must predict the same miRNA-gene interaction in order for it to be considered a plausible result. miRWalk includes experimentally validated data from EntrezGene PubMed and predictions occur at the promoter region, ORF,3'and/or 5' UTR as long as the miRNA seed region binds to the sequence of interest (Dweep *et al.*, 2011); and it assigns a p-value according to the number of miRNA base pairs that would bind to an mRNA sequence.

The bioinformatic analysis with the 13 Se-sensitive miRNAs using miRWalk, described in Chapter 2 section 2.2.6, resulted in 8878 targets. These data were then analysed by IPA to predict physiological functions of this group of miRNAs and thus of Se activity through miRNAs. Three lists of genes were generated from miRWalk, using 1) miRWalk and the comparison of 5 bioinformatic tools; using 2) miRWalk validated targets and 3) miRNA target prediction on the 3'UTR.

For each analysis, a graph showing the most significant pathways of the data set and a table showing the numerical value of the significance of the association between the data set and the canonical pathway measured in 2 ways: 1) The p-value calculated using Fisher's exact test determining the probability that the association between the genes in the dataset and the canonical pathway is explained by chance alone. 2) A ratio of the number of molecules from the data set that map to the pathway divided by the total number of molecules that map to the canonical pathway is displayed.

The top 5 canonical pathways from each of the three lists are as below:

1. Predictions on the 5’ and 3’ UTR:

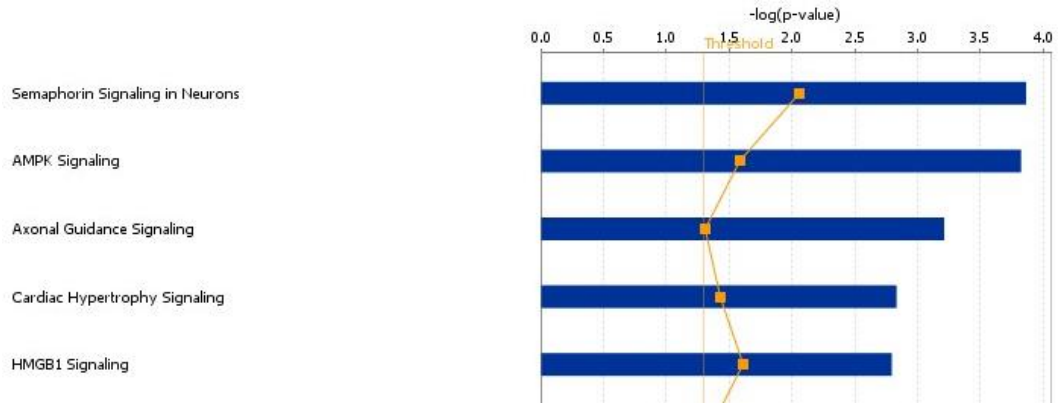


Figure 40 Pathways generated from of the 13 Se-sensitive miRNA targets
Semaphorin signalling in neurons, AMPK signalling, axonal guidance signalling, cardiac hypertrophy signalling, and HMGB1 signalling were the IPA generated networks using the predicted targets from the 13 Se-sensitive miRNAs interacting with the 5’ and 3’UTR.

Pathway	p-value	Ratio
Semaphorin signalling in neurons	1.37E-04	29/52 (0.558)
AMPK signalling	1.5E-04	59/137 (0.431)
Axonal guidance signalling	6.14E-04	140/395 (0.354)
Cardiac hypertrophy signalling	1.47E-03	88/227 (0.388)
HMGB1 signalling	1.61E-03	42/96 (0.438)

Table 5 p-value and ratios of the pathways generated from of the 13 Se-sensitive miRNA targets shown on Figure 40.

2. Comparison of 5 target prediction tools

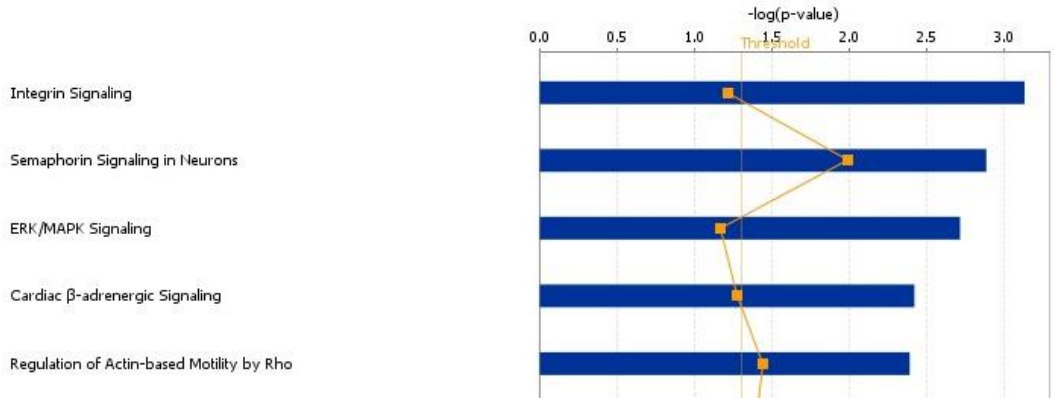


Figure 41 Comparison the pathways generated from the 13 Se-sensitive miRNAs targets predicted by 5 bioinformatic tools.
Integrin signalling, Semaphorin signalling in neurons, ERK/MAPK signalling, cardiac b-adrenergic signalling, and regulation of actin-based motility by Rho are the top 5 canonical pathways predicted by miRWalk and 4 other bioinformatic tools.

Pathway	p-value	Ratio
Integrin signalling	7.46E-04	21/198 (0.106)
Semaphorin signalling in neurons	1.32E-03	9/52 (0.173)
ERK/MAPK signalling	1.94E-03	19/188 (0.101)
Cardiac b-adrenergic signalling	3.82E-03	14/126 (0.111)
Regulation of actin-based motility by Rho	4.1E-03	11/88 (0.125)

Table 6 p-values and ratios of the pathways generated by the comparison the pathways generated from the 13 Se-sensitive miRNAs targets predicted by 5 bioinformatic tools shown on Figure 41.

3. miRWalk validated targets in the 3'UTR

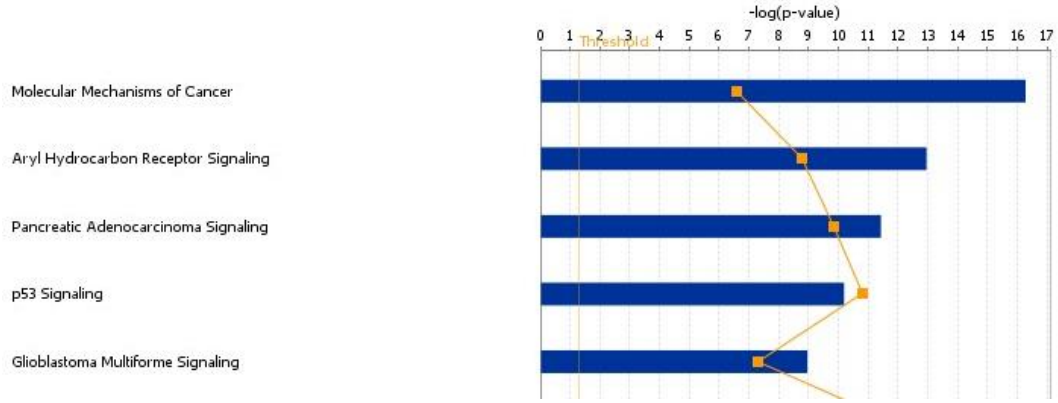


Figure 42 Pathways generated from the validated targets of the 13 Se-sensitive miRNAs.
Molecular mechanisms of cancer, aryl hydrocarbon receptor signalling p53, pancreatic adenocarcinoma signalling, p53 signalling, and glioblastoma multiform signalling were IPA generated networks from the 13 Se-sensitive miRNA validated targets interacting with the 3'UTR according to miRWalk.

Pathway	p-value	Ratio
Molecular mechanisms of cancer	5.31E-17	50/358 (0.14)
Aryl hydrocarbon receptor signalling	1.11E-13	28/150 (0.187)
Pancreatic adenocarcinoma signalling	3.78E-12	23/110 (0.209)
p53 signalling	6.5E-11	20/87 (0.23)
Glioblastoma multiform signalling	1.1E-09	24/155 (0.155)

Table 7 p-values and ratios of the pathways generated from the validated targets of the 13 Se-sensitive miRNAs shown on Figure 42.

Because these networks were the result of the pathway analysis of the predicted mRNA targets from the Se-sensitive miRNAs, it was expected that some pathways would be directly involved with Se biological roles. This was not observed, however an alternative interpretation of these pathways is addressed in the discussion of the present chapter.

4.2 Discussion

The first step in miRNA research is generally target prediction, and it is followed by experimental validation. High-throughput analysis of miRNA expression in the Caco-2 cell line and the effect on the expression of Se was tested using a miRNA transcriptomic approach. The microarray comprised probes and predicted target genes involved in the PBP, which was identified as regulated by Se in previous studies (Pagmantidis *et al.*, 2008), miRNAs previously identified as expressed in human intestinal cells or tissues and therefore with “tissue specificity”. Because these miRNAs have been experimentally validated elsewhere, most of them were expected to be expressed; however, only about 37% of them were detected in the Caco-2 cell line in this study. This is not an unusual phenomenon; authors of other studies have reported that mature miRNAs already validated experimentally in the same or a similar cell line have not been detected; for example the hsa-miR-145 (Bandres *et al.*, 2006) was not expressed neither in Banderas, Cubedo *et al*, nor in the present study. Also, only 101 miRNAs were found to be bioinformatically predicted out of the 292; and 90 out of the141 from PCR and microarrays platforms.

Mature miRNA and miRNA* sequences were included and analysed. Interestingly, of the two corresponding miRNA mature and * sequences, hsa-miR 375 and hsa-miR 375*, the latter was up-regulated in the presence of Se, whereas hsa-miR 375 was expressed but no Se-dependant changes in expression were detected suggesting than for this particular case, has-miR-375* responsiveness to Se must be mediated by the miRNA* rather than the mature mRNA species.

In terms of bioinformatic tools, from the detected miRNAs in Caco-2 cells, a higher number of miRNAs from IPA were expressed compared with the miRNAs predicted by GeneSte2miRNA, microRNA.org, or MicroCosm. This is not surprising since it is a tool with experimentally validated mRNA-miRNA interaction information. The association between the predicted miRNAs and the miRNAs expressed in Caco-2 was Microcosm > GeneSet2miRNA > microrna.org.

Appendix A shows, from the 145 expressed miRNAs, 71 correspond to those obtained through literature review, and 69 of them were the result of bioinformatic prediction tools.

One of the aims of this chapter was to investigate how Se supplementation and depletion affect the miRNA expression, the model for the present experiment was taken from (Pagmantidis *et al.*, 2005) who aimed to investigate the effects of Se on selenoprotein expression. In this model, Caco-2 cells are grown in medium without serum with or without Se. the level of *GPX1* expression in the Se deficient cells is very low and comparable to that found in cell grown in normal medium (Gautrey *et al.*, 2011) this reflects the low Se content of normal media (Leist *et al.*, 1996) and the sensitivity of *GPX1* expression Se supply.

In future work a more complete picture of the miRNA profile might be obtained by also analysing miRNAs in cells grown in normal medium. However, using the Se deficient and supplemented media, I was able to identify miRNA potentially sensitive to Se supply.

Interestingly, the Se sensitive miRNAs identified by the array were all up-regulated, and in some instances “on-off” effect between Se supplemented and depleted cells was indicated by QRT-PCR. Widely used miRNAs as reference species were examined within the data obtained to identify which were the most stable miRNAs for the for the purposes of use in normalizing QRT-PCR data to validate the microarray data, from which has-let-7a emerged as a candidate. Literature review identified the 44nt long snoRNA, RNU6B, as a suitable external second reference miRNA for data normalization. QRT-PCR data generated from the original samples analysed by miRNA microarray hybridization plus 2 biological replicas generated were normalized to these two reference miRNAs. Results of these analyses were very similar for both reference miRNAs, indicating that they were suitable controls for normalization. According to a study in which small nucleolar RNAs –but not RNU6B- and several miRNAs including let-7a and miR-16, two reference miRNA candidates for this study, miR-16 were explored as potential reference

miRNA candidates, let-7a was chosen over miR-16 according to the study on the role of miRNAs in colorectal cancer and miR-16 is a tumour suppressor (Chang *et al.*, 2010). The miRNAs that were not detected by microarray hybridization of sample obtained from cells grown under Se-depleted conditions were detected by QRT-PCR probably reflecting greater sensitivity of QRT-PCR compared to microarray. Nevertheless, the direction of change in miRNA expression measured between samples from Se-supplemented to Se-depleted cells was consistent between the two techniques.

A study using QRT-PCR and miRNA chip found down regulation of miR-373* in a study on bone marrow from children with B-cell precursor acute lymphoblastic leukaemia (Ju *et al.*, 2009); on the other hand, miR-373* was overexpressed in retinoblastoma detected by microarray and northern blot (Zhao *et al.*, 2009). And finally reported as expressed in a QRT-PCR analysis where no significant difference of the expression of miR-373* was reported between melanoma cells and primary human melanocyte (Satzger *et al.*, 2010).

mir-429 is a member of the miR-200 family, together with miR-141, 200a, 200b, and 200c, it has been suggested to prevent the transition of the normal stratified squamous epithelium to metaplastic columnar epithelium target ZEB1 and ZEB2 in order to maintain the oesophagus phenotype (Smith *et al.*, 2011; Stratmann *et al.*, 2011). It was reported as down-regulated in colon cell lines and tissues (Stratmann *et al.*, 2011; Sun *et al.*, 2011b; Zhang *et al.*, 2011b). Its role was studied by overexpressing with a mimic miR-429 in colon cancer biopsy and cell viability was found to be decreased as well as cell attachment; both which were inversely proportional to the concentration of mimic miR, suggesting that miR429 could play a role in cancer proliferation and progression. Also, downregulation of c-myc by a reporter assay targeting c-myc 3’UTR, gene responsible for cell cycle progression, apoptosis and cellular transformation (Sun *et al.*, 2011b). Overexpression of miR-429 is suggested to be linked to cancer on cells expressing CD133 which is a well-known colorectal cancer biomarker (Zhang *et al.*, 2011b).

Microarray data from colorectal adenocarcinoma samples linked a high rate in mortality with miR-185 overexpression despite cancer stage, gender and age of colorectal cancer Swedish patients. A correlation between the downreguation of miR-133b while miR-185 is up-regulated seemed to facilitate colorectal cancer progression (Akcakaya *et al.*, 2011). miR-185's role in cell proliferation in the colon cell line SW1116 mimicked RhoA/Cdc42 downregulation which could be translated into cancer survival due to a regulation of cell cycle and facilitating malignant cell apoptosis, reducing cancer progression (Liu *et al.*, 2011b). Limited research on miRNA-185 in colorectal cancer suggested to be an oncogene and tumour suppressor, the expression of this miRNA on kidney, ovarian and breast cancer models associated miR-185 downregulation with inhibition of cancer proliferation regulated by the a reduction of expression of a sine oculis homeobox, Six1, that acts as cell cycle regulator (Imam *et al.*, 2010)

A study of colorectal cancer samples from 58 Irish patients found that the expression of miR-10b is implicated in cancer progression after a QRT-PCR analysis related the increasing expression level of miR-10b to the depth of invasion progressing from Tumour (T)1 T1 to T4, where the most significant change was from stage T2 to T3 (Chang *et al.*, 2011). *In silico* hypermethylated CpG sites were found upstream miR-10b which were then observed in a methylation array, linking this miRNA with epigenetic changes and reducing colony formation of cancer cells by targeting and downregulating MAPRE1 (Kim *et al.*, 2011b).

miR-106b is part of the miR-106b-25 oncogene cluster (Nishida *et al.*, 2012) and it was the most studied miRNA at the time that the miRNA microarray in this present work was performed. Experiments on gastrointestinal samples are limited, but they include the microarray data from a large number of gastric cancer tissues from Germany comparing healthy against cancer tissues; miR-106b was overexpressed 9/10 times in patients where the cancer had invaded the lymph-nodes (Tchernitsa *et al.*, 2010b); and there was a similar pattern in biopsies of colon cancer and

its surrounding areas (Wang *et al.*, 2010b). In neoplastic gastric cancer, the expression on human peripheral blood of miR-106b is up-regulated; and after resection of the cancerous tissue, miR-106b expression dropped (Tsujiura *et al.*, 2010). Transfection of miR-106b into colon cell lines inactivates the Cyclin-dependent kinases (CDK) inhibitors, responsible for cell cycle activation (Cicenas and Valius, 2011) facilitating cancer progression (Kim *et al.*, 2009b).

There is no previous evidence of the gastrointestinal expression of miR-30b; it has been reported as differentially expressed in invasive urothelial carcinoma of the bladder detected by microarray; and its upregulation is associated with low survival when miR-30b together with 2 other miRNAs are overexpressed (Wszolek *et al.*, 2011). Targets of miR-30b are CELSR3, GALNT1, GALNT7, SEMA3, and TWF (in melanoma cells), genes promoting immunosuppression and tumour progression and transmembrane modifications in favour of cancer progression (Gaziel-Sovran *et al.*, 2011); another oncogene role of miR-30b was found in medulloblastoma (Lu *et al.*, 2009). However, it inhibits epithelial to mesenchymal transition in hepatocytes decreasing cell apoptosis and invasion by targeting TGF-β1 (Zhang *et al.*, 2012).

miR-203 promotes tumour suppression by targeting and degrading Akt2 mRNA on the human colon adenocarcinoma cell line HT29 with a p53 mutation but not with the wild type (Li *et al.*, 2011). This is supported by a study on gastric or colorectal cancer as well as colonic cell lines in which the downregulation of this miRNA directly correlated with the tumour growth. (Chiang *et al.*, 2011).

miR-28-5p expression in colorectal samples was down-regulated on the tumorous samples after pairing 24 biopsies from normal and cancerous biopsies from Italian patients. Following the assessment of miR-28-5p expression, transfection of miR-28-5p in HCT116 and RKO human colorectal adenocarcinoma cell lines decreased cell proliferation due to a 2.2 (HCT116) and 1.8 (RKO) fold increase in the apoptosis promoter protein poly(adenosine diphosphate-ribose) polymerase 1.

Controversially, miR-28-5p reduced cell invasion while its isoform miR-28-3p had the opposite effect. This could be explained by the target genes of miR-28-5p being CCND1 and HOX83 with cell cycle and proliferation functions as opposed to the miR-28-3p target NM23-H1 having a metastatic role miR-28-3p target (Almeida *et al.*, 2012).

The abundantly expressed miR-22 was detected by northern blot in several tissues and recently its expression was assessed by QRT-PCR in HCT116, HCT116 p53 KO and HT29 colon cancer cell lines. HIF-1α was identified as a miR-22 target, its translation but not transcription being affected. The protein level of the hypoxia regulator HIF-1α changed inversely to miR22 expression (Yamakuchi *et al.*, 2010; Yamakuchi *et al.*, 2011). Cell viability was reduced in HT-29 after miR-22 vector transfection, with a ~30% increased in apoptotic nucleis and a 95% decrease in cell survival (Zhang *et al.*, 2011a).

miR-532-5p had no references when the thesis was being written.

The role of miR-184 in the gastrointestinal system has been suggested to be as a regulator of apoptosis and inhibitor of cancer proliferation under miR-494 overexpression according to studies on gastrointestinal stromal tumours (GIST) cells with a KIT mutation (Kim *et al.*, 2011c) -a transmembrane protein inhibits apoptosis when mutated (Casar-Borota *et al.*, 2012)- and a target of miR-494 verified by luciferase reporter assays.

miR-494 transfection resulted in a decreased expression of KIT protein expression and a slightly lower effect on mRNA expression, suppression of cell proliferation and a 25% increase in apoptosis assessed by flow cytometry (Kim *et al.*, 2011c).

In 2011, miR-625 overexpression was found in human oocytes (Xu *et al.*, 2011); more recently, the expression of miR-625 was identified in neoplastic and healthy pancreatic biopsies with overexpression in the

former detected by a PCR assay (Yu *et al.*, 2012) but with no further correlation to the progression of the disease or functional analysis.

From the initial bioinformatic analysis in which the targets from the 13 Se-sensitive miRNAs were predicted by miRWalk and the networks identified by IPA function of these group of miRNAs could be elucidated. Some of these pathways such as the semaphorins were related to neuronal function; these are a membrane-associated secreted proteins associated with axon guidance and nervous system development (Pasterkamp, 2012). Semaphoring signalling mechanisms have been suggested to play important roles in neural injuries and diseases, and proposed to be able to regenerate central nervous system failure (Pasterkamp, 2012). Another function of the semaphorin system is the guidance of neuronal molecules. They may have a similar function in the immune system, (Schmidt and Rathjen, 2010) which they have been identified as being expressed in T and B cells, increasing antibody production and acting on macrophages to produce cytokines (Mizui *et al.*, 2009). A possible function of miRNA regulation of selenoproteins in human intestine might be linked to enhancement of the immune system observed under Se supplementation.

Cardiac electrophysiology, adaptability and heart failure are determined by the cardiac β-adrenergic signalling, another pathway identified by IPA as being Se-sensitive. β-adrenergic receptors are also responsible for calcium transport and glycogen catabolism to generate strong cardiac contractions and provide glucose; when these systems fail, the cardiac muscle loses contractibility and congestive heart failure occurs (Saucerman and McCulloch 2006); Keshan disease is a well known cardiomyopathy observed under low Se intake; and more recently, an increased expression of calcium channels genes observed in Se deficiency has been reported to induce cardiomyocyte injury (Cui *et al.*, 2012).

The identification of ERK/MAPK signalling pathway as being Se-sensitive miRNA regulated suggests that miRNAs could also regulate

extracellular communication by activation of intracellular responses for an adequate biological response achieved due to extracellular signal-reductase kinase (ERK) and mitogen-activated kinase (MAPK) balance in plasma and intracellular compartments (McKay and Morrison, 2007).

Most pathways identified by IPA were related to cancer proliferation and invasiveness. This includes integrin signalling which has an identified role in cancer proliferation and invasiveness by cell cycle control. It is regulated by the expression of integrin adaptors the regulation of which is known to contribute to cancer development and metastasis (Cas proteins). In epithelial-rich tissues, tumour aggressiveness is modified by Crk proteins and, tumour progression by the IPP complex, expressed and tumour growth, cell-cell adhesion and cancer progression and invasiveness facilitated by the Cap family (p140CAP). Interestingly, high levels of p140CAP and NEDD9 lead to resistance to cancer therapy.(Cabodi *et al.*, 2010).

A crosstalk signalling model has been proposed (Gong *et al.*, 2010) between HMGB1 and cancer pathways since overexpression of HMGB1 protein is associated with cancer proliferation and activates pathways like MAPK and AKT (Gong *et al.*, 2010) which were identified by the present analysis. HMGB1 overexpression is associated with tumour growth, inflammation and apoptosis regulators (Jia *et al.*, 2013). HMGB1 in Caco-2 cells might be regulated by the G1-S phase transition cell cycle regulator CyclinE rather than by p53.

Aryl hydrocarbons activate metabolic and detoxification pathways, cellular proliferation and migration; and regulate immune and neuronal systems (Barouki *et al.*, 2012).

p53 is a crucial DNA damage response mediator (Reinhardt and Schumacher, 2012). Despite the fact that Caco-2 cells lack the p17 chromosome where the p53 gene is located (Yu *et al.*, 2010), DNA damage response pathways were predicted to be affected by the Se-sensitive miRNAs. p53 regulates more than 2,500 tumorigenesis and tumour invasion genes (Mao *et al.*, 2012). Loss of heterozygosity is the

main cause of TP53 impaired activity triggering cancer (Bellini *et al.*, 2012), a loss of p53 function has pro-apoptotic effects through mitochondria outer membrane permeabilization (Stegh, 2012), possibly by caspase 8 (Liu *et al.*, 2011a).

There is evidence for DNA damage response genes being targets of miRNAs; for example SFN2H and miR-99 family in MCF7 cells (Mueller *et al.*, 2012) and p53 dependent upregulation of miR-192, miR-194, and miR-215 in U2OS, SJSA, HT29, and A549 cells In addition, p53 promoting pri-miRNA to pre-miRNA processing (Wan *et al.*, 2011) has linked miRNAs and p53 signalling pathways.

The pancreatic signalling pathways, predicted to be miRNA-sensitive is also related to p53. This pathway affects the epithelial cell compartment and the stromal cell compartment (transforming growth factor-β), cell apoptosis triggered by inadequate DNA damage response (TP53) activation of cell cycle progression and apoptosis, high production of undifferentiated cells (McCleary-Wheeler *et al.*, 2012). In addition, IPA identified glioblastoma multiform as another miRNA-sensitive pathway. This pathways is expressed in aggressive brain tumours characterized by vascular proliferation, tumour chemoresistance and cell death and with a general genomic instability that leads to a common mutation of the p53 pathway, cancer cycle progression by inactivation of the pRB signalling pathway, cancer development by PTEN, proliferation, invasion and apoptosis by transcription target activation from plasma membrane to the nucleus via STAT protein complexes (Mao *et al.*, 2012).

Interestingly, another classification of the p53 and cardiac signalling pathways identified by IPA could also be as metabolic pathways; recently p53 has been suggested to regulate metabolic functions like oxidative stress and glucose and fatty acid metabolism by glycolysis, activation of β-oxidation or mitochondrial oxygen consumption (Zhang *et al.*, 2010). Similarly, the cardiac hypertrophy signalling is regulated in the event of heart contractibility failure where phenotypic cardiomyocytes exhibit metabolic changes like utilization of free fatty acids which

decreases ATP availability in cardiomyopathy; in later stages of heart disease, this shifts back to glucose to produce energy. Another metabolic related change observed in the pathways analysis is Akt signalling, which responds to cardiac failure and helps maintain the homeostasis through protein synthesis, inhibition of apoptosis and metabolism also occurs in colon. An increased apoptosis and fibrosis is observed when the heart cannot adapt and activation of the apoptotic pathways, caspase activities, β-adrenergic, mitochondria cytochrome C occurs (Harvey and Leinwand, 2011). This could then link the APMK signalling pathway normalising the energetic balance ATP generation and consumption, cell growth under low nutrient availability, as well as regulating glucose metabolism. It also triggers ULK-1 dependent mytophagy of defective mitochondria controlling mitochondrial homeostasis (Mihaylova and Shaw, 2011).

Overall, it is possible that the p53, cardiac hypertrophy signalling Akt signalling pathway that are predicted to be targets of Se-sensitive miRNAs may have metabolic regulatory functions in Caco-2 cells.

In summary, results from QRT-PCR thus validated the miRNA microarray data and so detection of the expression of miRNAs in Caco-2 cells and more specifically, measurement of changes in expression under difference Se availability was accomplished. Finally, a newly developed algorithm was used as a guide towards prediction of effects of the miRNAs found to be Se regulated. The large number of gene targets predicted made possible the identification of several pathways using IPA, some of which include cancer signalling pathways and biological functions representative of other organs other than in the intestine. In order to find closer biological functions between these Se-sensitive miRNAs in Caco-2 cells and since these findings of Se were identified based on studies in different experimental models or species, we reasoned that a role for Se regulated miRNAs may be identified more robustly by examining effects of Se on gene expression pathways in the same cell lines as used for the miRNA microarray analysis; thus the response of the Caco-2 cell transcriptome was determined by microarray analysis as described in Chapter 5.

Chapter 5 Transcriptomic array characterisation of selenium sensitive genes on Caco-2 cells

5.1 Introduction

Several microarray experiments to address gene expression changes after alterations in Se status have been done using samples from rats, humans and mice, from different tissues and with varying dose and time of exposure. Overall, the response of selenoproteins has been similar, but the response to downstream targets to Se supplementation has varied when organic or inorganic forms of Se were compared (Bosse *et al.*, 2009; Liao *et al.*, 2011). Because the available microarray data from Se supplementation do not match the conditions of the present study and in order to link the miRNA results with the mRNA, a transcriptomic experiment was performed with the same RNA as was used for the previous miRNA microarray experiment.

In order to analyse the data generated by microarrays or other large sets of data, different approaches have be taken, for example the use of statistics and hypergeometric probabilities algorithms when the data to be compared have not been generated under the same conditions (Falcon and Gentleman, 2007) or the use of software that integrates molecules into networks for example Ingenuity Pathway Analysis from Ingenuity systems.

In Chapter 4, 13 miRNAs were identified as altered by Se-supply using a miRNA microarray, and these changes were confirmed by QRT-PCR. This chapter describes a global transcriptomic array experiment was carried out to investigate genes which shared expression changes under the same conditions under which the miRNAs were found earlier to show altered expression in Chapter 4 so a “match” between these two data sets could be done.

The expression at mRNA level of downstream targets of Se, as well as the expression of selenoproteins, after Se supplementation or Se deprivation has been investigated previously in several tissues and organisms. This was of particular importance because miRNA data are limited especially in gut cells. These studies have shown similarities and differences among the data sets. Nevertheless, these data were not suitable for the present study because the conditions of each experiment are different, such as tissue or cell line, organisms or gender in which the experiments have been conducted, time and dose of exposure/deficiency and finally the use of either organic or inorganic form of Se (or a combination of both). And since the study started from a group of targets from an experiment carried out under different circumstances to the one the miRNA microarray described in Chapter 4. Therefore to achieve the aim of finding a relationship between the miRNAs identified as being sensitive to Se supply and expressed mRNAs in Caco-2 cells, the same RNA was sent for a total transcriptomic experiment.

Prior to the work described in this thesis, no published data were available describing whether miRNA expression changes in response to altered Se supply. However, during the course of this work, although nothing has been published on miRNA expression responses to Se in Caco-2 cells a few reports indicate possible relationships between Se or selenoproteins and expression of specific miRNA. The first study linking Se with miRNA expression was an investigation which showed that the expression of the miR34 and its isoforms in prostate cancer cells was sensitive to Se concentration in the culture medium (Sarveswaran *et al.*, 2010). Furthermore, due to previous evidence linking Se deficiency and bone disorders (Sun *et al.*, 2011a), an investigation in patients suffering from Kashin-Beck disease showed that Se partially protected against chondrocyte cell death caused by T-2 toxin and also found changes in expression of a group of miRNAs that have been suggested to regulate differentiation targeting genes that encode fibril-forming collagen (Li *et al.*, 2008; Chen *et al.*, 2011b). Another study of adipose tissue–derived stem cells (ATSCs) showed that silencing of miR-10b and miR-23b increased the mRNA expression of the selenoprotein SENP1, and in a

complementary way, after SENP1 knockdown, an upregulation of miR10b and miR323b was observed (Kim, Jung et al. 2011); the same study found that miR-10b and miR-23b knockdown reduced cell death (Kim *et al.*, 2011a). Finally, in human lung adenocarcinoma cells TRX1 expression was silenced by a specific miRNA resulting in cell toxicity, ROS production, and decreased in cell viability (Poerschke and Moos, 2011).

Overall, the above studies indicate that selenoprotein gene expression can be regulated by miRNA and that changes in Se supply can alter miR expression. Work described in Chapter 4 and this chapter of this thesis has identified a number of miRNA for which expression is altered in Caco-2 cells by Se supply in the culture medium. The aim of the experiments described in the present chapter was to investigate the possible functional significance of the observed changes. To achieve this, the samples from which the 13 Se-sensitive miRNAs were identified were sent hybridized to a HumanRef-8 Expression Illumina BeadChip to investigate the transcripts sensitive to Se supply. I also carried out a two-stage analysis. Initially, since data were also available describing those mRNA in Caco-2 sensitive to exactly the same changes in Se supply, a bioinformatic approach was used to link the two datasets by using the target prediction tools previously described (Chapter 2) to look for MRE in the Se-sensitive transcripts. Second, having identified miRNA that were predicted to have target sequences in Se-sensitive genes, effects of silencing one specific miRNA was investigated using anti-miR. Antisense and miRNA mimics are the used as means of understanding miRNA role in gene expression regulation, disease diagnosis and therapy (Gambari *et al.*, 2011).

The aims and objectives of this chapter are to identify the mRNAs changing upon 40.5nM sodium selenite supplementation on Caco-2 cells for 72 h, for which the same samples that triggered the upregulation of the 13 Se-sensitive miRNAs described in Chapter 4 were sent for a total transcriptomic experiment in order investigate if there is a link between the Se-sensitive miRNAs and the Se-sensitive mRNAs through a target

prediction between the two data sets, to find candidate miRNA(s) and /or mRNA(s) for a functional analysis.

A summary of the work carried out in this chapter is illustrated in Figure 43.

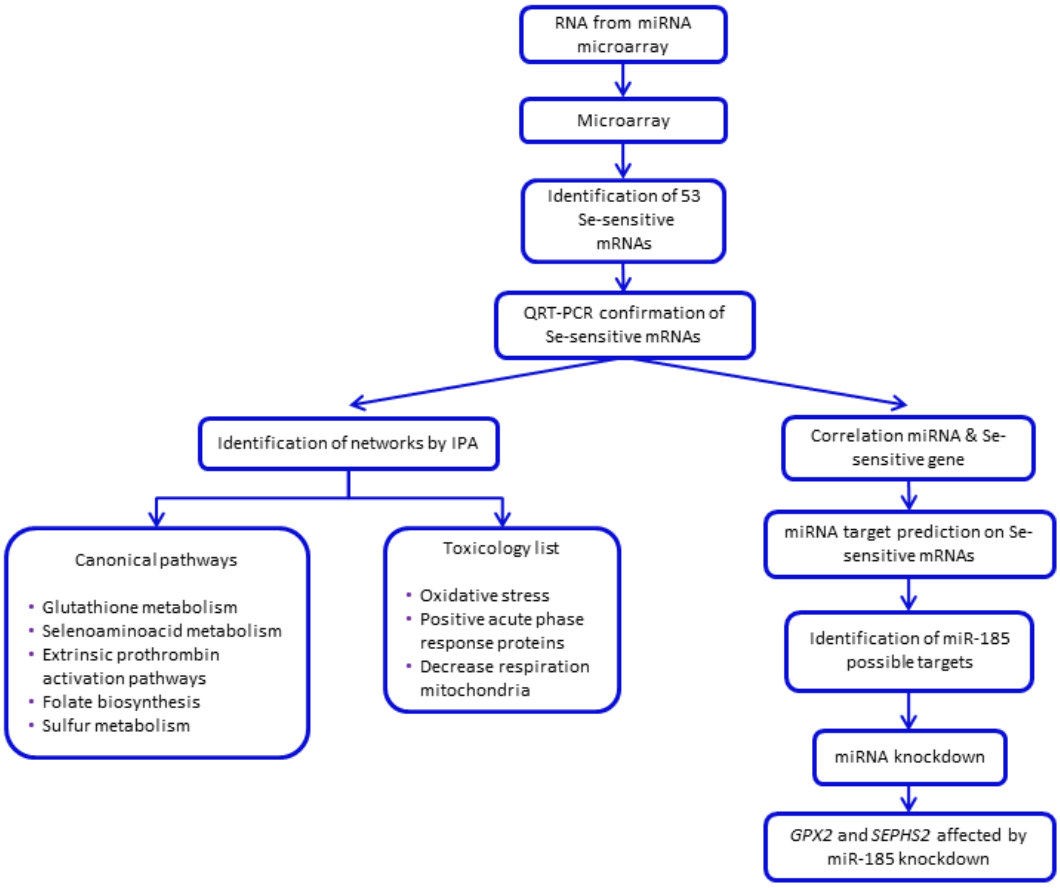


Figure 43 Chapter 5 workflow
RNA samples were sent for a total transcriptome experiment in order to identify changes in expression of mRNAs under the same conditions the Se-sensitive miRNAs were identified. 53 transcripts were differentially expressed after 72 h of 40.5 nM sodium selenite treatment. After similar levels of expression were obtained by QRT-PCR, a pathway analysis by IPA generated metabolic related pathways and mitochondrial respiration related toxicological lists. On the other hand, a miRNA target prediction analysis between the miRNA and the mRNA microarrays revealed matches between these two data sets; one of them was miR-185. After miR-185 knockdown, the levels of expression of two of its predicted targets (*GPX2* and *SEPHS2*) changed.

5.2 Results

5.2.1 *Transcriptomic analysis of gene experiment on Caco-2 cells*

As described in Chapter 4, RNA quality was assessed before samples were sent for total transcriptomic analysis. Only those samples of sufficient quality for hybridization and for which the Se supplementation had been proved to be effective by looking at the expression of *GPX1* and *SELW* were analysed. The RNA was subjected to Service XS' quality control assessment that is carried out as part of a regular procedure prior to sample hybridization to ensure reliable data is obtained.

Aliquots of the RNA samples previously analysed by miRNA microarray were sent to Service XS. Prior to gene microarray analysis, and according to Service XS standard operation procedures, quality control was determined by Nanodrop ND100 spectrophotometry and capillarity electrophoresis on the Agilent Bioanalyzer. Hybridization was carried out as previously described using a Human genome V2 Agilent chip in Chapter 2 section 2.2.16 microarray hybridization.

BeadChips were scanned using the Illumina iScan array scanner, where the fluorescent signal intensity from the single stranded fluor excited products of the beads was located and quantified from TIFF files corresponding to each assay in order to obtain a raw signal from a high resolution image which were then normalized with no background adjustment using Illuminas's Genomestudio v.2010.1 software. Internal controls (spiked-in controls and sample dependent controls) complied with specifications, gene signals were strong enough and all highly expressed genes were present and therefore signal generation was considered "passed" for analysis.

Data exported from Illumina Bead Studio in a GeneSpring compatible format was loaded into GeneSpring GX 11 for analysis and analysed using the Illumina.SingleColor.HumanRef-8_V3_0_R1_11282963_A technology annotation package. Expression values below 1 were thresholded to 1.0 on import and transformed to log base 2 prior to analysis. Quantile normalisation was applied across the samples with a baseline transformation to the median of all samples. To support flag based filtering, detection p-values <0.6 were marked as 'absent' and p-values >0.8 marked as 'present'. Detection p-values >0.6 but <0.8 were marked as 'marginal'.

Of 23206 entities, 13670 remained after filtering on flags so that 100% of the samples in any given group (treated or untreated) had 'present' or 'marginal' flags associated with each probe. Hierarchical clustering was used to assess the similarity of samples using a Euclidian distance metric with a centroid linkage rule. Samples from cells supplemented with selenium clustered as a separate group to those without selenium supplementation.

Differential gene expression analysis was performed using t-tests in GeneSpring (Benjamini-Hochberg FDR multiple testing correction applied) where probes had a corrected p-value of <0.05 and supplemented with RankProducts analysis using the RankProd package in BioConductor (probes with a pfp<0.05).

A cut-off of 1.0 was decided in order to obtain all the entities that changed under Se supplementation/ deprivation and because a bioinformatic pathway analysis was planned with gene lists; a higher number of genes for pathway analysis on IPA is preferred. This resulted in a quality controlled probe list of 53 significantly differentially expressed entities as shown in Table 8.

Symbol	Entrez Gene Name	Accession	Fold change	Regulation (Se+ vs Se-)
GPX1	Glutathione peroxidase 1	NM_201397.1	3.321	↑
SELW1	Selenoprotein W, 1	NM_003009.2	2.651	↑
MT1E	Metallothionein 1E	NM_175617.3	1.823	↑
MT1G	Metallothionein 1G	NM_005950.1	1.804	↑
MT1F	Metallothionein 1F	NM_005949.2	1.777	↑
GPX3	Glutathione peroxidase 3	NM_002084.3	1.716	↑
SELN1	Selenoprotein N, 1	NM_206926.1	1.58	↑
MT1X	Metallothionein 1X	NM_005952.2	1.527	↑
MT1A	Metallothionein 1A	NM_005946.2	1.522	↑
MT1H	Metallothionein 1H	NM_005951.2	1.509	↑
LOC728643		NR_003277.1	1.428	↑
SEPHS2	Selenophosphate synthetase 2	NM_012248.2	1.392	↓
CMIP	c-Maf-inducing protein	NM_030629.1	1.378	↑
GPX2	Glutathione peroxidase 2 (gastrointestinal)	NM_002083.2	1.375	↓
PIR	Pirin (iron-binding nuclear protein)	NM_001018109.1	1.37	↓
MT2A	Metallothionein 2A	NM_005953.2	1.357	↑
FABP7	Fatty acid binding protein 7, brain	NM_001446.3	1.324	↑
CGB5	Chorionic gonadotropin, beta polypeptide 5	NM_033043.1	1.323	↑
DCUN1D4	DCN1, defective in cullin neddylation 1, domain containing 4	NM_015115.1	1.306	↓
EPHX2	Epoxide hydrolase 2, cytoplasmic	NM_001979.4	1.305	↑
ME1	Malic enzyme 1, NADP(+)-dependent, cytosolic	NM_002395.3	1.301	↓
PIR	Pirin (iron-binding nuclear protein)	NM_001018109.1	1.298	↓
C7ORF59	Chromosome 7 open reading frame 59	NM_001008395.2	1.294	↑
CD63	CD63 molecule	NM_001040034.1	1.292	↑
HTATIP2	HIV-1 Tat interactive protein 2, 30kDa	NM_006410.3	1.286	↓
IFI6	Interferon, alpha-inducible protein 6	NM_022872.2	1.284	↑
ORM1	Orosomucoid 1	NM_000607.1	1.283	↑
GNG10	Guanine nucleotide binding protein (G protein), gamma 10	NM_001017998.2	1.282	↑
KLF5	Kruppel-like factor 5 (intestinal)	NM_001730.3	1.28	↓
GCNT3	Glucosaminyl (N-acetyl) transferase 3, mucin type	NM_004751.1	1.276	↓
HAND1	Heart and neural crest derivatives expressed 1	NM_004821.1	1.272	↑
C6ORF192	Chromosome 6 open reading frame 192	NM_002483.3	1.27	↓

Symbol	Entrez Gene Name	Accession	Fold change	Regulation (Se+ vs Se-)
CEACAM6	Carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen)	NM_052831.2	1.27	↓
SELK	Selenoprotein K	NM_021237.3	1.269	↑
GPX4	Glutathione peroxidase 4 (phospholipid hydroperoxidase)	NM_001039847.1	1.255	↑
GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	NM_003878.1	1.25	↓
GAB2	GRB2-associated binding protein 2	NM_080491.1	1.245	↑
LAMB1	Laminin, beta 1	NM_002291.1	1.245	↓
ATP6V1B1	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B1	NM_001692.3	1.242	↑
FGG	Fibrinogen gamma chain	NM_000509.4	1.239	↑
OIP5	Opa interacting protein 5	NM_007280.1	1.235	↓
IFITM3	Interferon induced transmembrane protein 3 (1-8U)	NM_021034.2	1.227	↑
CACYBP	Calcyclin binding protein	NM_014412.2	1.225	↑
SULT1A1	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1	NM_177530.1	1.22	↑
OSTC	Oligosaccharyltransferase complex subunit	NM_021227.2	1.217	↓
SLC1A3	Solute carrier family 1 (glial high affinity glutamate transporter), member 3	NM_004172.3	1.217	↓
IFITM2	Interferon induced transmembrane protein 2 (1-8D)	NM_006435.2	1.211	↑
LY6E	Lymphocyte antigen 6 complex, locus E	NM_002346.1	1.208	↑
ATP1B3		XM_001133534.1	1.205	↓
GPT2	Glutamic pyruvate transaminase (alanine aminotransferase) 2	NM_133443.1	1.204	↓
PAGE4	P antigen family, member 4 (prostate associated)	NM_007003.2	1.197	↑
GSDMB	Gasdermin B	NM_018530.2	1.179	↑
MAT2B	Methionine adenosyltransferase II, β	NM_013283.3	1.135	↑

Table 8 Differentially expressed genes in Caco-2 cells grown under Se supplemented/depleted conditions.
List of genes differentially expressed from RNA extracted from Caco-2 cells grown in Se supplemented/depleted media. Transcriptomic experiment was carried out using Illumina HumanRef-8 v.3 Expression BeadChip as platform. Expression values and normalization were obtained using GeneSpring with a cut-off of 1.0 to investigate change in expression. The mRNA levels of selenoproteins GPX1, SELW, GPX3, SEPN1, SELK, and GPX4 were up-regulated under Se supplementation conditions; whereas SEPHS2 and GPX2 were down-regulated under the same conditions.

Selenoproteins GPX1, SELW, GPX3, SEPN1, SELK, and GPX4 were up-regulated when cells were cultured under Se supplemented conditions, across species and tissues.

The most affected selenoproteins were GPX1 and SELW. This is consistent with the hierarchy for selenoprotein protection against Se deficiency (Crane *et al.*, 2009; McCann and Ames, 2011), these are low hierarchy selenoproteins and therefore more susceptible to changes in dietary Se (Bellinger *et al.*, 2009; Kasaikina *et al.*, 2011b).

5.2.2 *Bioinformatic analysis of network building on selenium sensitive genes*

The 53 Se sensitive genes were analysed by IPA both to identify the major altered pathways and to provide a graphic representation of gene-gene interactions and a representation of regulatory processes within the cell that could identify the roles of the up or down regulated genes changed upon Se supplementation. IPA aims to group a large number of molecules and compares the number of molecules within our experiment in relationship to a specific pathway to calculate the threshold and test the statistical significance. However, pathway analysis with the short list of 53 genes resulted in the identification of a small number of networks with a score above the threshold.

Canonical pathways identified as Se-sensitive were arachidonic acid metabolism, glutathione metabolism, and selenoamino acid metabolism. The P value is calculated according to Fisher's exact test to associate the number of genes that could be associated to that network by chance Figure 44; the generated networks are ranked according to their significance.

Toxicity list analysis in IPA identified oxidative stress, positive acute phase response proteins, and decrease respiration of mitochondria as being Se-sensitive pathways (Figure 45).

As described in Chapter 4, a graph showing the most significant canonical pathways (Figure 44) and toxicity lists (Figure 45) of the microarray and a table showing the p-value and ratios as a measurement of the significance of the association between the data set and the canonical pathways (Table 9) or toxicology lists (Table 10Table 10 p-value and ratio from toxicity list from IPA on the Se-sensitive genes from Figure 45.).

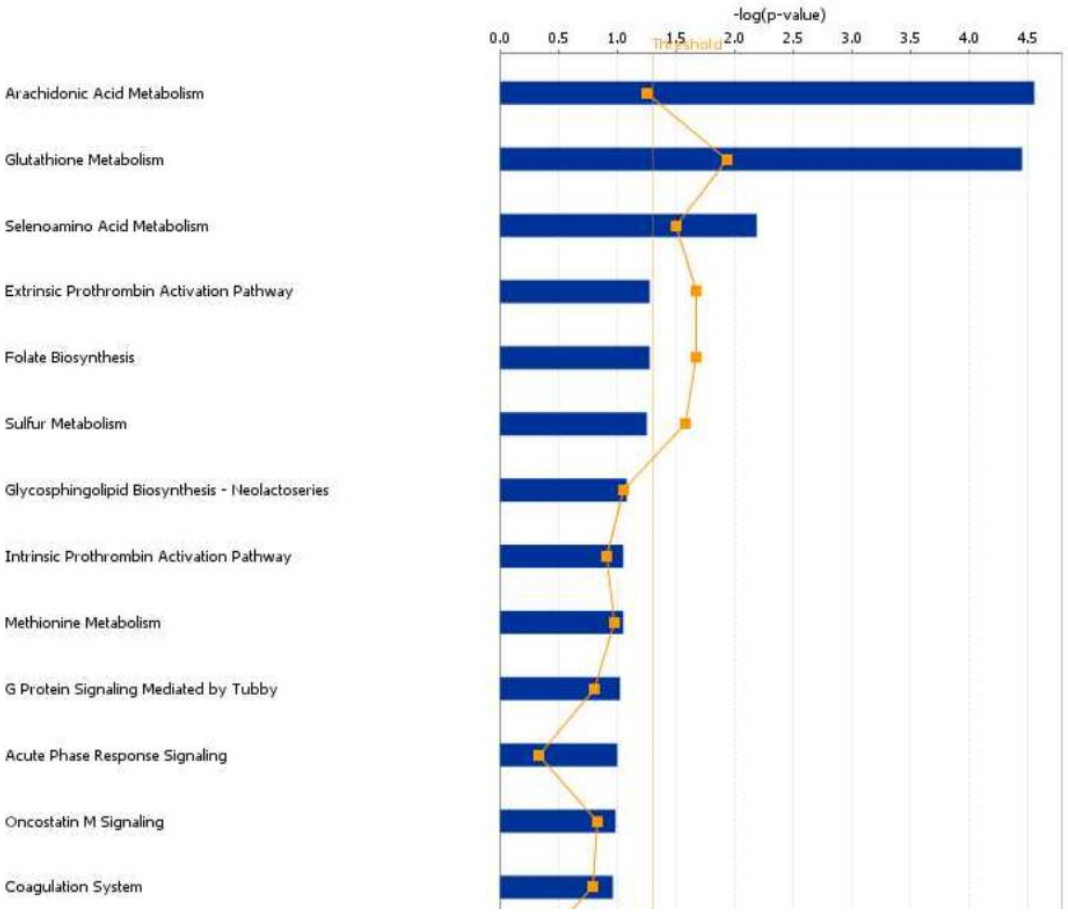


Figure 44 Canonical pathways from IPA on the Se-sensitive genes.

Dataset of canonical pathways from IPA using 53 Se sensitive genes identified by microarray from Caco-2 cells. This figure shows a histogram ordered according to the p-value which is calculated by dividing the number of genes from the differentially expressed genes from Caco-2 cells grown under Se depleted/supplemented conditions listed on Table 8 by the number of genes that forms each pathway. The ratio of Se-sensitive genes/genes from a pathway represented by the orange line that connects one rectangle with another. Ratios above the threshold line are significant pathways for the set of genes analysed in IPA.

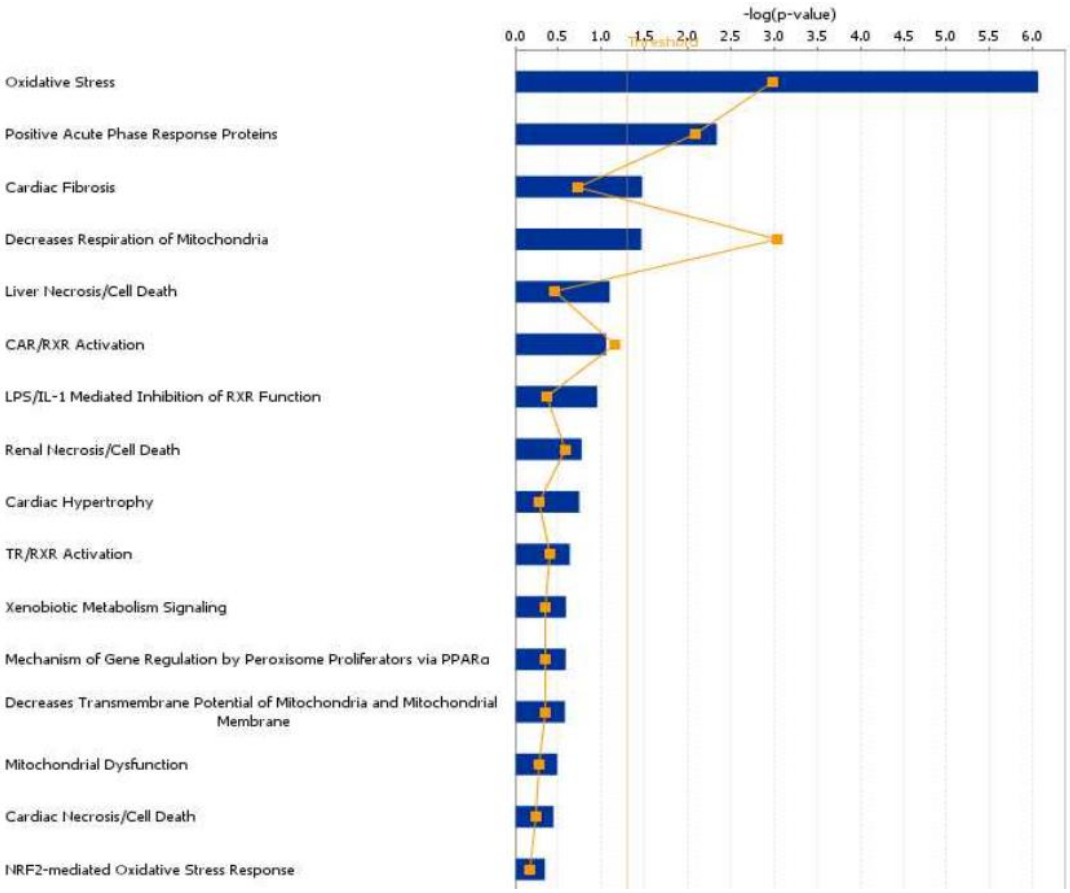


Figure 45 Toxicity list from IPA on the Se-sensitive genes.
Ingenuity Pathway Analysis showing the mechanisms of toxicity from the 53 Se sensitive genes after Se supplemented/ depleted treatment in Caco-2 cells. The toxicity list histogram, follows the same rules as the Canonical Pathways on Figure 44, the p-value is used to rank the order of the pathways and the orange line that connects the orange rectangles is calculated by dividing the number of genes from the differentially expressed genes from Caco-2 cells cultured under Se depleted/supplemented conditions listed on Table 8 by the number of genes that forms each pathway.

Pathway	p-value	Ratio
Arachidonic Acid Metabolism	3.6E-06	5/54 (0.093)
Glutathione Metabolism	1.59E-05	4/35 (0.114)
Selenoamino Acid Metabolism	7.28E-03	2/29 (0.069)
Sulfur Metabolism	4.37E-02	1/10 (0.1)
Folate Biosynthesis	4.37E-02	1/10 (0.1)

Table 9 p-values and ratios of the top 5 canonical pathways from IPA on the Se-sensitive genes from Figure 44.

Toxicity list	p-value	Ratio
Oxidative Stress	5.24E-07	5/37 (0.135)
Cardiac Fibrosis	2.82E-02	2/59 (0.034)
Decreases Respiration of Mitochondria	3.51E-02	1/8 (0.125)
CAR/RXR Activation	6.91E-02	1/16 (0.062)
Liver Necrosis/Cell Death	8.87E-02	2/112 (0.018)

Table 10 p-value and ratio from toxicity list from IPA on the Se-sensitive genes from Figure 45.

5.2.3 *Microarray validation by QRT-PCR*

For verification of the microarray data, 8 genes were chosen to be amplified by QRT-PCR. PCR primers used for this purpose were designed using the accession numbers of the genes of interest and NCBI sequences. NCBI BLAST software was used to ensure specificity.

RNA from 6 samples (3 Se depleted and 3 Se supplemented) in 3 biological replicates was isolated and reverse transcribed as previously described in chapter 2.

Because SYBR Green from Roche was used as the fluorescent dye, quality control of the amplification primers was carried out by of measuring melting temperature of each sample. For each set of primers, a melting temperature analysis was done and they all displayed one peak at the melting peak chart which shows that the fluorescent signal read by the LightCycler originated from only one amplified product.

A summary of the efficiency, error and slope for every gene amplified with SYBR green on the Light cycler is shown on Table 11. In Appendix B, the standard curve their slope, amplification curves as well as the melting temperature analysis as means of quality control assessment for QRT-PCR (Bustin *et al.*, 2009) analysis from LightCycler480 Software Version 1.5 are displayed for the genes used for microarray confirmation by QRT-PCR. As shown in Table 11 the genes of interest had an error of less than 0.1 which is less than the minimum required according to the LightCycler 480 Operator's manual for Software Version 1.5. In addition, the efficiency coefficient was close to 2, a good indicator that the amplification efficiency is close to 100%.

Gene of interest	Error	Efficiency
ACBT	0.0345	2.057
GAPDH	0.0418	1.972
GPX1	0.0192	1.953
GPX2	0.0305	2.002
GPX3	0.0416	1.968
SELK	0.0366	1.968
SEPHS2	0.0385	1.965
PIR	0.0992	1.763

Table 11 Summary of QRT-PCR amplification conditions for the Standard curves of each of the amplified genes by QRT-PCR using SYBR green as a fluorescent dye and Light Cycler as PCR system.
PCR efficiency was assessed by the LightCycler480 software using serial dilutions of 10 from a concentrated PCR product. QRT-PCR should be carried out conditions that give an efficiency close to 2 and a minimal error (<0.2). These criteria were achieved for all transcripts.

The expression of the six selected target genes normalized either to the expression of either ACBT or GAPDH is shown in Figure 46 and Figure 47. Using either reference gene, expression of GPX1, GPX3, and SELK are found to be increased in Se supplemented cells compared to cells from Se deficient medium, whereas expression of SEPHS2, and PIR were decreased. Relative expression of the targets showed similar change to those observed in the transcriptomic experiment although the difference in expression of SELK and PIR between the Se treatments was not statistically significant, but showed a similar trend to the microarray data.

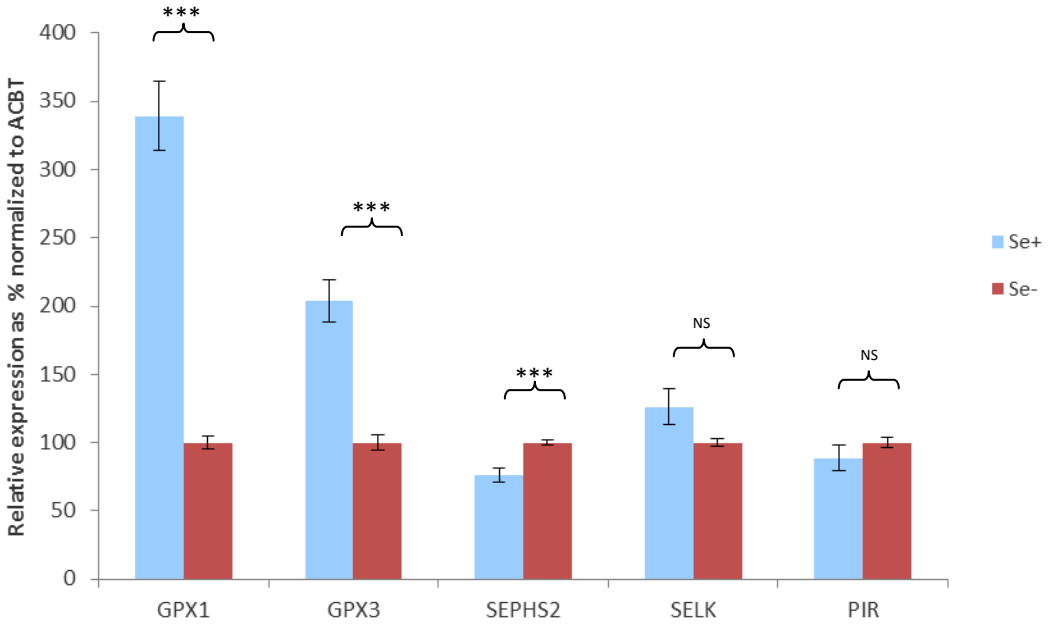


Figure 46 Expression of selected target genes by QRT-PCR using normalization to ACBT.

GPX1, GPX3, SEPHS2, SELK, and PIR primers were designed and specificity was assessed empirically by melting curve analysis in which a single melting peak was produced. Efficiency for each gene was close to 2 and error was <0.2. Cp values were extracted from the LightCycler and data was normalized to ACBT. GPX1, GPX3, and SELK were found to be up-regulated in cells grown under Se sufficient media, whereas SEPHS2 and PIR were down-regulated in cells cultured under the same conditions.

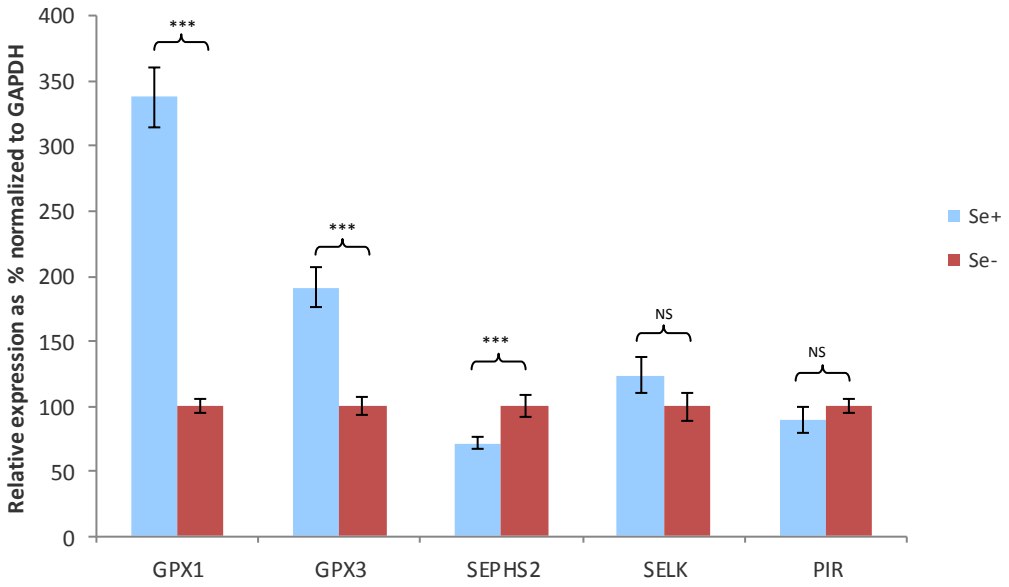


Figure 47 Assessment of relative expression of genes normalized to GAPDH for microarray confirmation expressed as percentage.
GPX1, GPX3, SEPHS2, SELK, and PIR primers were designed and specificity was assessed empirically by melting curve analysis in which a single melting peak was produced. Efficiency for each gene was close to 2 and error was <0.2. Cp values were extracted from the LightCycler and data was normalized to ACBT. GPX1, GPX3, and SELK were found to be up-regulated in cells grown under Se sufficient media, whereas SEPHS2 and PIR were down-regulated in cells cultured under the same conditions.

5.2.1 *Bioinformatic association of microarray results*

To assess potential MRE within the Se-sensitive genes identified in the transcriptomic experiment (Chapter 5), the target prediction tools previously described (Chapter 2) were used. Because miRNA-mRNA cleavage occurs not only within the 3' or 5'UTR but also in the ORF, no restrictions on the MRE localization were made when this was allowed by the database (e.g. microCosm and miRWalk). When using databases such as GenSet2miRNA the search of the mRNA sequence was within the 3'UTR. Thus, essentially the search for MRE in the Se sensitive genes

was carried out as described in Chapter 2 but without restriction of the regions of the nucleotide sequences. Briefly each HUGO ID (microRNA.org and microCosm) or group of HUGO IDs (miRWalk and GeneSet2miRNA) of the genes were loaded into the bioinformatics tools, in this case, the default setting were changed in order to allow the scan of the gene throughout the entire sequence; thermodynamics, seed region and orthology parameters were the same as for the rest of the bioinformatics predictions described in this thesis. However, during the course of this study, because some of the bioinformatics tools were upgraded (microRNA.org and GeneSet2miRNA) some parameters were modified. The miRNA target genes identified are listed in Table 12 as well as the bioinformatics tool which predicted them. As shown in Table 12 miR-185, miR-625, miR429, miR373*, miR22, miR-532-5p, miR106b, and miR-30b were predicted to recognise target sequences in mRNAs showing differential expression in Caco-2 cells following altered Se supply. Thus, of the 13 Se-sensitive miRNA identified by miRNA-microarray analysis (Chapter 4), 8 were predicted to have targets in Se-sensitive mRNAs.

miRNA	miRNA target	Bioinformatic tool
hsa-miR-185	GPX3	miRWalk microRNA.org microCosm
	GPX2	microRNA.org
	SEPHS2	microRNA.org
hsa-miR-625	ME1 CACYBP	miRWalk GeneSet2miRNA
	ATP6V1B1 CMIP	miRWalk
	GPX1	microCosm
hsa-miR-429	SELK	miRWalk microRNA.org
hsa-miR-373*	HTATIP2 GPT2	miRWalk
hsa-miR-22	GTP2	miRWalk
hsa-miR-532-5p	EPHX2	miRWalk
hsa-miR-106b	GPX2	microCosm
	MT1F	microRNA.org
hsa-miR-30b	GPX3	microCosm

Table 12 Bioinformatic correlation of Se sensitive miRNA-mRNA interaction. MRE from the Se sensitive genes identified in Chapter 4 were matched with the miRNAs from Chapter 4 by microRNA.org, miRWalk, microCosm and GeneSet2miRNA.

Prior to my work, only limited research had been published about any of these miRNAs. The most studied miRNA was miR-106b although there are few data available in gastrointestinal models and more specifically, colorectal models. This miRNA has been identified as an oncogene with the potential to develop metastasis (Tchernitsa *et al.*, 2010a; Wang *et al.*, 2010b; Nishida *et al.*, 2012). With regards to the rest of the Se-sensitive miRNA function in the colon, had been proposed to have a role in miR-10b with cancer progression (Chang *et al.*, 2011), but

the opposite effect was reported in human gastric cancer cells(Kim *et al.*, 2011b; Kim *et al.*, 2011c). miR-373* has been found to be expressed in bone marrow (Ju *et al.*, 2009), melanoma cells, primary human melanocyte (Satzger *et al.*, 2010) and retinoblastoma (Zhao *et al.*, 2009). So far, some Se-sensitive miRNAs have been only been found to be tumour suppressors on colorectal tissues or cell lines when they are up-regulated; this is the case of miR-429 (Sun *et al.*, 2011b) miR-203 (Chiang *et al.*, 2011; Li *et al.*, 2011), and miR-28-5p which has been reported to reduce invasion of tumour cells (Almeida *et al.*, 2012). miR-22 overexpression by transfection (Zhang *et al.*, 2011a) and miR-494 upregulation (Kim *et al.*, 2011c) have been categorized as oncogenes.

Neither miR-30b nor miR-625 have been studied in the gastrointestinal epithelium and there is no evidence for expression miR-532-5p at all. However, miR-30 expression has been reported as an oncogene in bladder carcinoma (Gaziel-Sovran *et al.*, 2011; Wszolek *et al.*, 2011) and medulloblastoma (Lu *et al.*, 2009), and controversially it suppresses of tumour progression in hepatocytes (Zhang *et al.*, 2012). miR-625 is overexpressed human oocytes (Xu *et al.*, 2011) and pancreatic biopsies (Yu *et al.*, 2012) but no further hypothesis of its function has been proposed.

miR-185 has been found to be expressed in colorectal adenocarcinoma samples (Akcakaya *et al.*, 2011). In addition mi185 has been suggested to have a pro-apoptotic effect on cancer cells, and a tumour suppressor effect (Liu *et al.*, 2011b), although oncogenic effects have been reported in kidney, ovarian and breast tumours (Imam *et al.*, 2010).

In the literature, only a few targets had previously been identified for the Se-sensitive miRNAs: miR-30b targets in melanoma cells are CELSR3, GALNT1, GALNT7, SEMA3, and TWF (Gaziel-Sovran *et al.*, 2011); HIF-1α was identified as a miR-22 target (Yamakuchi *et al.*, 2011); finally, Six1, a tumour progression regulator, was indicated to be a target for miR-185 in kidney, breast, and ovaries (Imam *et al.*, 2010). As shown

in Table 12, bioinformatic analysis with microRNA.org, MicroCosm, miRWalk and GeneSet2miRNA identified MRE for miR-185, miR-625, miR429, miR373*, miR22, miR-532-5p, miR106b and miR-30b within Se-sensitive mRNAs GPX1, GPX2, GPX3, SEPHS2, ME1, CACYBP, ATP6V1B1, CMIP, SELK, HTATIP2, GPT2, EPHX2 and MT1F. According to the network analysis from Chapter 5, GPX1, 2, 3, and SEPHS2 which is necessary for the synthesis of selenophosphate synthase, were part of the Glutathione metabolism pathway.

miR-185 was chosen as the focus for subsequent functional analysis using anti-miR silencing. This choice was based on its previously reported expression in colorectal tissue (Akcakaya *et al.*, 2011; Liu *et al.*, 2011b) and the prediction that it has MRE targets in the three selenoprotein mRNAs GPX2, GPX3 and SEPHS2. In addition to miRNA target prediction, the SECIS element was also predicted in the above mentioned selenoproteins sequences using SECISearch 2.19 (Kryukov *et al.*, 2003). Table 13 shows the MRE start sites of miR-185 and complementary information showing where SECIS element is predicted. For GPX3 and SEPHS2, two miR-185 recognition elements were predicted, whereas GPX2 had only one; one of these predictions was of particular interest since SEPHS2 MRE for miR-185 started at the nucleotide 570, which is within the predicted SECIS element (550-647). Since Se supplementation led to increased expression of miR-185, the aim of the miR silencing experiments was to investigate if knock-down of miR-185 under Se supplementation conditions led to an effect on its predicted targets similar to at observed during Se depletion.

miR-185 target	MRE start	SECIS	Bioinformatic tool
GPX2	145	861-955	microRNA.org
GPX3	314	1533-1625	microCosm microRNA.org
	1234		miRWalk
SEPHS2	570	550-647	microRNA.org
	2112		miRWalk

Table 13 miR-185 predicted MRE in the selenoproteins genes.
Targets for miR-185 were identified by microRNA.org, microcosm, and miRWalk bioinformatics tools. SECIS site was identified using SECISearch 2.19 on *GPX2*, *GPX3* and *SEPSH2* sequences, and then related to the MRE with miR-185 cleavage site.

5.2.1 *Expression of mRNAs identified as miR-185 targets in silico*

The bioinformatic analysis described in section 5.2.1 indicated that the transcripts for, GPX3 GPX2 and SEPHS2 had MRE for miR-185. Having determined conditions for miR-185 silencing, RNA from cells that had been treated with the anti-miR-for 48 was reverse transcribed and the level of expression of miR-185 targets assessed by QRT-QPCR. After 48h treatment with the ant-miR, the expression of the potential miR-185 targets GPX2 and SEPHS2 was significantly increased by approximately 60% (Figure 48) compared with negative control; in contrast there was no significant effect on GPX3 expression.

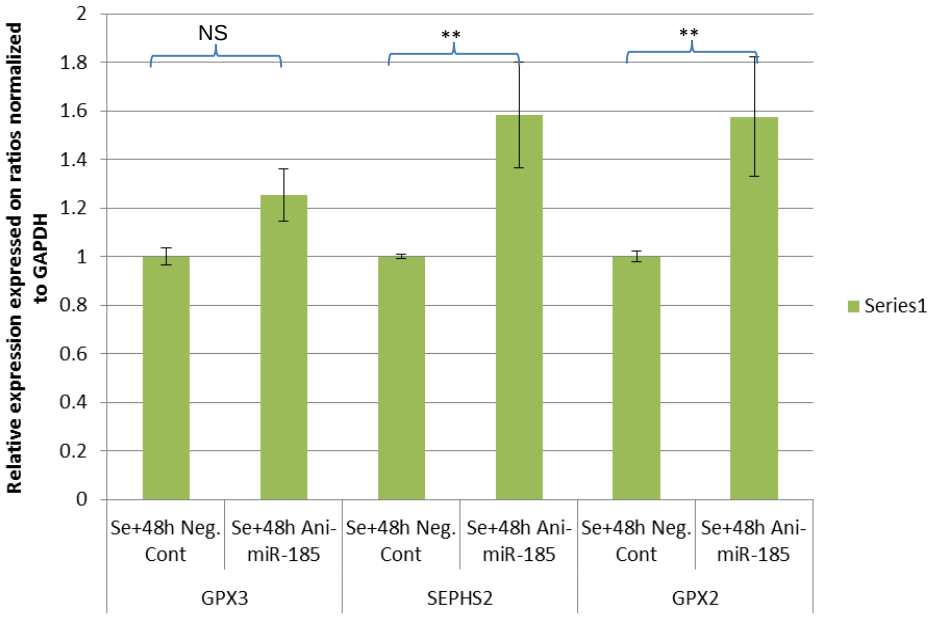


Figure 48 mRNA levels of miR-185 target genes after 48h of anti-miR exposure.
Cells Se-supplemented for 72h and transfection was performed 24h after planted with 70nM of anti-miR185. The expression of the miR-185 was assessed by QRT-PCR and normalized to GAPDH. The expression of GPX3 and SEPHS2 changed significantly at 48h of anti-miR185 exposure; the effects on GPX2 were no significant.

5.3 Discussion

This chapter describes a microarray analysis on the same RNA from Caco-2 cells analysed for miRNA. None of the transcripts from the PBP that were used for the bioinfomatic analysis to predict miRNAs described in chapter 3, were found to form altered expression. Despite the fact that the PBP was previously identified in two tissues from different organisms, changes in gene expression can be triggered differently.

The expression at mRNA level of downstream targets of Se as well as the expression of selenoproteins after Se supplementation or Se deprivation has been investigated previously in several tissues and organisms finding similarities and differences among the data sets. When comparing these data, we have to take into account the fact that the conditions of each experiment are different such as; tissue or cell line organisms and gender in which the experiments have been conducted, time and dose of exposure/deficiency and finally the use of either organic or inorganic form of Se (or a combination of both).

In this chapter, pathway analysis shows a significant effect on the arachidonic acid pathway (Figure 44). Previously it has been suggested that the overall effect of Se supplementation on the arachidonic acid metabolism is the decreased oxidation of arachidonic acid in 15-HpETE, a lipo hydroperoxidase activity metabolite. (Rock and Moos, 2010); from the list of Se sensitive genes listed in Table 8. GPX1, GPX2, GPX3, GPX4 and EPXH2 were part of the inflammatory pathway and arachidonic acid metabolism pathway according to IPA. However, inhibition of Trx in the human umbilical vein cell line, EA.hy926, released arachidonic acid inducing cell toxicity (Crane *et al.*, 2009); this effect was also found in normal subcutaneous areolar and adipose mouse tissue L929 (Kurosawa *et al.*, 2009). The enrichment of a selenoprotein, SELP, as well as Se supplementation of human embryonic kidney cells (HEK-293) prevented oxidation of arachidonic acid metabolism by reducing lipid hydroperoxidases (Rock and Moos, 2010). Another gene from IPA's

arachidonic acid metabolism list, EPHX2, which encodes for the enzyme soluble epoxide hydrolase, was identified as one of the 36 genes of arachidonic acid metabolism from a bioinformatics analysis on prostate cancer databases and up-regulated in prostate cancer samples, its function is to degrade the epoxides formed from the arachidonic acid metabolism (Vainio *et al.*, 2011; Morisseau *et al.*, 2012).

The production of γ-glutamylcysteine from cysteine and glutamate by glutamate cysteine ligase is the first step in glutathione metabolism; after the addition of a glycine by glutathione synthetase, γ-glutamylcysteinylglycine is formed and transported outside the cell and can act as an extra source of cysteine for the cells (Forman *et al.*, 2009). The function of glutathione peroxidases is as a redox regulator but activity of glutamate cysteine ligase and glutathione synthetase also determine the levels of GPX (Bentley *et al.*, 2008). Nevertheless, IPA listed GPX1, GPX2, GPX 3, and GPX 4 on the glutathione metabolism pathway.

Selenophosphate synthetase 2 (SEPHS2) and Methionine adenosyltransferase II β (MAT2B) were identified by IPA as part of the Selenoamino Acid Metabolism pathway. SEPHS2 is directly involved in Se incorporation to selenoproteins through selenophosphate synthesis from selenide prior to Sec- $tRNA^{Sec}$ formation and Sec incorporation facilitated by the SECIS element in the 3'UTR of the selenoprotein mRNA. Also SEPHS2 contains a C-terminal binding domain that might promote ribosome binding (Caban and Copeland, 2012). On the other hand, MAT2B, a methionine adenosyltransferase, forma a heterodimer with MAT2A forming a complex that catalyses the synthesis of the enzyme S-adenosyl methionine (Nordgren *et al.*, 2011). A phylogenetic study of bacterial DNA for S-adenosyl methionine domain containing proteins found a sequence alignment of such proteins with the SECIS element. In one of them, the domain of S-adenosyl methionine interacts with the C-terminal aligned to the stop codon of the S-adenosyl methionine protein family, containing a Cys-rich motif that could potentially regulate SBP2 and selenoprotein incorporation (Haft and Basu, 2011). MAT2B silencing in human hepatocytes led to apoptosis and MAT2B was found to be

overexpressed in cancer cells (Wang *et al.*, 2008). The PBP identified on mouse colon (Kipp *et al.*, 2009) and human lymphocytes (Pagmantidis *et al.*, 2008) was not differentially expressed in the present study except for LAMB1, this is differentially expressed in the human lymphocyte and codes for lamin beta 1, with a -1.24 fold change and its overexpression is linked with the pathogenesis of epilepsy (Wu *et al.*, 2008; Wu *et al.*, 2011b).

IPA toxicology pathways identified oxidative stress and the mitochondrial respiratory pathway as being sensitive to Se supply. Mitochondrial dysfunction and oxidative stress were listed as 3 different toxicology pathways; they are linked by reactive oxygen species (ROS). Se has previously being suggested to have has protective effect on mitochondrial function (Tirosh *et al.*, 2007; Wojewoda *et al.*, 2012) and antioxidant effects through selenoproteins. Production of ROS, oxidizing agents that interact with and modify cellular macromolecules, is triggered by exogenous and endogenous mechanisms as part normal cell physiology where the aim is homeostasis. Nevertheless, ROS production can be pathological if the cell is unable to remove these damaging molecules, resulting in oxidative stress (Lenaz, 2012). The mitochondria respiratory chain is an endogenous mechanism of ROS production, complexes I and III producing the most oxidising molecules. Since glutathione reduces hydrogen peroxide and other radicals, the ROS concentration is glutathione dependent. GPX4 exists as a mitochondrial isoform and knockdown of GPX4 expression lead to increase to mitochondrial ROS production (Cole-Ezea *et al.*, 2012). TrxR also helps ROS balance (Lenaz, 2012) with which GPX 1, GPX2 GPX3 and GPX4 interact in the oxidative stress pathway from the toxicity list. A group of oxidative stress protectors (Vainio *et al.*, 2011) metallothions (MT1E, MT1G, MT1F, MT1X, MT1A, and MT1H), which were up-regulated upon Se treatment, function like glutathione peroxidases; e.g. on cells with insulin sensitivity, metallothioneins act as antioxidant enzymes protecting cells from insulin damage (Park *et al.*, 2011); they have also been suggested to protect colorectal cancer progression caused by oxidative damage (Hamouda *et al.*, 2011). Metallothionein 1F (MT1F) was a

component of the Mitochondria respiration pathway. Evidence of oxidative stress protection by SELP has been observed in HEK-293 cells, where SELP enrichment and Se supplementation protected from oxidative stress, although the effect was lower than the observed for GPX4 (Rock and Moos, 2010).

Even though Fibrinogen Gamma Chain (FFG) expression was up-regulated by of 1.24, this gene was part of several canonical and toxicological pathways, like extrinsic prothrombotic activation pathway, coagulation system, and acute response signalling from the former; and positive acute response proteins for the latter. The only study of Se and fibrinogen levels is the cohort study Coronary Artery Risk Development in Young Adults, in which Se assessment in toenail from 18-30 year old healthy participants found no statistical difference between inflammation biomarkers and the levels of Se in toenails (Xun *et al.*, 2010).

The gene expression responses observed in microarray data from various tissues and cell lines to Se compounds show differences. For example when supplementing a cell model for studying wound healing because of their improper regeneration, keloid fibroblasts, COL15A1, COL5A3, COL3A1, COL1A1, COL6A3 and FLRT3 showed significant downregulation (De Felice *et al.*, 2011); in the human thyroid cancer cell line FRO, a group of genes, mostly growth factors (GADD153, GADD34, EGFR, SGSH, SQSTM1, CAL2, PRKCH, DUSP5, GSK3B, IGF2R, and TRAF1) was up-regulated, while most of the down-regulated genes corresponded to splicing factors (RFX5, SFRS7, ALDHSA1, ZNF207, CDC2, FUBP1, SFPQ, H2AFV, and SFRS2) (Kato *et al.*, 2010). None of these transcripts were detected as differentially expressed in the present study. This response of the genes when supplementation with different selenium compounds has been previously investigated in mice (Irons *et al.*, 2006; Barnes *et al.*, 2009; Kipp *et al.*, 2009) and rats with a particular interest on a comparison between organic and inorganic forms of Se that were supplemented under similar conditions and showed different mRNA expression on certain genes (Barnes *et al.*, 2009). The microarray data sets show different patterns of regulation genes depending on the Se

compound used but overall, the expression of selenoproteins showed. Interestingly on a human intervention study where plasma Se showed high concentrations, SELW, a low hierarchy selenoprotein, showed a downregulation after 10 days of Se supplementation on a human intervention study (Goldson *et al.*, 2011).

In order to get a broad idea of the similarities and differences among Se involving microarrays, the easiest way is to compare the expression of selenium-containing protein up regulation of TRXR-1, PGX1 and GPX4 were found to be up-regulated in human colon carcinoma, no changes in expression were found in TRXR-2 (Yagublu *et al.*, 2011) when TRXR1 and 2 were no expressed on Caco-2 cells. Effects in Gpx1 and on intervention studies where colorectal biopsies from 21 volunteers were supplemented with dairy-Se (150 µg/d) or yeast-Se (150 µg/d) for 6 weeks showed an mRNA upregulation of GPX 1 and 2 only when Se-dairy supplementation. SelP had a higher expression (regardless of the Se compound) that GPX1 and TRXR-1 was the least expressed. (Hu *et al.*, 2011).

Comparison of microarray data intestinal tissue of mice fed either a Se-deficient diet or a Se sufficient diet containing selenomethionine showed a significant downregulation of *Selw*, *Gpx1* and *Gpx2*, similar to the present study, but also Sel*h* and *Selm* which were not differentially expressed in Caco-2 cells. When the expression of *Gpx3*, *Selk*, *Sels*, *Txnrd1*, *Txnrd2*, *Txnrd3*, *Sep15*, and *Selt* expression was investigated with QRT-PCR, a significant change expression was reported that was not evident on the microarray; similar trends in the expression of *Gpx3*, *Selk* and *GPX3* and *SELK* were both were up-regulated by Se in mice and Caco-2 cells, although *Gpx2* and *Sephs2* in mice were down-regulated after feeding Se deficient diet (Kipp *et al.*, 2009), while *GPX2* and *SEPHS2* showed the opposite effect in Caco-2 cells than the one reported in mouse colon (Kipp *et al.*, 2009).

In this chapter using IPA, canonical pathways and toxicological pathways were identified with the group of 53 Se-sensitive genes. Since

IPA is based on published data, the pathways built should represent an accurate picture of the gene interaction for the specific organism chosen for this study, *H. sapiens*. When searching for research papers that would link Se with the networks analysed by IPA, the links between the physiological functions of the genes identified on the transcriptomic experiment for *H. sapiens* were predominantly oxidative stress but also included selenocysteine, glutathione and arachidonic acid metabolism. However, calcium mobilization (Reeves *et al.*, 2010), glucose metabolism (Olsson *et al.*, 2011), and the protein folded response (Kasaikina *et al.*, 2011a), roles of Se previously reported were not identified as Se-sensitive pathways by IPA.

To conclude, the trend in the expression of low hierarchy selenoproteins was observed on the microarray and QRT-PCR, where *GPX1* and *SELW* were up-regulated under Se supplementation. Further supported earlier observed changes in the pattern of gene expression evidence for Se roles in oxidative stress, mitochondria respiration chain, and lipid metabolism when the Se-sensitive genes were grouped into networks. Finally, the changes in expression of the 53 Se-sensitive transcripts can now be matched with the Se-sensitive miRNAs.

With the increasing evidence of the regulatory role of miRNA, several methods have been developed to bioinformatically predict mRNA-miRNA interactions and protein-RNA binding. In this chapter a bioinformatic analysis predicted that MRE corresponding to several Se-sensitive miRNAs are present within the Se-sensitive transcripts identified in Chapter 5. Interestingly, the three genes GPX2, GPX3 and SEHPS2 were predicted to contain MRE for miR-185. GPX2 and GPX3 are known selenoproteins and SEHPS2 is necessary for two steps in incorporation of selenide into selenocysteyl-tRNA^{Sec} during selenoprotein biosynthesis. On the basis that miR-185 was predicted to have multiple targets, that these targets were in the selenoprotein biosynthesis pathway (according to IPA nomenclature) and that miR-185 has been previously reported to be expressed in colorectal tissue (Akcakaya *et al.*, 2011; Liu *et al.*, 2011b),

functional effects of miR knockdown were experimentally investigated for miR-185.

Knockdown of miR-185 was achieved by exposure to anti-miR for 24h, 48h and 72h. This is the amount of time over which miRNA knockdown is generally assessed (Felgner *et al.*, 1987; Hackl *et al.*, 2010; Chiang *et al.*, 2011). The extent of knockdown was of the order of 40-50%. miRNA silencing using antisense miRNAs often uses concentrations of 40 nM when using an Anti-miR or 100 nM when LNA anti-miR oligonucleotides. The effects on the expression of a miRNA and their biological functions when miRNA knockdown experiment are conducted are never predictable; e.g. a miR-675 knockdown of ~50% was achieved on colon cancer cell lines 24h after transfection (with 40nM of Anti-miR-675 in HT-29 cells and this reduced cell proliferation by ~30%; a ~40% silencing of the same miRNA on MIP101cells and decreased cell proliferation by ~35%) more than that reported in HT29 cells with a higher miRNA silencing; a similar phenomenon was observed in SW480 cells where 80% miR-675 silencing only reduced cell proliferation by ~30% (Tsang *et al.*, 2010). Comparison between miRNA silencing on colon cell lines reported in the literature and the one presented in this thesis is often difficult. For example, even though the level of knockdown of miR-140 was not reported, when silencing of miR-140 with 100 nM of LNA anti-miR140 for 72 h in $CD133^{+hi}CD44^{+hi}$ colon cancer stem-like cells, blocked the beneficial effects of miR-140 which lead to DNA damage and increase in proliferation (Song *et al.*, 2009). Also, assuming that miR-34a silencing was achieved (level of knockdown was not registered) with 150pmol of Anti-miR-34a in HT29 cells, miR-34a knockdown progressively increased its target BCL2 protein between 24 h and 72 h (Wang *et al.*, 2010a).

Analysis of the target transcripts after miR knockdown showed that expression of both GPX2 and SEPHS2 were increased after 48h of anti-miR treatment.

From the miRNA target prediction, the MRE for miR-185 in GPX3 and SEPHS2 were within the predicted SECIS element (Table 13)

suggesting that miR-185 might have an effect on transcript stability and/or Sec incorporation. However, miR-185 knockdown did not have a significant effect on GPX3 mRNA expression.

Under Se supplementation conditions, *GPX2* and *SEPHS2* were found to be down-regulated whereas *GPX3* is up-regulated (Chapter 4). miR-185 knockdown at 48h showed a trend towards upregulation of two of its predicted targets (GPX2 and SEPHS2) whereas there was no significant change in expression of *GPX3*. The changes in *GPX2* and *SEPHS2* expression after knockdown are consistent with the observed down-regulation of miR-185 under Se supplementation. Thus, it appears that when miR-185 is down-regulated (either by Se supplementation or anti-miR) expression of GPX2 and SEPHS2 is increased. It is generally recognised that the role of miRNAs is to cleave their target transcripts and down-regulate their expression (Kong *et al.*, 2008; Wahid *et al.*, 2010; Huntzinger and Izaurralde, 2011). The results described in this chapter concur with this mechanism of action, because by inhibiting the miR-185, GPX2 and SEPHS2 transcripts would not have been cleaved and therefore their expression would have been up-regulated. Furthermore, when Caco-2 cells were grown under Se depleted medium, miR-185 was overexpressed while miR-185 targets *GPX2* and *SEPHS2* were down-regulated; so miR-185 silencing in Caco-2 cells during Se supplementation would mimic the expression of the transcripts from cells grown under Se deficient conditions.

Thus, the observed changes in expression of *GPX2* and *SEPHS2* following knockdown of miR-185 suggest that there is a link between miR-185 and these selenoproteins; especially in the case of *SEPHS2*, because the SEPHS2 MRE lies on the SECIS (Table 13), a UGA-SEC codon and a 3′ UTR stem loop which is has a unique sequence and structure (Banerjee *et al.*, 2012) and it is the key element that binds to SBP2 and Ser-tRNASec for selenoprotein incorporation, it is plausible that miR-185 regulates *SEPHS2* expression through selenoprotein incorporation and that *GPX2* change in expression is either due to the mRNA degradation by

miR-185 or because of a secondary effect from the SEPHS2 feedback mechanism.

Further evidence for this being due to miR-mRNA interactions could be obtained by means of a luciferase reporter assay in which the MRE could be inserted and after endogenous miR-185 overexpression supplementing cells with Se, miR-185 would bind to the MRE and repress the expression of luciferase gene as illustrated in Figure 49.

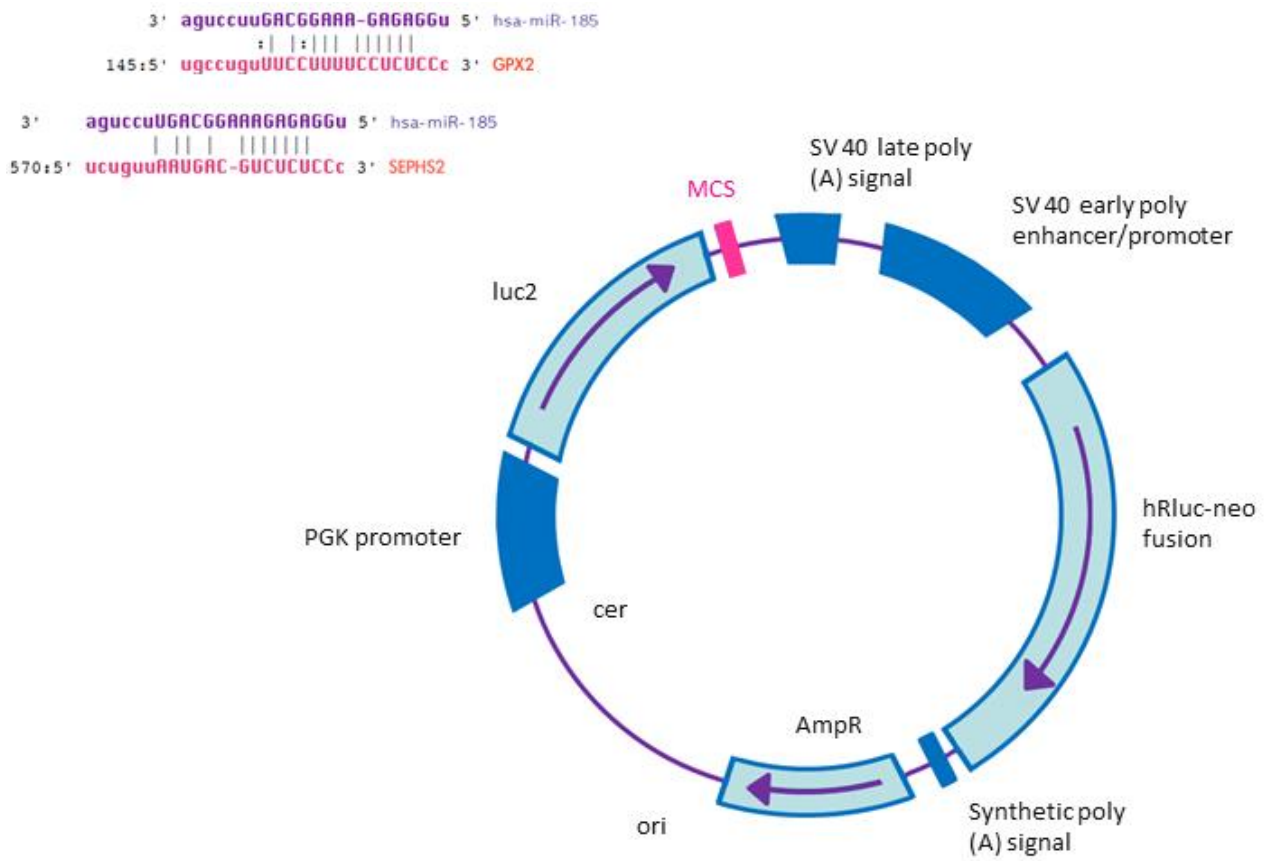


Figure 49 Luciferase miRNA target expression vector.
Luciferase reporter vector in which the MRE (magenta sequence) in inserted into the multiple cloning site (MCS). As the primary reporter gene is firefly luciferase, its expression indicates the cleavage of the miRNA to its target when miRNA upregulation is induced by Se. upon miRNA up-regulation. If GPX2 and SEHPS2 were miR-185 targets, a reduced or no luciferase activity would be expected.

Chapter 6 Final discussion

The work described in this thesis began with a bioinformatic prediction using a group of downstream targets of Se identified in mouse colon (Kipp *et al.*, 2009) and human lymphocytes (Pagmantidis *et al.*, 2008) to identify a miRNA or group of miRNAs that could regulate genes encoding mitochondrial translation factors, ribosomal and elongation factor identified as a protein biosynthetic pathway (Pagmantidis *et al.*, 2008). Those bioinformatics predictions, miRNAs expressed in colon cell lines and tissues, and miRNAs used in PCR arrays were all included in the design of a miRNA microarray Human genome V2 Agilent chip. Using this chip and RNA isolated from Caco-2 cells grown in either Se sufficient/ Se deficient medium, 13 miRNAs were found to be regulated by 40nM sodium selenite supplementation after 72 h. Because the present study had a similar aim, namely to investigate the effects of Se on miRNA and mRNA expression, this cellular model of Se depletion was adopted from Pagmantidis *et al.*, 2005; although the Se content from fetal calf serum varies, the expression of *GPX1* and *SELW* of Caco-2 cells grown in Se-depleted medium are similar to fetal calf serum containing medium (Gautrey *et al.*, 2011) suggesting that normal, unsupplemented medium has a similar Se content to the SE deficient medium used here.

The subsequent step was a transcriptomic experiment using the same RNA from which the 13 Se-sensitive miRNAs were identified and hybridized to Illumina HumanRef-8 v.3 Expression BeadChip; this resulted in the identification of 53 differentially expressed transcripts when Se-supplemented and control cells were compared. In order to find a correlation between these two data sets, miRNA prediction tools described in Chapter 2 were used to predict Se-sensitive miRNA MREs on the Se-sensitive mRNA, and a further pathway analysis with IPA identified the following pathways potentially Se sensitive semaphoring signalling pathway, axon guidance and nervous system development and regeneration associated proteins (Pasterkamp, 2012) with interesting immune functions such as antibody and cytokine production and expression in T and B cells (Mizui *et al.*, 2009). This may explain the link

between Se supplementation and survival of HIV patients (Fawzi *et al.*, 2005; Hurwitz *et al.*, 2007; Harthill, 2011). This pathway analysis suggested as well a link between Se-sensitive miRNAs and the cardiomyopathy observed in Se deficiency (Cui *et al.*, 2012) (possibly Keshan disease), through the cardiac β-adrenergic signalling; the ability of the cardiac muscle to contract when this signalling pathway is compromised (Saucerman and McCulloch 2006).

While cellular functions such as extracellular communication processes by activated by intracellular responses via ERK/MAPK signalling pathway (McKay and Morrison, 2007) were also identified by IPA; most pathways were related to cancer proliferation and invasiveness. This included integrin signalling pathway (cell cycle regulation that controls cancer progression through Cas, Crk proteins and IPP complex) (Cabodi *et al.*, 2010), HMG1 signalling (cancer proliferation, inflammation and apoptosis regulator) (Jia *et al.*, 2013); glioblastoma multiform pathway which is expressed in brain tumours where vascular proliferation, tumour chemoresistance and apoptosis are characteristics (Mao *et al.*, 2012) and finally, pancreatic signalling pathways which has been linked to cell apoptosis and dysplasia in response to DNA damage (McCleary-Wheeler *et al.*, 2012).

Pathway analysis also suggested that the Se-sensitive miRNAs could have metabolic regulatory functions in Caco-2 cells. For example, the pathways potentially affected were: aryl hydrocarbon receptor signalling (a metabolic pathway activator) (Barouki *et al.*, 2012); alternative carbohydrate and lipid metabolism functions of the p53 signalling pathway (Zhang *et al.*, 2010), and fatty acid (Harvey and Leinwand, 2011) and glucose (Mihaylova and Shaw, 2011) metabolism from the cardiac hypertrophy signalling.

Based on the bioinformatics predictions described in Chapter 5 and the biological functions of the group of transcripts in which MRE were identified, miR-185 and its targets *GPX3*, *SEPHS2*, and *GPX2* were chosen for functional analysis. In order to mimic endogenous expression

of miR-185 in Caco-2 cells, cells were supplemented with Se and miR-185 knockdown achieved with an antisense oligonucleotide complementary to miR-185. The knockdown was monitored at 24, 36, and 72h and the expression of *GPX3*, *SEPHS2*, and *GPX2* was assessed by QRT-PCR from which data, at 48h, the expression of *SEPHS2*, and *GPX2* was found to be up-regulated compared to the control.

miRNA are flexible in their recognition of target mRNA sequences through an imperfect base pairing. This gives them the possibility to regulate a group of genes (Brodersen and Voinnet, 2009; Schnall-Levin *et al.*, 2011; Subramanyam and Blelloch, 2011). Several algorithms and pathway analysis tools have been developed in order to identify possible miRNA-mRNA interactions (Grocock *et al.*; Betel *et al.*, 2008; Ingenuity, 2008; Antonov *et al.*, 2009; Dweep *et al.*, 2011); however, tissue localization, biological functions, and miRNA-mRNA interactions that are do not occur in the seed region or are not conserved among species will not be predicted by those tools because those are part of the criteria of exclusion to avoid false positives.

From the 750 entities included on the miRNA chip design, 145 miRNA were expressed in Caco-2 cells. 71 of these were part of the 194 miRNAs on the chip that had previously been found to be expressed in intestinal cells or tissues; 15 miRNAs expressed were bioinformatically predicted and do not correspond to either to any previously described in the intestinal microRNAome or included on PCR arrays or microarrays. Discussing which would be the most efficient bioinformatics tool for this research would not be relevant because these predictions were done originally in a group of transcripts that were not differentially expressed upon Se supplementation/depletion.

Little or nothing (in the case of miR-532-5p) is known about the Se-sensitive miRNAs identified in Chapter 4. As regards miR-625 and miR-429, mir-429 has been reported to be expressed at low levels in colon cell lines (Stratmann *et al.*, 2011; Sun *et al.*, 2011b; Zhang *et al.*, 2011b) and suggested to be a colorectal cancer biomarker (Zhang *et al.*, 2011b).

The experimentally validated target of an isoform of miR-429 in colorectal tissues is C-MYC which is a cell cycle and apoptosis promoter silenced by miR-429b (Sun *et al.*, 2011b), which could classify it as a cancer suppressor miRNA. Recently, miR-625 was detected as at high expression levels in human oocytes (Xu *et al.*, 2011) and overexpressed in neoplastic pancreatic biopsies compared to healthy tissue counterparts (Yu *et al.*, 2012).

Interestingly, Se supplementation led to the expression of miR-375* whereas the mature miRNA, miR-375, was expressed in the Caco-2 cell line but there was no change of expression between Se supplemented cells and controls. miR-375* is down-regulated in bone marrow from leukaemia patients (Ju *et al.*, 2009), melanoma cells and primary human melanocyte (Satzger *et al.*, 2010), and overexpressed in retinoblastoma (Zhao *et al.*, 2009). Oncogene miR-30b is highly expressed in invasive urothelial carcinoma of the bladder, (Wszolek *et al.*, 2011) and medulloblastoma (Lu *et al.*, 2009), but is a tumour suppressor in hepatocytes (Gaziel-Sovran *et al.*, 2011).

There is even more limited information about the roles of the Se-sensitive miRNA in the colon. High levels of miR-10b were related to malignancy in Irish colorectal cancer samples (Chang *et al.*, 2011). Overexpression of miR-196b cancer in gastric biopsies (Tchernitsa *et al.*, 2010b); in colon cancer biopsies (Wang *et al.*, 2010b), as well as in circulating blood from cancer patients (Tsujiura *et al.*, 2010) was found in infiltrative cancer. Tumour suppression properties of miR203 were observed in HT29 cells and gastric and colorectal biopsies(Li *et al.*, 2011). Down-regulation of miR10b has been related to lower tumour growth (Chiang *et al.*, 2011).

The exogenous overexpression of miR-28-5p in HCT116 and RKO cells led to a possible downregulation of CCND1 and HOX83 and decreased tumour proliferation (Almeida *et al.*, 2012). In HT29 cells, miR-22 overexpression was linked to increased apoptosis (Zhang *et al.*, 2011a); and cell viability was decreased by the up-regulation of miR-494

followed by the downregulation of its target KIT, an apoptotic inhibitor (Kim *et al.*, 2011c; Casar-Borota *et al.*, 2012).

miR-185 expression was not detected on the miRNA microarray under Se depletion conditions, but the addition of Se had a large an effect on this miRNA. Using QRT-PCR, miR-185 showed one of the biggest changes after Se supplementation when normalized to let-7b and the change was statistically significant when normalized to RNU6B. miR-185 roles in colorectal progression cancer were recently proposed; for example metastasis and low cancer survival were found during miR-185 overexpression (Akcakaya *et al.*, 2011) and miR-185 downregulation SW1116 cells mimicked RhoA/Cdc42 (Liu *et al.*, 2011b).

Because miRNA roles are different from tissue/cell line to tissue/cell line, it was important to determine the expression of the miRNAs on our experimental model Caco-2. Identifying their targets was then the next step for which a total transcriptome experiment hybridized to Illumina HumanRef-8 v.3 was performed so causation between the Se-sensitive miRNAs and the Se-sensitive genes could be found among these dataset. 53 genes were differentially expressed on the microarray from which *SEPHS2*, *GPX2*, *DCUN1D4*, *ME1*, *PIR*, *HTATIP2*, *KLF5*, *GCNT3*, *C6ORF192*, *CEACAM6*, *GGH*, *LAMB1*, *OIP5*, *OSTC*, *SLC1A3*, *ATP1B3*, and *GPT2* were down-regulated; from this group, only *GPX2* and *SEPHS2* were known to be linked to Se metabolism.

Transcripts other than Se proteins were identified after Se supplementation in the human lymphocyte, including 61 which are ribosomal, mitochondrial, elongation factors were differentially expressed after 100 µg sodium selenite (Pagmantidis *et al.*, 2008), except for *LAMB1* which codes for an extracellular matrix glycoprotein, laminin beta 1 (Wu *et al.*, 2011b). In a different study, *COL15A1, COL5A3*, *COL3A1*, *COL1A1*, *COL6A3* and *FLRT3* showed significant downregulation in keloid fibroblasts supplemented with 20 µg/day of selenocystine (De Felice *et al.*, 2011); human thyroid cancer cells supplemented with 150 µM seleno-l-methionine *GADD153*, *GADD34*, *EGFR*, *SGSH*, *SQSTM1*, *CAL2*, *PRKCH*,

DUSP5, *GSK3B*, *IGF2R*, and *TRAF1* (growth factors) were up-regulated whereas splicing factors *RFX5*, *SFRS7*, *ALDHSA1*, *ZNF207*, *CDC2*, *FUBP1*, *SFPQ*, *H2AFV*, and *SFRS2* were down-regulated (Kato *et al.*, 2010). None of the mentioned mRNA changes were found in the present study.

Of the differentially expressed selenoproteins identified in different gastrointestinal tissues or cell lines, *GPX1* and *GPX4* were up-regulated (Yagublu *et al.*, 2011) when cancer and healthy colonic tissue were compared. Effects on *SELP* were greater than *GPX1* in an intervention study of 21 volunteers supplemented with dairy-Se (150 µg/d) or yeast-Se (150 µg/d) (Hu *et al.*, 2011). In the present study, SELP was not differentially expressed, possibly because tissues are more exposed to circulating blood in which SELP is localized (Rock and Moos, 2010). In a human intervention study, SELW was surprisingly down-regulated after 10 days of Se supplementation (Goldson *et al.*, 2011) where this selenoprotein is rapidly saturated when Se is abundant. Selenoproteins SELP, TRXR1 and 2 were not expressed in the present Caco-2 study, but have been reported to be expressed in colon carcinoma tissue (Yagublu *et al.*, 2011) and colorectal biopsies (Hu *et al.*, 2011).

Further bioinformatic analysis of the 53 Se-sensitive transcripts revealed some networks related with biological roles of Se, as described in Chapter 5; one of these was arachidonic acid metabolism. Previous work has suggested that Se-antioxidant systems, particularly TrxR, influence regulation of arachidonic acid metabolism (Kurosawa *et al.*, 2009). Interestingly, there are similarities in pathways affected by dehydroepiandrosterone (DHEA), a chemopreventive and antiproliferative steroid, and the ones found after pathway analysis with the Se-sensitive genes from the present study. DHEA was used to incubate Sk-Gi and Sk-Sc cells (derived from the liver adenocarcinoma cell line SK-HEP-1) and a number of metabolites were chromatographically obtained, from which pathways were mapped with GeneSpring MS software revealing lipid metabolism, glutathione metabolism, and *s*-adenosylmethionine (SAM) metabolism to be affected (Cheng *et al.*, 2011). In addition, in BAEC cells

Se depletion accelerated oxidation of lipids associated with arachidonic acid metabolism (Rock and Moos, 2010), supporting the hypothesis that Se has a role in arachidonic metabolism. GPX1, GPX2, GPX3, GPX4 and EPXH2 are listed as part of the arachidonic acid metabolism as according to IPA. Indeed there is experimental evidence for a role of EPXH2 in arachidonic metabolism (Vainio *et al.*, 2011; Morisseau *et al.*, 2012).

Bioinformatic analysis linked Se-sensitive miRNAs with Se-sensitive mRNAs. ME1, CACYBP, ATP6V1B1 and GPX1had MRE for miR-625, SELK for miR-492, HTATIP2 and GPT2 for miR-373*; GTP2 for miR-22, EPHX2 for miR-532-5p, GPX2 and MT1F for miR-106b, GPX3 for miR-30b; and finally GPX2, GPX3, and SEPHS2 for miR-185. miRNA targets for miR-625, miR429, miR-373*, miR-22, miR-532-5p do not have any evident link to Se metabolism or biological functions. miRNA targets for miR-106b, miR-30b, and miR-185 showed a more evident link.

From those three, miR-185 targets are part of the selenoprotein metabolism and the MRE for GPX3 and SEPHS2 are within the sequence predicted to form the SECIS element according to the SECIS prediction software (Kryukov *et al.*, 2003). The latter observation suggests that GPX3 and SEPHS2 biosynthesis may be mediated by miR185. Since the SECIS functions by recruiting binding proteins into the Secys-tRNA incorporation complex, it is possible that miR-185 could regulate GPX3 and SEPHS2 expression not only by regulating mRNA degradation through the RISC complex but also by regulating their translation.

Knockdown of miR-185 was achieved in Se-supplemented cells after incubation for 24, 36, and 72 h. In spite of a greater miR-185 knockdown being obtained at 24h compared with 48 h, miRNA targets only showed statistically significant changes in expression at 48 h, with an up-regulation of *GPX2* and *SEPHS2*. This is consistent with the predicted miRNA target cleavage and miRNA silencing reported mechanism of action (Kong *et al.*, 2008; Wahid *et al.*, 2010; Huntzinger and Izaurralde, 2011). The changes in expression only at 48 h with no effect at 24 or 72 h could be due to the experimental design in which up-regulation of miR-185

by Se was done 48 h prior to miRNA knockdown and as a result knockdown over 24 h might not give enough time to the cells to trigger any feedback mechanism or for the mRNA to recover from miRNA cleavage. Even though knockdown was achieved at 72 h, the miR-185 expression was not down-regulated to as great an extent as at 24 or 48 h and this phenomenon could be explained by a weak effect of the anti-miR at 72 h because at 72 h the antisense oligonucleotide has been degraded or because during this simultaneous treatment, the addition of Se that upregulates miR-185 expression is more efficient than miR-185 silencing despite the high Anti-miR concentration of 70nM.

In conclusion, the main findings of the study described in this thesis are that:

1) Se supplementation of Caco-2 cells did not only change the expression of specific selenoproteins, but also altered a number of non-selenoprotein genes, pathways, downstream targets, and 13 miRNAs.
2) The miRNAs are predicted to have MRE in some of the transcripts affected by Se supply.
3) Knockdown of miR-185 affected *GPX2* and *SEPHS2* expression consistent with Se depleted cells microarray data.

On the basis of these finding, I propose that after Se supplementation, changes in expression lead to an unknown biological events that modulates expression of *GPX2* and *SEPHS2* through miRNAs. This is illustrated in Figure 50, Se mechanisms of action are represented by selenium incorporation into selenoproteins through a selenoprotein hierarchy which could be regulated through miRNAs targeting the SECIS element and/or regulating genes involved in selenoprotein biosynthesis.

As a result of the target prediction of the Se-sensitive miRNAs and further pathways analysis, metabolic pathways like fatty acid and glucose; physiological pathways like β-adrenergic receptors disruption that leads to cardiomyopathy and immune system enhancement; and at a cellular level,

cell cycle regulatory function and cancer signalling pathways that control cancer invasiveness, proliferation and apoptosis, were recognised as possible Se-sensitive miRNA mechanisms of action.

From the IPA analysis from the transcriptomic experiment, Se-related pathways like selenoamino acid and glutathione metabolism and the toxicological responses like oxidative stress and decrease mitochondrial respiration.

If more time had been available, this hypothesis could have been tested by assessing the expression GPX3, GPX2, and SEPHS2 at the protein level so as to investigate the effects of miR-185 on translation; future work could also include the cloning of the MRE followed by overexpression of miR-185 with Se to test directly whether the observed effects are due to a miRNA-mRNA interaction.

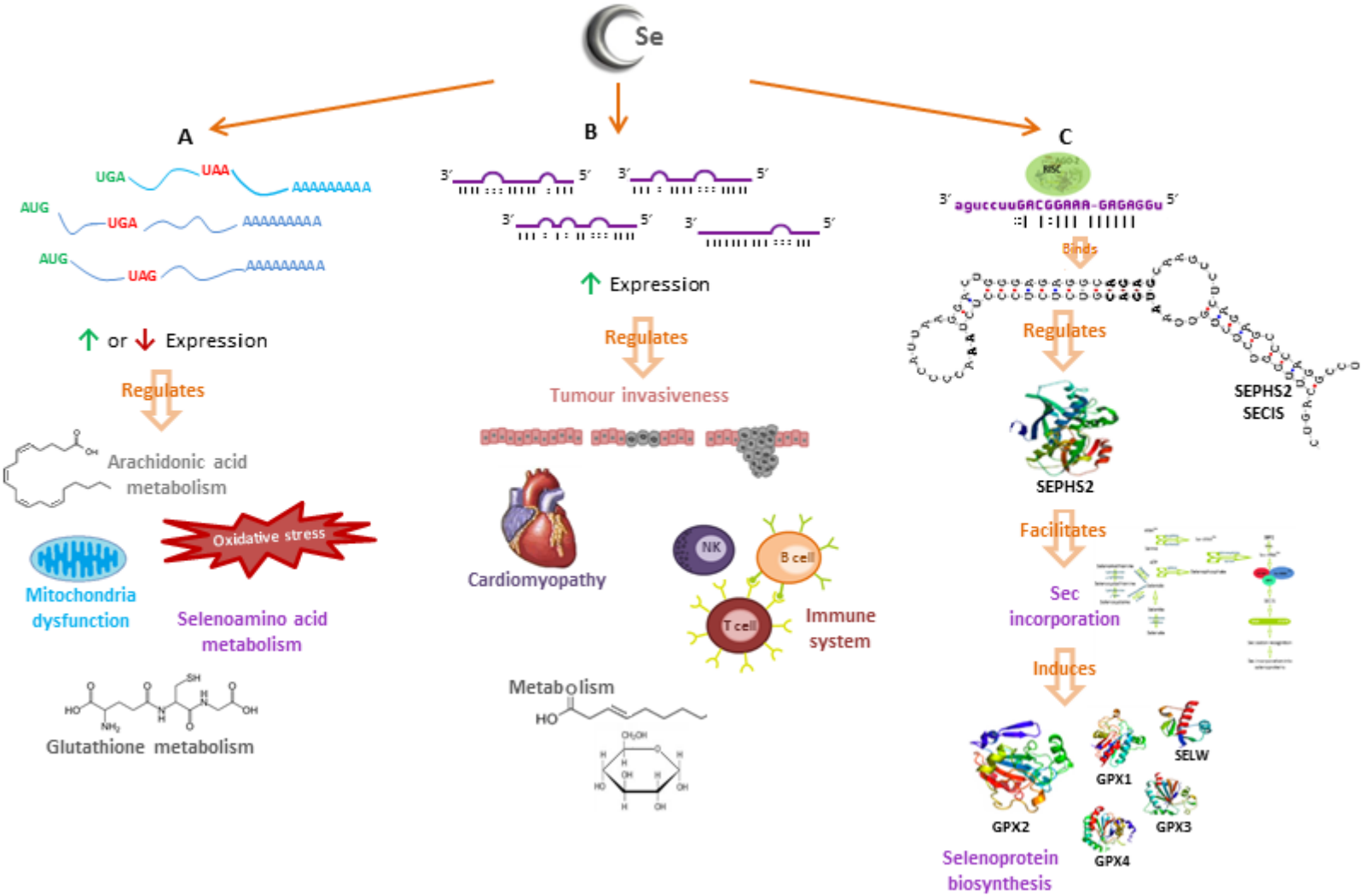

Se
A
UGA
UAA
AAAAAAAAA
AUG
UGA
AAAAAAAAA
AUG
UAG
AAAAAAAAA
↑ or ↓ Expression
Regulates
Arachidonic acid metabolism
Oxidative stress
Mitochondria dysfunction
Selenoamino acid metabolism
Glutathione metabolism
B
↑ Expression
Regulates
Tumour invasiveness
Cardiomyopathy
NK
B cell
T cell
Immune system
Metabolism
C
RISC
3′ aguccuuGACGGAAA-GAGAGGu 5′
Binds
Regulates
SEPHS2
SECIS
SEPHS2
Facilitates
Sec incorporation
Induces
GPX1
SELW
GPX2
GPX3
GPX4
Selenoprotein biosynthesis

Figure 50 Mechanisms of action of Se involving miRNAs: A hypothesis.
Se supplementation on Caco-2 cells affects the expression of miRNAs and mRNAs which may regulate cellular functions through three possible mechanisms:
A) The 53 genes for which expression was up or down-regulated upon Se supplementation were analysed by IPA, and the canonical pathways identified were arachidonic acid metabolism linked to decreased oxidation of arachidonic acid; glutathione metabolism, selenoamino acid metabolism; toxicity lists identified oxidative stress and mitochondrial dysfunction both of which are processes related to ROS and the antioxidant effects of selenoproteins related cellular processes.
B) Se supplementation up-regulated 13 miRNAs; pathway analysis with the predicted targets of these Se-sensitive miRNAs suggested that survival of immune compromised patients, myocardial lesion development (under Se deficiency), cancer proliferation and invasiveness and lipid and carbohydrate metabolism (alternative functions of the p53 and Akt signalling pathways) as well as extracellular communication were suggested Se-sensitive miRNA regulatory mechanisms.
C) Knockdown of miR-185, a Se-sensitive miRNA, affected the expression of SEPHS2 and GPX2 two selenoproteins predicted to be miR-185 targets. MRE for SEPHS2 matches the SECIS, which suggests that miR-185 may regulate selenoprotein incorporation; and that GPX2 overexpression in response to miR-185 knockdown could be either a direct interaction between the mRNA-miRNA cleavage that would lead to GPX2 mRNA degradation under Se-supplementation conditions (which up-regulates miR-185), or an indirect interaction between SEPHS2 SECIS' MRE and miR-185.

Appendix A miRNA microarray chip design

Table 14 lists the miRNAs included on the miRNA microarray design human genome V2 Agilent customised 8x15K slide. Chapter 3 describes the bioinformatics analysis, literature review and array platforms from which Table 14 was obtained.

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	MicroCosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-let-7a	MIMAT0000062	*				*		[1-3]
hsa-let-7a*	MIMAT0004481						*	
hsa-let-7b	MIMAT0000063	*				*		[1,2]
hsa-let-7b*	MIMAT0004482						*	
hsa-let-7c	MIMAT0000064	*				*		[1,3]
hsa-let-7c*	MIMAT0004483						*	
hsa-let-7d	MIMAT0000065	*				*		[1, 3, 4]
hsa-let-7d*	MIMAT0004484						*	
hsa-let-7e	MIMAT0000066	*				*		[1, 3]
hsa-let-7e*	MIMAT0004485						*	
hsa-let-7f	MIMAT0000067	*				*		[1-3]
hsa-let-7f-1*	MIMAT0004486						*	
hsa-let-7f-2*	MIMAT0004487						*	
hsa-let-7g	MIMAT0000414	*				*		[1-3, 5]
hsa-let-7i	MIMAT0000415	*				*		[1-3]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	MicroCosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-1	MIMAT0000416	*	*					[3, 6]
hsa-miR-100	MIMAT0000098	*						[3, 4]
hsa-miR-101	MIMAT0000099	*			*	*		[3]
hsa-miR-103	MI000108							[3-5]
hsa-miR-103a	MIMAT0000101							
hsa-miR-105	MIMAT0000102				*	*		[5]
hsa-miR-106a	MIMAT0000680	*		*	*	*		[3-5]
hsa-miR-106b	MIMAT0004672							[3]
hsa-miR-107	MIMAT0000104					*		[3]
hsa-miR-10a	MIMAT0000253	*			*			[3-5]
hsa-miR-10a*	MIMAT0004555						*	
hsa-miR-10b	MIMAT0000254	*		*	*			[3, 6]
hsa-miR-10b*	MIMAT0004556						*	
hsa-miR-1178	MIMAT0005823						*	
hsa-miR-1180	MIMAT0005825						*	
hsa-miR-1185	MIMAT0005798						*	
hsa-miR-1197	MIMAT0005955						*	
hsa-miR-1201	MIMAT0005864						*	
hsa-miR-1202	MIMAT0005865						*	
hsa-miR-1204	MIMAT0005868						*	
hsa-miR-1205	MIMAT0005869						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-1206	MIMAT0005870						*	
hsa-miR-122	MIMAT0000421	*	*					[5]
hsa-miR-1224-3p	MIMAT0005459						*	
hsa-miR-1224-5p	MIMAT0005458						*	
hsa-miR-1225-3p	MIMAT0005573						*	
hsa-miR-1225-5p	MIMAT0005572						*	
hsa-miR-1226	MIMAT0005577						*	
hsa-miR-1231	MIMAT0005586						*	
hsa-miR-1237	MIMAT0005592						*	
hsa-miR-1238	MIMAT0005593						*	
hsa-miR-124	MIMAT0000422	*						[3, 7]
hsa-miR-124*	MIMAT0004591						*	
hsa-miR-1243	MIMAT0005894						*	
hsa-miR-1244	MIMAT0005896						*	
hsa-miR-1245	MIMAT0005897						*	
hsa-miR-1246	MIMAT0005898						*	
hsa-miR-1247	MIMAT0005899							
hsa-miR-1251	MIMAT0005903							
hsa-miR-1257	MIMAT0005908						*	
hsa-miR-1258	MIMAT0005909						*	
hsa-miR-1259	MIMAT0005910						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-125a-3p	MIMAT0004602						*	
hsa-miR-125a-5p	MIMAT0000443	*						
hsa-miR-125b	MIMAT0000423	*				*		[3-5]
hsa-miR-125b-1*	MIMAT0004592						*	
hsa-miR-125b-2*	MIMAT0004603						*	
hsa-miR-126	MIMAT0000445	*						[3, 5, 6]
hsa-miR-126*	MIMAT0000444							
hsa-miR-1260	MIMAT0005911						*	
hsa-miR-1261	MIMAT0005913						*	
hsa-miR-1262	MIMAT0005914						*	
hsa-miR-1263	MIMAT0005915						*	
hsa-miR-1265	MIMAT0005918						*	
hsa-miR-1266	MIMAT0005920						*	
hsa-miR-1268	MIMAT0005922						*	
hsa-miR-1269	MIMAT0005923						*	
hsa-miR-1270	MIMAT0005924						*	
hsa-miR-1272	MIMAT0005925						*	
hsa-miR-1273	MIMAT0005926						*	
hsa-miR-127-3p	MIMAT0000446	*						
hsa-miR-1274a	MIMAT0005927						*	
hsa-miR-1275	MIMAT0005929						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-128a	MIMAT0000424							[3, 5]
hsa-miR-128b	MIMAT0005799						*	[3, 5]
hsa-miR-1284	MIMAT0005941						*	
hsa-miR-1287	MIMAT0005878						*	
hsa-miR-1288	MIMAT0005942						*	
hsa-miR-1289	MIMAT0005879						*	
hsa-miR-129*	MIMAT0004548						*	
hsa-miR-1290	MIMAT0005880						*	
hsa-miR-1291	MIMAT0005881			*				
hsa-miR-1292	MIMAT0005943						*	
hsa-miR-1293	MIMAT0005883						*	
hsa-miR-129-3p	MIMAT0004605	*						
hsa-miR-1294	MIMAT0005884	*					*	
hsa-miR-129-5p	MIMAT0000242							
hsa-miR-1298	MIMAT0005800							
hsa-miR-1303	MIMAT0005891							
hsa-miR-1308	MIMAT0005947			*	*		*	
hsa-miR-130a	MIMAT0000425	*			*			[3-6]
hsa-miR-130a*	MIMAT0004593						*	
hsa-miR-130b	MIMAT0000691	*	*					[3, 5]
hsa-miR-130b*	MIMAT0004680						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-132	MIMAT0000426	*	*					[3]
hsa-miR-132*	MIMAT0004594							[7]
hsa-miR-1321	MIMAT0005952							
hsa-miR-1322	MIMAT0005953							
hsa-miR-1324	MIMAT0005956							
hsa-miR-133a	MIMAT0000427							[3, 6]
hsa-miR-133b	MIMAT0000770							[1, 2, 3, 7]
hsa-miR-134	MIMAT0000447	*						[5]
hsa-miR-135a	MIMAT0000428	*	*					[3, 5]
hsa-miR-135a*	MIMAT0004595							
hsa-miR-135b	MIMAT0000758	*	*					[1, 3, 6, 7]
hsa-miR-135b*	MIMAT0004698							
hsa-miR-136	MIMAT0000448					*		[3, 6]
hsa-miR-136*	MIMAT0004606						*	
hsa-miR-137	MIMAT0000429	*				*		
hsa-miR-138	MIMAT0000430	*						[3]
hsa-miR-138-1*	MIMAT0004607						*	
hsa-miR-138-2*	MIMAT0004596						*	
hsa-miR-139-3p	MIMAT0004552						*	
hsa-miR-139-5p	MIMAT0000250	*				*		
hsa-miR-140-3p	MIMAT0004597	*						

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-140-5p	MIMAT0000431	*						
hsa-miR-141	MIMAT0000432	*	*		*	*		[2, 3, 4]
hsa-miR-141*	MIMAT0004598				*			
hsa-miR-142-3p	MIMAT0000434	*						[3, 5]
hsa-miR-142-5p	MIMAT0000433	*				*		[3, 5, 6]
hsa-miR-143	MIMAT0000435	*						[1-3]
hsa-miR-143*	MIMAT0004599						*	
hsa-miR-144	MIMAT0000436	*	*		*		*	[3]
hsa-miR-144*	MIMAT0004600				*			
hsa-miR-145	MIMAT0000437						*	[1, 2, 3, 5, 7]
hsa-miR-145*	MIMAT0004601							
hsa-miR-1468	MIMAT0006789						*	
hsa-miR-146a	MIMAT0000449	*						[3]
hsa-miR-146a*	MIMAT0004608						*	
hsa-miR-146b-3p	MIMAT0004766						*	
hsa-miR-146b-5p	MIMAT0002809							
hsa-miR-147	MIMAT0000251					*		[5]
hsa-miR-147b	MIMAT0004928						*	
hsa-miR-148a	MIMAT0000243	*						[2, 3, 5]
hsa-miR-148a*	MIMAT0004549						*	
hsa-miR-148b	MIMAT0000759							[3]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-148b*	MIMAT0004699							
hsa-miR-149	MIMAT0000450	*				*		[3, 5]
hsa-miR-149*	MIMAT0004609						*	
hsa-miR-150	MIMAT0000451	*				*		[3]
hsa-miR-150*	MIMAT0004610						*	
hsa-miR-151-3p	MIMAT0000757	*						
hsa-miR-151-5p	MIMAT0004697	*						
hsa-miR-152	MIMAT0000438	*	*					[3]
hsa-miR-153	MIMAT0000439					*		[3]
hsa-miR-1537	MIMAT0007399							
hsa-miR-154	MIMAT0000452							[3]
hsa-miR-154*	MIMAT0000453							[5]
hsa-miR-155	MIMAT0000646	*	*			*		[1, 3]
hsa-miR-155*	MIMAT0004658						*	
hsa-miR-15a	MIMAT0000068	*		*		*		[1, 3, 5]
hsa-miR-15a*	MIMAT0004488							
hsa-miR-15b	MIMAT0000417	*		*		*		[2, 3, 5]
hsa-miR-15b*	MIMAT0004586						*	
hsa-miR-16	MIMAT0000069	*	*	*		*		[1-3, 5]
hsa-miR-16-1*	MIMAT0004489						*	
hsa-miR-16-2*	MIMAT0004518							

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-17	MIMAT0000070	*	*		*	*		[1]
hsa-miR-17-5p	MIMAT0000071	*			*			[4, 5]
hsa-miR-181a	MIMAT0000256	*						[3, 5]
hsa-miR-181a-2*	MIMAT0004558						*	
hsa-miR-181b	MIMAT0000257	*						[3, 5, 6]
hsa-miR-181c	MIMAT0000258	*						[3]
hsa-miR-181c*	MIMAT0004559						*	
hsa-miR-181d	MIMAT0002821							[3]
hsa-miR-182	MIMAT0000259	*				*		[3, 5, 6]
hsa-miR-182*	MIMAT0000260							[5]
hsa-miR-1825	MIMAT0006765						*	
hsa-miR-1826	MIMAT0006766						*	
hsa-miR-1827	MIMAT0006767						*	
hsa-miR-183	MIMAT0000261	*			*	*		[3, 7]
hsa-miR-183*	MIMAT0004560						*	
hsa-miR-184	MIMAT0000454							[3, 5]
hsa-miR-185	MIMAT0000455	*		*	*			
hsa-miR-185*	MIMAT0004611						*	
hsa-miR-186	MIMAT0000456	*	*			*		[3, 5]
hsa-miR-186*	MIMAT0004612							
hsa-miR-187	MIMAT0000262		*					[5, 6]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-187*	MIMAT0004561						*	
hsa-miR-188-3p	MIMAT0004613						*	
hsa-miR-188-5p	MIMAT0000457							
hsa-miR-18a	MIMAT0000072							[3]
hsa-miR-18b	MIMAT0001412	*						[3]
hsa-miR-190	MIMAT0000458		*					[3, 5]
hsa-miR-191	MIMAT0000440	*						[3, 5]
hsa-miR-192	MIMAT0000222	*						[3, 6]
hsa-miR-192*	MIMAT0004543							
hsa-miR-193a-3p	MIMAT0000459				*			
hsa-miR-193a-5p	MIMAT0004614	*				*		
hsa-miR-193b	MIMAT0002819	*			*			[3]
hsa-miR-194	MIMAT0000460	*						[3, 5, 6]
hsa-miR-195	MIMAT0000461	*				*		[3, 5]
hsa-miR-196a	MIMAT0000226	*	*		*	*		[3, 6]
hsa-miR-196a*	MIMAT0004562				*			
hsa-miR-196b	MIMAT0001080	*	*			*		[3]
hsa-miR-197	MIMAT0000227	*				*		[3, 5, 6]
hsa-miR-198	MIMAT0000228		*					
hsa-miR-199a-3p	MIMAT0000232	*					*	
hsa-miR-199a-5p	MIMAT0000231					*		

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-199b-5p	MIMAT0000263					*		
hsa-miR-19a	MIMAT0000073	*	*					[3, 5, 7]
hsa-miR-19a*	MIMAT0004490						*	
hsa-miR-19b	MIMAT0000074	*				*		[3, 6]
hsa-miR-19b-1*	MIMAT0004491							
hsa-miR-19b-2*	MIMAT0004492						*	
hsa-miR-200a	MIMAT0000682				*	*		[2-5]
hsa-miR-200a*	MIMAT0001620				*			
hsa-miR-200b	MIMAT0000318				*	*		[2-5]
hsa-miR-200b*	MIMAT0004571				*			
hsa-miR-200c	MIMAT0000617	*			*	*		[2, 3, 5]
hsa-miR-200c*	MIMAT0004657				*			
hsa-miR-202	MIMAT0002811	*					*	[6]
hsa-miR-203	MIMAT0000264	*		*	*	*		[3, 5]
hsa-miR-204	MIMAT0000265					*		[3]
hsa-miR-205	MIMAT0000266	*						[3]
hsa-miR-206	MIMAT0000462		*					
hsa-miR-208a	MIMAT0000241						*	
hsa-miR-208b	MIMAT0004960						*	
hsa-miR-20a	MIMAT0000075	*	*			*		[2-4]
hsa-miR-20a*	MIMAT0004493							

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-20b	MIMAT0001413	*				*		[3]
hsa-miR-20b*	MIMAT0004752						*	
hsa-miR-21	MIMAT0000076	*	*		*			[1, 2, 3, 5, 7]
hsa-miR-21*	MIMAT0004494				*			
hsa-miR-210	MIMAT0000267	*						[3, 5]
hsa-miR-211	MIMAT0000268					*		
hsa-miR-212	MIMAT0000269						*	
hsa-miR-214	MIMAT0000271	*	*			*		[3, 5]
hsa-miR-214*	MIMAT0004564						*	
hsa-miR-215	MIMAT0000272							[3]
hsa-miR-216	MIMAT0000273							[5]
hsa-miR-216a	MIMAT0004959						*	
hsa-miR-216b	MIMAT0000274		*			*		
hsa-miR-218	MIMAT0000275	*						[3, 5]
hsa-miR-218-1*	MIMAT0004565		*				*	
hsa-miR-218-2*	MIMAT0004566		*				*	[4]
hsa-miR-219-1-3p	MIMAT0004567					*		
hsa-miR-219-2-3p	MIMAT0004675					*		
hsa-miR-219-5p	MIMAT0000276				*	*		
hsa-miR-22	MIMAT0000077	*		*	*			[2, 3, 6]
hsa-miR-22*	MIMAT0004495							

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-220b	MIMAT0004908					*		
hsa-miR-220c	MIMAT0004915						*	
hsa-miR-221	MIMAT0000278							[3-5]
hsa-miR-222	MIMAT0000279	*	*		*			[3-5]
hsa-miR-222*	MIMAT0004569				*			
hsa-miR-223	MIMAT0000280	*			*		*	[2, 3, 5]
hsa-miR-223*	MIMAT0004570							
hsa-miR-224	MIMAT0000281							[3, 5]
hsa-miR-23a	MIMAT0000078	*	*			*	*	[2, 3]
hsa-miR-23a*	MIMAT0004496							
hsa-miR-23b	MIMAT0000418	*	*			*		[2-4]
hsa-miR-23b*	MIMAT0004587						*	
hsa-miR-24	MIMAT0000080	*						[2, 3, 6]
hsa-miR-24-1*	MIMAT0000079						*	
hsa-miR-24-2*	MIMAT0004497						*	
hsa-miR-25	MIMAT0000081	*	*					[3, 5]
hsa-miR-25*	MIMAT0004498						*	
hsa-miR-26a	MIMAT0000082	*				*		[2, 3, 5]
hsa-miR-26a-1*	MIMAT0004499						*	
hsa-miR-26a-2*	MIMAT0004681						*	
hsa-miR-26b	MIMAT0000083	*				*		[2, 3]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-26b*	MIMAT0004500						*	
hsa-miR-27a	MIMAT0000084	*	*		*			[2, 3, 5]
hsa-miR-27a*	MIMAT0004501				*			
hsa-miR-27b	MIMAT0000419	*	*					[2, 3]
hsa-miR-27b*	MIMAT0004588							
hsa-miR-28-3p	MIMAT0004502	*						
hsa-miR-28-5p	MIMAT0000085	*		*	*			
hsa-miR-296-3p	MIMAT0004679						*	
hsa-miR-296-5p	MIMAT0000690					*		
hsa-miR-298	MIMAT0004901					*		
hsa-miR-299-3p	MIMAT0000687							[3]
hsa-miR-299-5p	MIMAT0002890							[3]
hsa-miR-29a	MIMAT0000086	*			*			[3, 4]
hsa-miR-29a*	MIMAT0004503				*			
hsa-miR-29b	MIMAT0000100	*						[2-4]
hsa-miR-29b-1*	MIMAT0004514						*	
hsa-miR-29b-2*	MIMAT0004515						*	
hsa-miR-29c	MIMAT0000681	*			*			[3]
hsa-miR-29c*	MIMAT0004673				*			
hsa-miR-300	MIMAT0004903			*		*		
hsa-miR-301a	MIMAT0000688	*						

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-301b	MIMAT0004958	*						
hsa-miR-302a	MIMAT0000684	*				*		[3]
hsa-miR-302a*	MIMAT0000683							
hsa-miR-302b	MIMAT0000715	*						[3]
hsa-miR-302c	MIMAT0000717	*				*		[3]
hsa-miR-302c*	MIMAT0000716							
hsa-miR-302d	MIMAT0000718					*		[3]
hsa-miR-302d*	MIMAT0004685						*	
hsa-miR-302e	MIMAT0005931						*	
hsa-miR-302f	MIMAT0005932						*	
hsa-miR-30a	MIMAT0000087	*						
hsa-miR-30a*	MIMAT0000088						*	
hsa-miR-30b	MIMAT0000420	*		*	*			[2, 3, 5]
hsa-miR-30b*	MIMAT0004589							
hsa-miR-30c	MIMAT0000244	*						[2, 3, 5]
hsa-miR-30c-1*	MIMAT0004674						*	
hsa-miR-30c-2*	MIMAT0004550						*	
hsa-miR-30d	MIMAT0000245	*						[3]
hsa-miR-30d*	MIMAT0004551							
hsa-miR-30e	MIMAT0000692	*						
hsa-miR-31	MIMAT0000089		*					[1, 3, 4, 5, 7]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-31*	MIMAT0004504							
hsa-miR-32	MIMAT0000090	*				*		[3]
hsa-miR-32*	MIMAT0004505						*	
hsa-miR-320a	MIMAT0000510		*				*	
hsa-miR-320b	MIMAT0005792		*				*	
hsa-miR-320d	MIMAT0006764		*				*	
hsa-miR-323-3p	MIMAT0000755				*	*		
hsa-miR-323b-5p	MIMAT0001630							
hsa-miR-324-3p	MIMAT0000762					*		[3]
hsa-miR-324-5p	MIMAT0000761	*						[3, 5]
hsa-miR-325	MIMAT0000771		*					[5]
hsa-miR-326	MIMAT0000756							[3, 5]
hsa-miR-328	MIMAT0000752							[5, 7]
hsa-miR-329	MIMAT0001629				*	*		
hsa-miR-330-3p	MIMAT0000751					*		
hsa-miR-330-5p	MIMAT0004693						*	
hsa-miR-331-3p	MIMAT0000760	*						
hsa-miR-331-5p	MIMAT0004700						*	
hsa-miR-335	MIMAT0000765	*						[3, 5]
hsa-miR-335*	MIMAT0004703							
hsa-miR-337-3p	MIMAT0000754						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-337-5p	MIMAT0004695						*	
hsa-miR-338-3p	MIMAT0000763	*						
hsa-miR-338-5p	MIMAT0004701					*		
hsa-miR-339-3p	MIMAT0004702	*						
hsa-miR-339-5p	MIMAT0000764	*				*		
hsa-miR-33a	MIMAT0000091							
hsa-miR-33a*	MIMAT0004506						*	
hsa-miR-33b	MIMAT0003301						*	
hsa-miR-33b*	MIMAT0004811						*	
hsa-miR-340	MIMAT0004692					*		[5]
hsa-miR-340*	MIMAT0000750						*	
hsa-miR-342-3p	MIMAT0000753					*		
hsa-miR-342-5p	MIMAT0004694							
hsa-miR-345	MIMAT0000772	*				*		[3]
hsa-miR-346	MIMAT0000773						*	
hsa-miR-34a	MIMAT0000255	*			*			[3, 5]
hsa-miR-34a*	MIMAT0004557						*	
hsa-miR-34b	MIMAT0004676							[6]
hsa-miR-34b*	MIMAT0000685						*	
hsa-miR-34c-3p	MIMAT0004677					*		
hsa-miR-34c-5p	MIMAT0000686	*				*		

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-361-3p	MIMAT0004682						*	
hsa-miR-361-5p	MIMAT0000703	*				*		
hsa-miR-362-3p	MIMAT0004683					*		
hsa-miR-362-5p	MIMAT0000705					*		
hsa-miR-363	MIMAT0000707	*				*		[3]
hsa-miR-363*	MIMAT0003385						*	
hsa-miR-365	MIMAT0000710	*						[3]
hsa-miR-367	MIMAT0000719				*			[3]
hsa-miR-367*	MIMAT0004686							
hsa-miR-369-3p	MIMAT0000721							
hsa-miR-369-5p	MIMAT0001621						*	
hsa-miR-370	MIMAT0000722							[5]
hsa-miR-371-3p	MIMAT0000723						*	
hsa-miR-371-5p	MIMAT0004687						*	
hsa-miR-372	MIMAT0000724	*						[3]
hsa-miR-373	MIMAT0000726	*						[3]
hsa-miR-373*	MIMAT0000725			*	*			
hsa-miR-374a	MIMAT0000727	*			*	*		
hsa-miR-374a*	MIMAT0004688						*	
hsa-miR-374b	MIMAT0004955	*				*		
hsa-miR-374b*	MIMAT0004956						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-375	MIMAT0000728	*						[3]
hsa-miR-376a	MIMAT0000729	*			*	*	*	
hsa-miR-376a*	MIMAT0003386						*	
hsa-miR-376b	MIMAT0002172					*		
hsa-miR-376c	MIMAT0000720	*						
hsa-miR-377	MIMAT0000730					*		[3]
hsa-miR-377*	MIMAT0004689						*	
hsa-miR-378	MIMAT0000732	*						[3]
hsa-miR-378*	MIMAT0000731						*	
hsa-miR-379	MIMAT0000733	*				*		[3]
hsa-miR-379*	MIMAT0004690						*	
hsa-miR-380	MIMAT0000735					*		
hsa-miR-381	MIMAT0000736			*	*	*		[3]
hsa-miR-382	MIMAT0000737							[3]
hsa-miR-383	MIMAT0000738					*		
hsa-miR-384	MIMAT0001075						*	
hsa-miR-409-3p	MIMAT0001639							[3]
hsa-miR-409-5p	MIMAT0001638						*	
hsa-miR-410	MIMAT0002171	*						[3]
hsa-miR-411	MIMAT0003329	*				*		
hsa-miR-411*	MIMAT0004813						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-412	MIMAT0002170					*		
hsa-miR-421	MIMAT0003339	*				*		
hsa-miR-422b	MIMAT0001339						*	[3, 4]
hsa-miR-423-3p	MIMAT0001340	*						
hsa-miR-423-5p	MIMAT0004748	*						
hsa-miR-424	MIMAT0001341	*		*	*	*		[3]
hsa-miR-424*	MIMAT0004749				*			
hsa-miR-425	MIMAT0003393	*				*		[3]
hsa-miR-425*	MIMAT0001343							
hsa-miR-429	MIMAT0001536			*	*	*		[2, 3]
hsa-miR-431	MIMAT0001625	*						
hsa-miR-431*	MIMAT0004757						*	
hsa-miR-432	MIMAT0002814					*		
hsa-miR-432*	MIMAT0002815						*	
hsa-miR-433	MIMAT0001627				*	*		
hsa-miR-448	MIMAT0001532					*		
hsa-miR-449a	MIMAT0001541						*	
hsa-miR-449b	MIMAT0003327						*	
hsa-miR-450a	MIMAT0001545						*	
hsa-miR-450b-3p	MIMAT0004910						*	
hsa-miR-450b-5p	MIMAT0004909						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-451	MIMAT0001631							[3]
hsa-miR-452	MIMAT0001635	*						[3]
hsa-miR-452*	MIMAT0001636							
hsa-miR-454	MIMAT0003885	*						[3]
hsa-miR-454*	MIMAT0003884						*	
hsa-miR-455-3p	MIMAT0004784	*						
hsa-miR-455-5p	MIMAT0003150				*			
hsa-miR-483-3p	MIMAT0002173					*		
hsa-miR-483-5p	MIMAT0004761					*	*	
hsa-miR-484	MIMAT0002174	*						
hsa-miR-485-3p	MIMAT0002176						*	
hsa-miR-485-5p	MIMAT0002175						*	
hsa-miR-486-3p	MIMAT0004762						*	
hsa-miR-486-5p	MIMAT0002177	*						
hsa-miR-487a	MIMAT0002178					*		
hsa-miR-487b	MIMAT0003180						*	
hsa-miR-488	MIMAT0004763	*			*	*		
hsa-miR-488*	MIMAT0002804				*			
hsa-miR-489	MIMAT0002805					*		
hsa-miR-490-3p	MIMAT0002806						*	
hsa-miR-490-5p	MIMAT0004764						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-491-5p	MIMAT0002807						*	
hsa-miR-492	MIMAT0002812					*		
hsa-miR-493	MIMAT0003161					*		
hsa-miR-493*	MIMAT0002813						*	
hsa-miR-494	MIMAT0002816			*	*			[3]
hsa-miR-495	MIMAT0002817				*			[3]
hsa-miR-497	MIMAT0002820	*		*	*	*		[3]
hsa-miR-497*	MIMAT0004768				*			
hsa-miR-498	MIMAT0002824						*	[3]
hsa-miR-499-3p	MIMAT0004772						*	
hsa-miR-499-5p	MIMAT0002870						*	
hsa-miR-500a	MIMAT0004773							
hsa-miR-500a*	MIMAT0002871							
hsa-miR-501-3p	MIMAT0004774					*		
hsa-miR-501-5p	MIMAT0002872						*	
hsa-miR-502-3p	MIMAT0004775					*	*	
hsa-miR-502-5p	MIMAT0002873						*	
hsa-miR-503	MIMAT0002874	*				*		
hsa-miR-504	MIMAT0002875						*	
hsa-miR-505	MIMAT0002876	*			*	*		[3]
hsa-miR-505*	MIMAT0004776				*			

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-506	MIMAT0002878						*	
hsa-miR-507	MIMAT0002879					*		[3]
hsa-miR-508-3p	MIMAT0002880						*	
hsa-miR-509-3-5p	MIMAT0004975					*	*	
hsa-miR-509-3p	MIMAT0002881					*		
hsa-miR-509-5p	MIMAT0004779				*	*		
hsa-miR-510	MIMAT0002882					*		
hsa-miR-511	MIMAT0002808						*	
hsa-miR-512-3p	MIMAT0002823					*		[3]
hsa-miR-512-5p	MIMAT0002822						*	[3]
hsa-miR-513a-3p	MIMAT0004777					*	*	
hsa-miR-513a-5p	MIMAT0002877					*		
hsa-miR-513b	MIMAT0005788						*	
hsa-miR-513c	MIMAT0005789						*	
hsa-miR-514	MIMAT0002883							[3]
hsa-miR-515-3p	MIMAT0002827					*		
hsa-miR-515-5p	MIMAT0002826			*		*		
hsa-miR-516a-3p	MIMAT0006778							
hsa-miR-516b	MIMAT0002859					*	*	
hsa-miR-517*	MIMAT0002851						*	
hsa-miR-517a	MIMAT0002852	*						[3]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-517b	MIMAT0002857							
hsa-miR-517c	MIMAT0002866							[3]
hsa-miR-518a-3p	MIMAT0002863						*	
hsa-miR-518a-5p	MIMAT0005457					*		
hsa-miR-518b	MIMAT0002844	*						[3]
hsa-miR-518c	MIMAT0002848	*			*			
hsa-miR-518c*	MIMAT0002847						*	
hsa-miR-518e	MIMAT0002861				*		*	
hsa-miR-518e*	MIMAT0005450				*			
hsa-miR-518f	MIMAT0002842						*	
hsa-miR-518f*	MIMAT0002841							
hsa-miR-519b-3p	MIMAT0002837						*	
hsa-miR-519c-3p	MIMAT0002832					*		
hsa-miR-519d	MIMAT0002853					*	*	
hsa-miR-519e	MIMAT0002829						*	
hsa-miR-519e*	MIMAT0002828						*	
hsa-miR-520a-3p	MIMAT0002834						*	
hsa-miR-520a-5p	MIMAT0002833						*	
hsa-miR-520b	MIMAT0002843							
hsa-miR-520c-3p	MIMAT0002846						*	
hsa-miR-520d-3p	MIMAT0002856						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-520e	MIMAT0002825							
hsa-miR-520f	MIMAT0002830					*		
hsa-miR-520g	MIMAT0002858					*		[3]
hsa-miR-520h	MIMAT0002867					*		[3]
hsa-miR-521	MIMAT0002854					*		[3]
hsa-miR-522	MIMAT0002868					*		
hsa-miR-523	MIMAT0002840						*	
hsa-miR-524-3p	MIMAT0002850						*	
hsa-miR-525-3p	MIMAT0002839						*	
hsa-miR-525-5p	MIMAT0002838					*	*	
hsa-miR-526b	MIMAT0002835						*	
hsa-miR-526b*	MIMAT0002836						*	
hsa-miR-532-5p	MIMAT0002888	*		*	*	*	*	
hsa-miR-539	MIMAT0003163					*		
hsa-miR-541	MIMAT0004920						*	
hsa-miR-541*	MIMAT0004919						*	
hsa-miR-542-3p	MIMAT0003389						*	
hsa-miR-542-5p	MIMAT0003340						*	
hsa-miR-543	MIMAT0004954					*		
hsa-miR-544	MIMAT0003164						*	
hsa-miR-545	MIMAT0003165	*			*	*		

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-545*	MIMAT0004785				*			
hsa-miR-548a-3p	MIMAT0003251					*		
hsa-miR-548b-3p	MIMAT0003254					*		
hsa-miR-548c-3p	MIMAT0003285			*		*		
hsa-miR-548c-5p	MIMAT0004806	*						
hsa-miR-548d-3p	MIMAT0003323				*	*		
hsa-miR-548d-5p	MIMAT0004812						*	
hsa-miR-548e	MIMAT0005874			*				
hsa-miR-548f	MIMAT0005895						*	
hsa-miR-548g	MIMAT0005912						*	
hsa-miR-548h	MIMAT0005928						*	
hsa-miR-548i	MIMAT0005935						*	
hsa-miR-548j	MIMAT0005875						*	
hsa-miR-548k	MIMAT0005882						*	
hsa-miR-548l	MIMAT0005889						*	
hsa-miR-548m	MIMAT0005917						*	
hsa-miR-548n	MIMAT0005916			*			*	
hsa-miR-548o	MIMAT0005919						*	
hsa-miR-548p	MIMAT0005934						*	
hsa-miR-549	MIMAT0003333					*		
hsa-miR-550a	MIMAT0004800							

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-551a	MIMAT0003214						*	
hsa-miR-551b	MIMAT0003233						*	
hsa-miR-551b*	MIMAT0004794						*	
hsa-miR-552	MIMAT0003215					*		
hsa-miR-553	MIMAT0003216						*	
hsa-miR-554	MIMAT0003217						*	
hsa-miR-555	MIMAT0003219					*		
hsa-miR-556-3p	MIMAT0004793						*	
hsa-miR-556-5p	MIMAT0003220						*	
hsa-miR-557	MIMAT0003221						*	
hsa-miR-558	MIMAT0003222					*	*	
hsa-miR-559	MIMAT0003223						*	
hsa-miR-561	MIMAT0003225					*		
hsa-miR-563	MIMAT0003227						*	
hsa-miR-564	MIMAT0003228					*		
hsa-miR-566	MIMAT0003230						*	
hsa-miR-567	MIMAT0003231							
hsa-miR-568	MIMAT0003232						*	
hsa-miR-569	MIMAT0003234						*	[2]
hsa-miR-570	MIMAT0003235			*		*		
hsa-miR-571	MIMAT0003236						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-572	MIMAT0003237					*		
hsa-miR-573	MIMAT0003238					*		
hsa-miR-574-3p	MIMAT0003239					*		
hsa-miR-574-5p	MIMAT0004795					*		
hsa-miR-575	MIMAT0003240					*		
hsa-miR-576-5p	MIMAT0003241					*		
hsa-miR-577	MIMAT0003242			*				
hsa-miR-578	MIMAT0003243					*		
hsa-miR-579	MIMAT0003244					*		
hsa-miR-580	MIMAT0003245					*		
hsa-miR-582-3p	MIMAT0004797					*		
hsa-miR-582-5p	MIMAT0003247					*		
hsa-miR-584	MIMAT0003249					*		
hsa-miR-585	MIMAT0003250						*	
hsa-miR-586	MIMAT0003252					*		
hsa-miR-587	MIMAT0003253					*		
hsa-miR-588	MIMAT0003255					*		
hsa-miR-589	MIMAT0004799						*	
hsa-miR-589*	MIMAT0003256						*	
hsa-miR-590-3p	MIMAT0004801	*		*				
hsa-miR-590-5p	MIMAT0003258						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-591	MIMAT0003259						*	[2]
hsa-miR-592	MIMAT0003260					*		
hsa-miR-593	MIMAT0004802						*	
hsa-miR-593*	MIMAT0003261						*	
hsa-miR-595	MIMAT0003263					*		
hsa-miR-596	MIMAT0003264					*		
hsa-miR-597	MIMAT0003265					*		
hsa-miR-598	MIMAT0003266						*	
hsa-miR-599	MIMAT0003267						*	
hsa-miR-600	MIMAT0003268					*		
hsa-miR-601	MIMAT0003269							
hsa-miR-602	MIMAT0003270					*		
hsa-miR-603	MIMAT0003271				*	*		
hsa-miR-604	MIMAT0003272					*		
hsa-miR-605	MIMAT0003273					*		
hsa-miR-606	MIMAT0003274						*	
hsa-miR-607	MIMAT0003275			*		*		
hsa-miR-608	MIMAT0003276					*		
hsa-miR-609	MIMAT0003277					*		
hsa-miR-610	MIMAT0003278					*		
hsa-miR-611	MIMAT0003279					*		

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-612	MIMAT0003280					*		
hsa-miR-613	MIMAT0003281						*	
hsa-miR-614	MIMAT0003282						*	
hsa-miR-615-3p	MIMAT0003283						*	
hsa-miR-615-5p	MIMAT0004804						*	
hsa-miR-616*	MIMAT0003284						*	
hsa-miR-617	MIMAT0003286				*	*		
hsa-miR-618	MIMAT0003287					*		
hsa-miR-619	MIMAT0003288					*		
hsa-miR-620	MIMAT0003289					*		
hsa-miR-621	MIMAT0003290					*		
hsa-miR-622	MIMAT0003291						*	
hsa-miR-623	MIMAT0003292						*	
hsa-miR-624	MIMAT0004807					*		
hsa-miR-624*	MIMAT0003293						*	
hsa-miR-625	MIMAT0003294	*		*	*	*		
hsa-miR-625*	MIMAT0004808						*	
hsa-miR-626	MIMAT0003295						*	
hsa-miR-627	MIMAT0003296					*		
hsa-miR-628-3p	MIMAT0003297					*		
hsa-miR-628-5p	MIMAT0004809						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-629	MIMAT0004810	*						
hsa-miR-629*	MIMAT0003298							
hsa-miR-630	MIMAT0003299					*		
hsa-miR-631	MIMAT0003300					*		
hsa-miR-632	MIMAT0003302							
hsa-miR-633	MIMAT0003303						*	
hsa-miR-634	MIMAT0003304					*		
hsa-miR-635	MIMAT0003305					*		
hsa-miR-636	MIMAT0003306					*		
hsa-miR-637	MIMAT0003307					*		
hsa-miR-638	MIMAT0003308					*		
hsa-miR-639	MIMAT0003309						*	
hsa-miR-640	MIMAT0003310						*	
hsa-miR-641	MIMAT0003311			*		*		
hsa-miR-642a	MIMAT0003312							
hsa-miR-643	MIMAT0003313						*	
hsa-miR-644	MIMAT0003314					*		
hsa-miR-645	MIMAT0003315					*		
hsa-miR-646	MIMAT0003316					*		
hsa-miR-647	MIMAT0003317					*		
hsa-miR-648	MIMAT0003318						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-649	MIMAT0003319					*		
hsa-miR-650	MIMAT0003320						*	
hsa-miR-651	MIMAT0003321						*	
hsa-miR-652	MIMAT0003322	*						
hsa-miR-653	MIMAT0003328					*	*	
hsa-miR-654-3p	MIMAT0004814							
hsa-miR-654-5p	MIMAT0003330				*	*		
hsa-miR-655	MIMAT0003331				*	*		
hsa-miR-657	MIMAT0003335					*		
hsa-miR-658	MIMAT0003336						*	
hsa-miR-659	MIMAT0003337					*		
hsa-miR-660	MIMAT0003338					*		
hsa-miR-661	MIMAT0003324					*		
hsa-miR-662	MIMAT0003325						*	
hsa-miR-663	MIMAT0003326					*		
hsa-miR-663b	MIMAT0005867						*	
hsa-miR-664	MIMAT0005949						*	
hsa-miR-664*	MIMAT0005948						*	
hsa-miR-665	MIMAT0004952						*	
hsa-miR-668	MIMAT0003881					*		
hsa-miR-671-3p	MIMAT0004819						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-671-5p	MIMAT0003880					*		
hsa-miR-675	MIMAT0004284						*	
hsa-miR-7	MIMAT0000252	*	*			*		[3]
hsa-miR-708	MIMAT0004926					*		
hsa-miR-708*	MIMAT0004927						*	
hsa-miR-7-1*	MIMAT0004553		*				*	
hsa-miR-7-2*	MIMAT0004554		*				*	
hsa-miR-720	MIMAT0005954						*	
hsa-miR-744	MIMAT0004945	*			*			
hsa-miR-744*	MIMAT0004946				*			
hsa-miR-760	MIMAT0004957						*	
hsa-miR-765	MIMAT0003945					*		
hsa-miR-766	MIMAT0003888					*		
hsa-miR-767-3p	MIMAT0003883					*		
hsa-miR-767-5p	MIMAT0003882					*		
hsa-miR-769-3p	MIMAT0003887					*		
hsa-miR-769-5p	MIMAT0003886					*		
hsa-miR-770-5p	MIMAT0003948							
hsa-miR-802	MIMAT0004185						*	
hsa-miR-873	MIMAT0004953						*	
hsa-miR-874	MIMAT0004911					*		

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-875-3p	MIMAT0004923				*	*		
hsa-miR-875-5p	MIMAT0004922						*	
hsa-miR-876-3p	MIMAT0004925						*	
hsa-miR-876-5p	MIMAT0004924						*	
hsa-miR-877	MIMAT0004949	*						
hsa-miR-877*	MIMAT0004950						*	
hsa-miR-885-3p	MIMAT0004948						*	
hsa-miR-885-5p	MIMAT0004947						*	
hsa-miR-886-3p	MIMAT0004906						*	
hsa-miR-886-5p	MIMAT0004905						*	
hsa-miR-888	MIMAT0004916					*		
hsa-miR-888*	MIMAT0004917						*	
hsa-miR-889	MIMAT0004921						*	
hsa-miR-890	MIMAT0004912						*	
hsa-miR-891a	MIMAT0004902						*	
hsa-miR-891b	MIMAT0004913					*		
hsa-miR-892a	MIMAT0004907						*	
hsa-miR-892b	MIMAT0004918						*	
hsa-miR-9	MIMAT0000441	*				*		[6]
hsa-miR-9*	MIMAT0000442							[3, 6]
hsa-miR-920	MIMAT0004970						*	

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-922	MIMAT0004972					*		
hsa-miR-924	MIMAT0004974						*	
hsa-miR-92a	MIMAT0000092	*			*	*		
hsa-miR-92a-1*	MIMAT0004507						*	
hsa-miR-92a-2*	MIMAT0004508						*	
hsa-miR-92b	MIMAT0003218				*			
hsa-miR-92b*	MIMAT0004792				*			
hsa-miR-93	MIMAT0000093	*				*		[2, 3]
hsa-miR-93*	MIMAT0004509							
hsa-miR-933	MIMAT0004976						*	
hsa-miR-935	MIMAT0004978					*		
hsa-miR-936	MIMAT0004979						*	
hsa-miR-937	MIMAT0004980						*	
hsa-miR-938	MIMAT0004981						*	
hsa-miR-939	MIMAT0004982						*	
hsa-miR-940	MIMAT0004983						*	
hsa-miR-941	MIMAT0004984						*	
hsa-miR-942	MIMAT0004985					*		
hsa-miR-944	MIMAT0004987			*		*		
hsa-miR-95	MIMAT0000094		*		*			[3, 5]
hsa-miR-96	MIMAT0000095	*	*			*		[1, 3, 5, 7]

miRNA ID	Accession	miRNA PCR arrays commercially available	IPA	microRNA.org	Microcosm	GeneSet2miRNA	Microarray platforms	Colorectal microRNAome
hsa-miR-96*	MIMAT0004510						*	
hsa-miR-98	MIMAT0000096	*						[3]
hsa-miR-99a	MIMAT0000097	*						[5,3,6]
hsa-miR-99a*	MIMAT0004511						*	
hsa-miR-99b	MIMAT0000689		*		*			[3]
hsa-miR-99b*	MIMAT0004678						*	

Table 14 miRNA microarray design
This table shows the miRNAs that were identified from scientific papers (Intestinal microRNAome), the databases used for miRNA prediction and the miRNAs from commercial arrays that were the sources for the chip design.

miRNA ID	Accession
mmu-miR-217	MIMAT0000679
mmu-miR-30b*	MIMAT0004524
mmu-miR-483	MIMAT0004782
mmu-miR-107	MIMAT00000647
mmu-miR-216b	MIMAT0003729
mmu-miR-135a	MIMAT0000147
mmu-miR-27b	MIMAT0000126
mmu-miR-187	MIMAT0004782
mmu-miR-150	MIMAT0000160
mmu-miR-302d	MIMAT0003377

Table 15 Negative control miRNAs
miRNAs from *M muluscus* included on the miRNA design to monitor false positives.

References

1. Ogier-Denis, E., et al., *MicroRNAs and intestinal pathophysiology.* M S-Medecine Sciences, 2007. **23**(5): p. 509-14.

2. Sarver, A., et al., *Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states.* BMC Cancer, 2009. **9**(1): p. 401.

3. Cummins, J.M., et al., *The colorectal microRNAome.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(10): p. 3687-92.

4. Han, L., et al., *DNA methylation regulates MicroRNA expression.* Cancer Biology & Therapy, 2007. **6**(8): p. 1284-8.

5. Monzo, M., et al., *Overlapping expression of microRNAs in human embryonic colon and colorectal cancer.* Cell Research, 2008. **18**(8): p. 823-33.

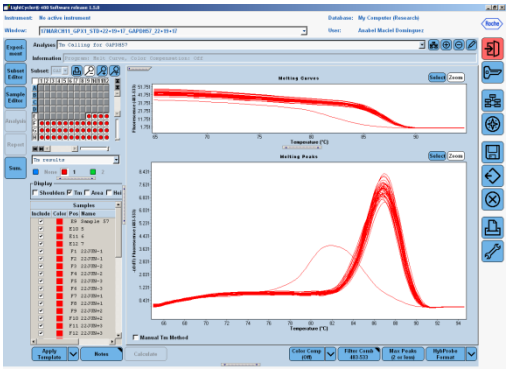
6. Peng, S., et al., *Multi-class cancer classification through gene expression profiles: microRNA versus mRNA.* Journal Of Genetics & Genomics = Yi Chuan Xue Bao, 2009. **36**(7): p. 409-16.

7. Bandres, E., et al., *Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues.* Molecular Cancer, 2006. **5**: p. 29.

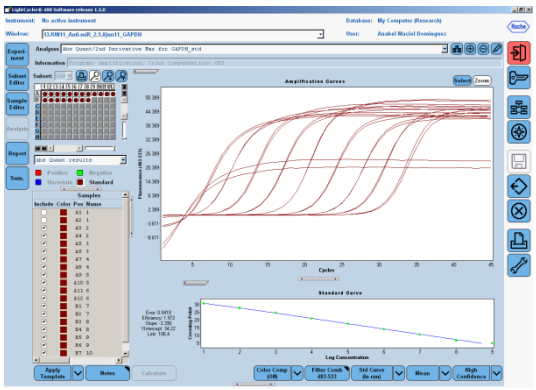
Appendix B QRT-PCR Melting curves and amplification curves

According to the guidelines for Minimum Information for Publication for Quantitative Real Time-PCR Experiments when using dyes like SYBR Green specificity has to be assessed by running the primer sequences thought BLAST on the NCBI website and by sunning a melting curve analysis on the amplification system used (LightCycler for this study); the efficiency by means of serial dilutions to produce a standard curve which would also ensure reproducible results (Bustin *et al.*, 2009).

A



B



C

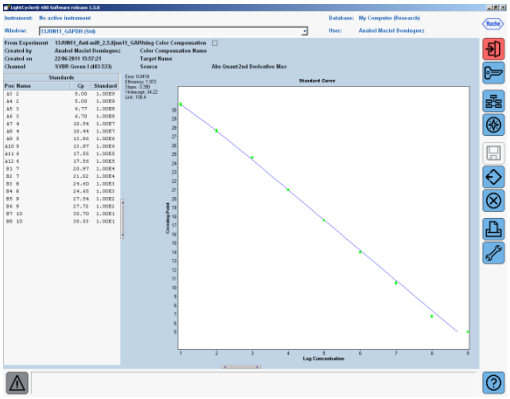
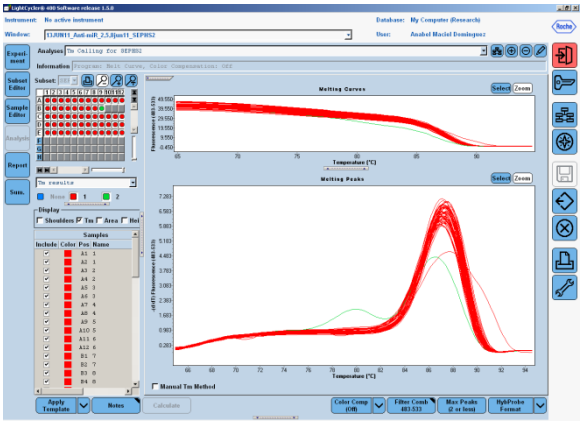
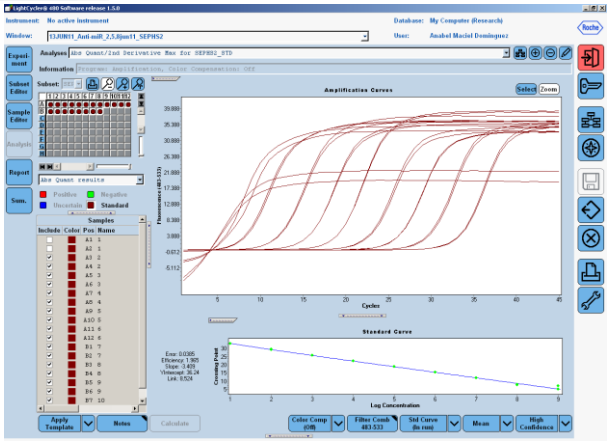


Figure 51 QRT-PCR analysis in the LightCycler480 of *GAPDH*.
Samples were amplified using GAPDH primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *GAPDH* was 1.972.

A



B



C

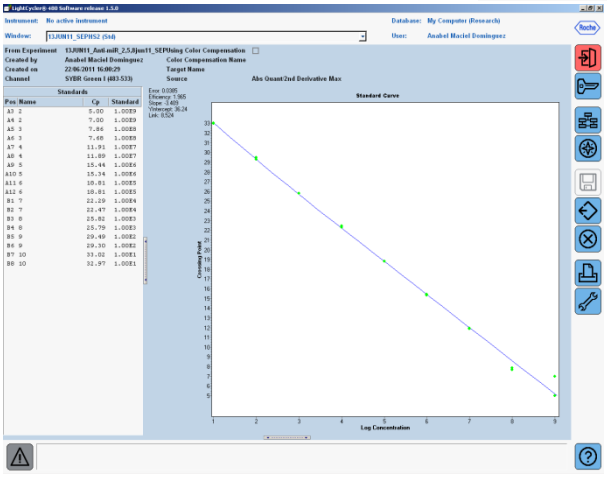
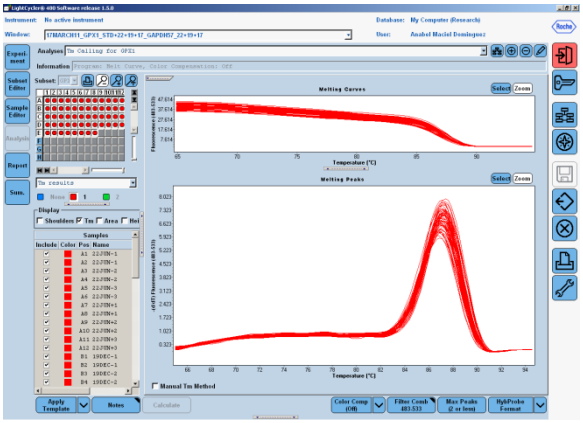
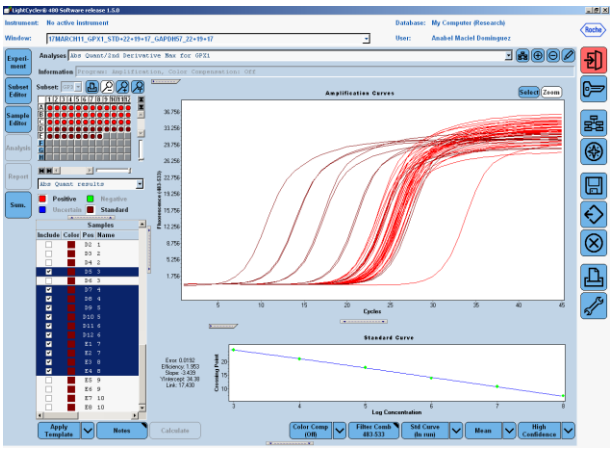


Figure 52 QRT-PCR analysis in the LightCycler480 of *SEPHS2*.
Samples were amplified using *SEPHS2* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *SEPHS2* was 1.965.

A



B



C

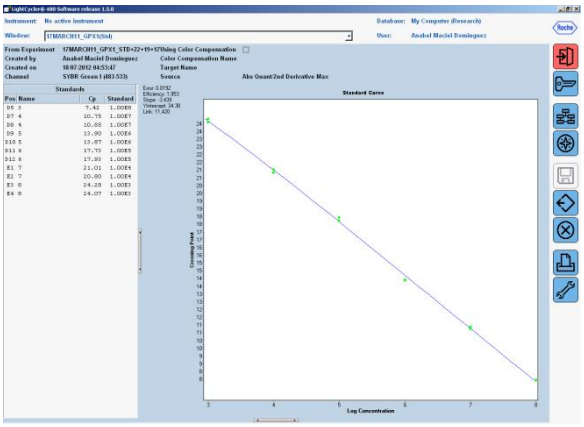
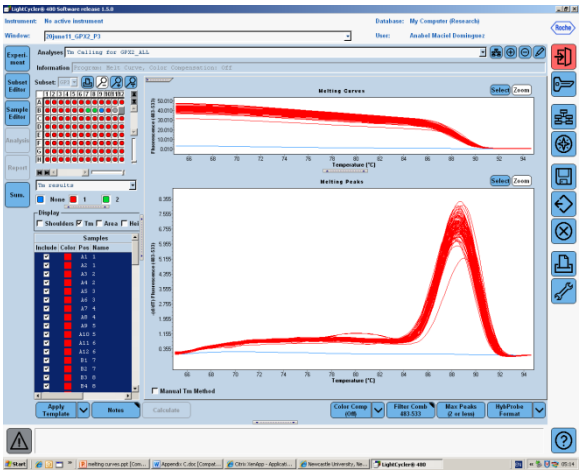
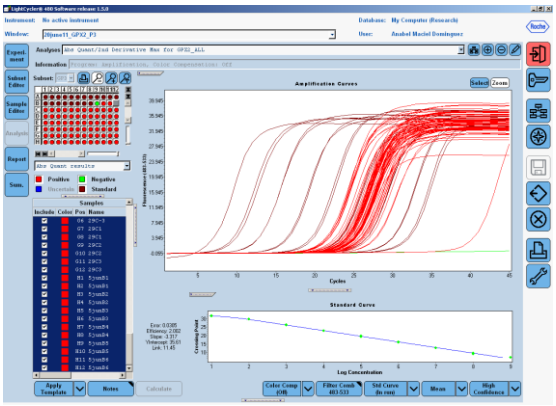


Figure 53 QRT-PCR analysis in the LightCycler480 of *GPX1*.
Samples were amplified using *GPX1* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *GPX1* was 1.953.

A



B



C

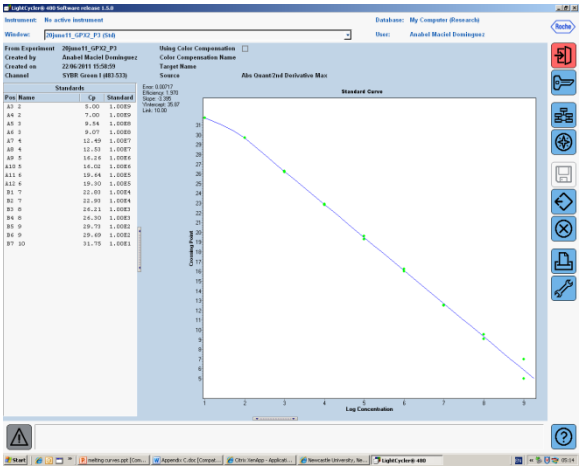
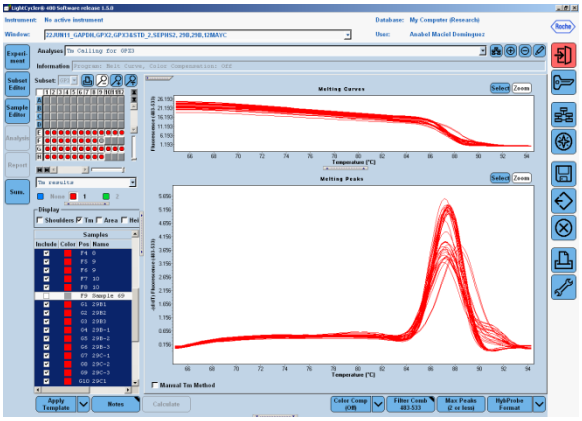
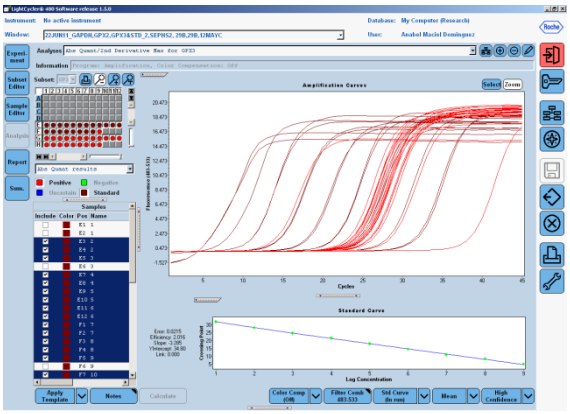


Figure 54 QRT-PCR analysis in the LightCycler480 of *GPX2*.
Samples were amplified using *GPX2* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *GPX2* was 2.002.

A



B



C

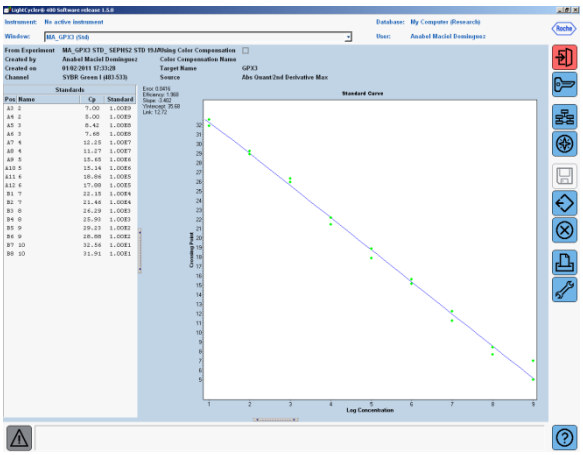
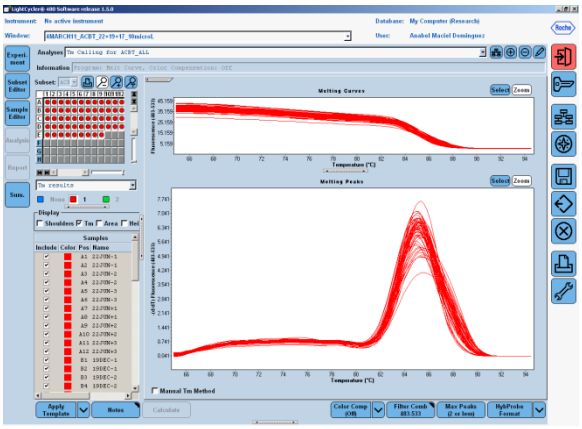
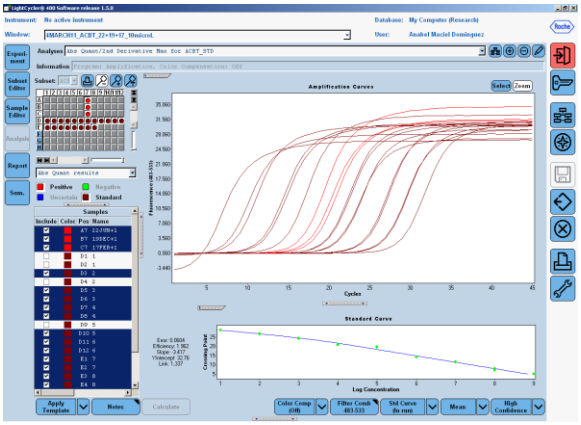


Figure 55 QRT-PCR analysis in the LightCycler480 of *GPX3*.
Samples were amplified using *GPX3* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *GPX3* was 1.968.

A



B



C

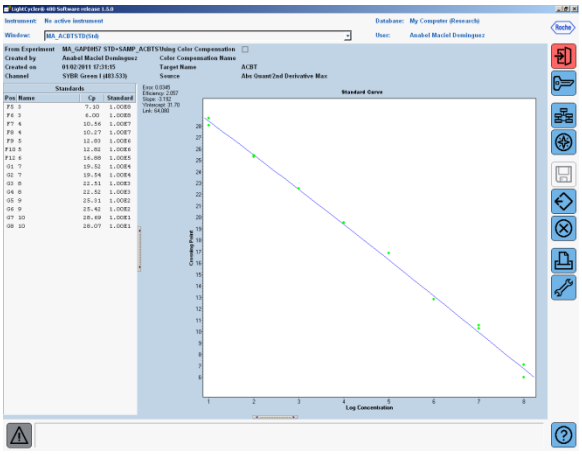
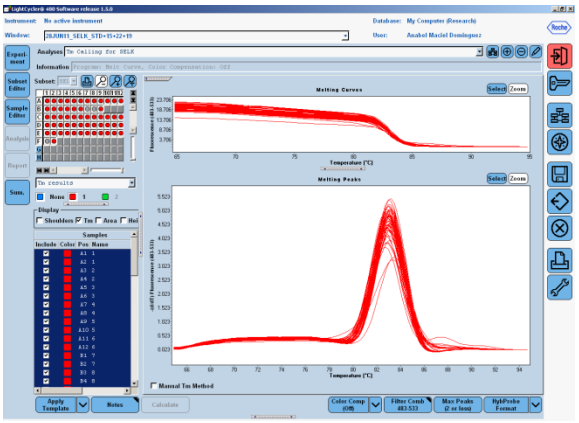
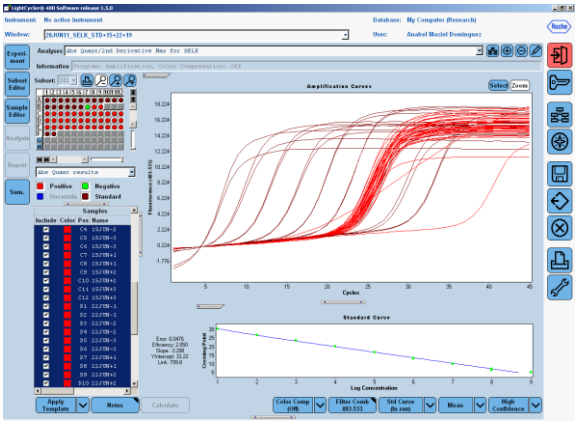


Figure 56 QRT-PCR analysis in the LightCycler480 of *ACBT*.
Samples were amplified using *ACBT* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *ACBT* was 2.057.

A



B



C

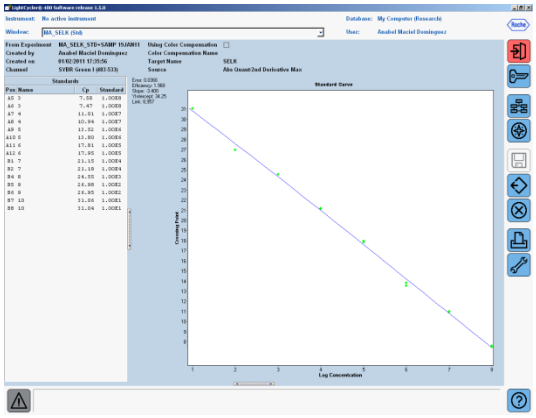
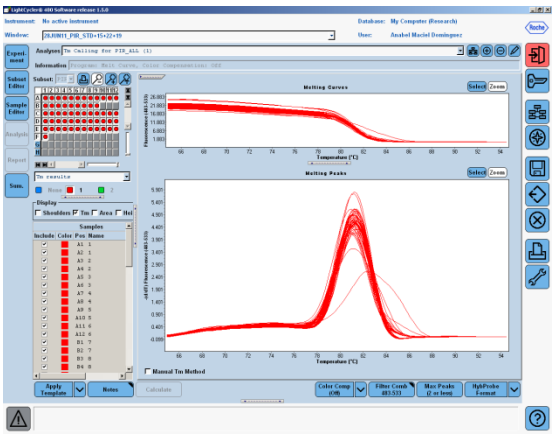
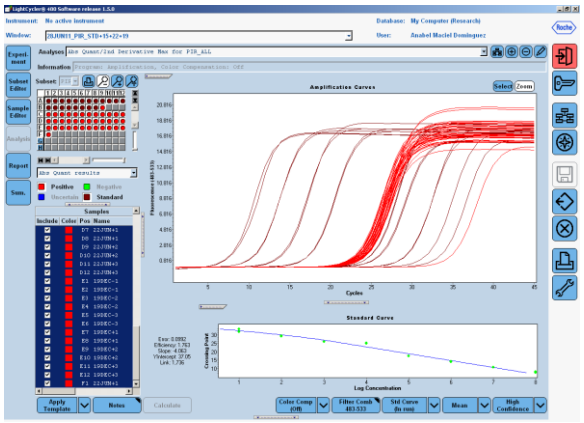


Figure 57 QRT-PCR analysis in the LightCycler480 of *SELK*.
Samples were amplified using *SELK* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *SELK* was 1.968.

A



B



C

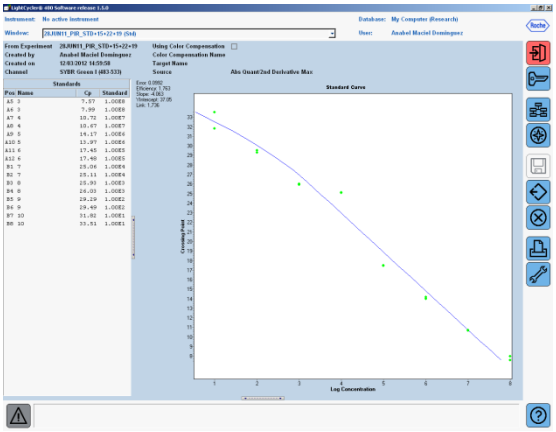


Figure 58 QRT-PCR analysis in the LightCycler480 of *PIR*.
Samples were amplified using *PIR* primers and with a fluorescent dye. Product accumulation was monitored by the LightCycler. A) PCR products generating a melting curve showed only one peak meaning sequence specificity; a different peak appears when a sample is over diluted samples or it is produced by the negative control. B) a ten-fold serial dilution is produced from PCR amplified cDNA in order to calculate C) the standard curve, from which efficiency of the enzymatic amplification is calculated and it is also used as external control to compare the expression of different experiments amplified by the same set of primers. Efficiency of *PIR* was 1.763.

7 References

Akcakaya, P., Ekelund, S., Kolosenko, I., Caramuta, S., Ozata, D.M., Xie, H., Lindforss, U., Olivecrona, H. and Lui, W.-O. (2011) 'miR-185 and miR-133b deregulation is associated with overall survival and metastasis in colorectal cancer', *International Journal of Oncology*, 39(2), pp. 311-8.

Almeida, M.I., Nicoloso, M.S., Zeng, L., Ivan, C., Spizzo, R., Gafa, R., Xiao, L., Zhang, X., Vannini, I., Fanini, F., Fabbri, M., Lanza, G., Reis, R.M., Zweidler-McKay, P.A. and Calin, G.A. (2012) 'Strand-specific miR-28-5p and miR-28-3p have distinct effects in colorectal cancer cells', *Gastroenterology*, 142(4), pp. 886-896.e9.

Antonov, A.V., Dietmann, S., Wong, P., Lutter, D. and Mewes, H.W. (2009) 'GeneSet2miRNA: finding the signature of cooperative miRNA activities in the gene lists', *Nucleic Acids Research*, 37(1), pp. W323–W328.

Arikan, D.C., Coskun, A., Ozer, A., Kilinc, M., Atalay, F. and Arikan, T. (2011) 'Plasma selenium, zinc, copper and lipid levels in postmenopausal Turkish women and their relation with osteoporosis', *Biological Trace Element Research*, 144(1-3), pp. 407-17.

Arthur, J.R., McKenzie, R.C. and Beckett, G.J. (2003) 'Selenium in the immune system', *Journal of Nutrition*, 133(5 Suppl 1), pp. 1457S-9S.

Azevedo, M.F., Barra, G.B., Naves, L.A., Ribeiro Velasco, L.F., Godoy Garcia Castro, P., de Castro, L.C., Amato, A.A., Miniard, A., Driscoll, D., Schomburg, L. and de Assis Rocha Neves, F. (2010) 'Selenoprotein-related disease in a young girl caused by nonsense mutations in the SBP2 gene', *J Clin Endocrinol Metab*, 95(8), pp. 4066-71.

Babbitt, C.C., Warner, L.R., Fedrigo, O., Wall, C.E. and Wray, G.A. (2011) 'Genomic signatures of diet-related shifts during human origins', *Proceedings of the Royal Society of London - Series B: Biological Sciences*, 278(1708), pp. 961-9.

Bandres, E., Cubedo, E., Agirre, X., Malumbres, R., Zarate, R., Ramirez, N., Abajo, A., Navarro, A., Moreno, I., Monzo, M. and Garcia-Foncillas, J. (2006) 'Identification by Real-time PCR of 13 mature

microRNAs differentially expressed in colorectal cancer and non-tumoral tissues', *Molecular Cancer*, 5, p. 29.

Banerjee, S., Yang, S. and Foster, C.B. (2012) 'A luciferase reporter assay to investigate the differential selenium-dependent stability of selenoprotein mRNAs', *The Journal of Nutritional Biochemistry*, 23(10), pp. 1294-1301.

Bao, B., Ali, S., Kong, D., Sarkar, S.H., Wang, Z., Banerjee, S., Aboukameel, A., Padhye, S., Philip, P.A. and Sarkar, F.H. (2011) 'Anti-tumor activity of a novel compound-CDF is mediated by regulating miR-21, miR-200, and PTEN in pancreatic cancer', *PLoS ONE [Electronic Resource]*, 6(3), p. e17850.

Barnes, K.M., Evenson, J.K., Raines, A.M. and Sunde, R.A. (2009) 'Transcript analysis of the selenoproteome indicates that dietary selenium requirements of rats based on selenium-regulated selenoprotein mRNA levels are uniformly less than those based on glutathione peroxidase activity', *Journal of Nutrition*, 139(2), pp. 199-206.

Barouki, R., Aggerbeck, M., Aggerbeck, L. and Coumoul, X. (2012) 'The aryl hydrocarbon receptor system', *Drug Metabol Drug Interact*, 27(1), pp. 3-8.

Barrett, L., Fletcher, S. and Wilton, S. (2012) 'Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements', *Cellular and Molecular Life Sciences*, pp. 1-22.

Bartel, D.P. (2004) 'MicroRNAs: genomics, biogenesis, mechanism, and function', *Cell*, 116(2), pp. 281-97.

Bartel, D.P. (2009) 'MicroRNAs: target recognition and regulatory functions', *Cell*, 136(2), pp. 215-33.

Bellinger, F.P., Raman, A.V., Reeves, M.A. and Berry, M.J. (2009) 'Regulation and function of selenoproteins in human disease', *Biochemical Journal*, 422(1), pp. 11-22.

Bellini, M.F., Cadamuro, A.C., Succi, M., Proenca, M.A. and Silva, A.E. (2012) 'Alterations of the TP53 gene in gastric and esophageal carcinogenesis', *J Biomed Biotechnol*, 2012, p. 891961.

Benes, V. and Castoldi, M. (2010) 'Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available', *Methods*, 50(4), pp. 244-9.

Bentley, A.R., Emrani, P. and Cassano, P.A. (2008) 'Genetic variation and gene expression in antioxidant related enzymes and risk of COPD: a systematic review', *Thorax*, 63(11), pp. 956-61.

Betel, D., Wilson, M., Gabow, A., Marks, D.S. and Sander, C. (2008) 'The microRNA.org resource: targets and expression', *Nucleic Acids Research*, 36(Database issue), pp. D149-53.

Bosse, A.C., Pallauf, J., Hommel, B., Sturm, M., Fischer, S., Wolf, N.M. and Mueller, A.S. (2009) 'Impact of selenite and selenate on differentially expressed genes in rat liver examined by microarray analysis', *Bioscience Reports*, 30(5), pp. 293-306.

Brigelius-Flohe, R. and Kipp, A. (2009) 'Glutathione peroxidases in different stages of carcinogenesis', *Biochimica et Biophysica Acta*, 1790(11), pp. 1555-68.

Brodersen, P. and Voinnet, O. (2009) 'Revisiting the principles of microRNA target recognition and mode of action', *Nature Reviews Molecular Cell Biology*, 10(2), pp. 141-8.

Brown, K.M. and Arthur, J.R. (2001) 'Selenium, selenoproteins and human health: a review', *Public Health Nutrition*, 4(2B), pp. 593-9.

Buermans, H., Ariyurek, Y., van Ommen, G., den Dunnen, J. and t Hoen, P. (2010) 'New methods for next generation sequencing based microRNA expression profiling', *BMC Genomics*, 11(1), p. 716.

Burk R. F., L.O.A. (1994) 'Selenium', in Wilkins, L.W. (ed.) *Modern Nutrition in Health and Disease* Ninth edn. London, pp. 265-274.

Burk, R.F. and Hill, K.E. (2009) 'Selenoprotein P-expression, functions, and roles in mammals', *Biochim Biophys Acta*, 11(7), p. 1.

Burk, R.F., Norsworthy, B.K., Hill, K.E., Motley, A.K. and Byrne, D.W. (2006) 'Effects of chemical form of selenium on plasma biomarkers in a high-dose human supplementation trial', *Cancer Epidemiology, Biomarkers & Prevention*, 15(4), pp. 804-10.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. and Wittwer, C.T. (2009) 'The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments', *Clinical Chemistry*, 55(4), pp. 611-22.

Caban, K. and Copeland, P.R. (2012) 'Selenocysteine insertion sequence (SECIS)-binding protein 2 alters conformational dynamics of residues involved in tRNA accommodation in 80 S ribosomes', *J Biol Chem*, 287(13), pp. 10664-73.

Cabodi, S., Camacho-Leal, M.d.P., Di Stefano, P. and Defilippi, P. (2010) 'Integrin signalling adaptors: not only figurants in the cancer story', *Nat Rev Cancer*, 10(12), pp. 858-870.

Cannell, I.G. and Bushell, M. (2010) 'Regulation of Myc by miR-34c: A mechanism to prevent genomic instability?', *Cell Cycle*, 9(14), pp. 2726-30.

Cannell, I.G., Kong, Y.W. and Bushell, M. (2008) 'How do microRNAs regulate gene expression?', *Biochemical Society Transactions*, 36(Pt 6), pp. 1224-31.

Carthew, R.W. and Sontheimer, E.J. (2009) 'Origins and Mechanisms of miRNAs and siRNAs', *Cell*, 136(4), pp. 642-55.

Casar-Borota, O., Fougner, S.L., Bollerslev, J. and Nesland, J.M. (2012) 'KIT protein expression and mutational status of KIT gene in pituitary adenomas', *Virchows Archiv*, 460(2), pp. 171-81.

Castañeda, J., Genzor, P. and Bortvin, A. (2011) 'piRNAs, transposon silencing, and germline genome integrity', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 714(1–2), pp. 95-104.

Castoldi, M., Schmidt, S., Benes, V., Noerholm, M., Kulozik, A.E., Hentze, M.W. and Muckenthaler, M.U. (2006) 'A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA)', *Rna-A Publication of the Rna Society*, 12(5), pp. 913-20.

Chang, K.H., Mestdagh, P., Vandesompele, J., Kerin, M.J. and Miller, N. (2010) 'MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer', *BMC Cancer*, 10, p. 173.

Chang, K.H., Miller, N., Kheirelseid, E.A., Lemetre, C., Ball, G.R., Smith, M.J., Regan, M., McAnena, O.J. and Kerin, M.J. (2011) 'MicroRNA signature analysis in colorectal cancer: identification of expression profiles in stage II tumors associated with aggressive disease', *Int J Colorectal Dis*, 26(11), pp. 1415-22.

Chen, B., Rao, X., House, M.G., Nephew, K.P., Cullen, K.J. and Guo, Z. (2011a) 'GPx3 promoter hypermethylation is a frequent event in human cancer and is associated with tumorigenesis and chemotherapy response', *Cancer Letters*, 309(1), pp. 37-45.

Chen, J., Chu, Y., Cao, J., Wang, W., Liu, J. and Wang, J. (2011b) 'Effects of T-2 toxin and selenium on chondrocyte expression of matrix metalloproteinases (MMP-1, MMP-13), alpha2-macroglobulin (alpha2M) and TIMPs', *Toxicol In Vitro*, 25(2), pp. 492-9.

Chen, K., Song, F., Calin, G.A., Wei, Q., Hao, X. and Zhang, W. (2008) 'Polymorphisms in microRNA targets: a gold mine for molecular epidemiology', *Carcinogenesis*, 29(7), pp. 1306-11.

Chen, X.J., Duan, F.D., Zhang, H.H., Xiong, Y. and Wang, J. (2012) 'Sodium selenite-induced apoptosis mediated by ROS attack in human osteosarcoma U2OS cells', *Biol Trace Elem Res*, 145(1), pp. 1-9.

Cheng, M.-L., Shiao, M.-S., Chiu, D.T.-Y., Weng, S.-F., Tang, H.-Y. and Ho, H.-Y. (2011) 'Biochemical disorders associated with antiproliferative effect of dehydroepiandrosterone in hepatoma cells as revealed by LC-based metabolomics', *Biochemical Pharmacology*, 82(11), pp. 1549-1561.

Chiang, Y., Song, Y., Wang, Z., Chen, Y., Yue, Z., Xu, H., Xing, C. and Liu, Z. (2011) 'Aberrant expression of miR-203 and its clinical significance in gastric and colorectal cancers', *Journal of Gastrointestinal Surgery*, 15(1), pp. 63-70.

Choudhuri, S. (2010) 'Small noncoding RNAs: Biogenesis, function, and emerging significance in toxicology', *Journal of Biochemical and Molecular Toxicology*, 24(3), pp. 195-216.

Cicenas, J. and Valius, M. (2011) 'The CDK inhibitors in cancer research and therapy', *J Cancer Res Clin Oncol*, 137(10), pp. 1409-18.

Cole-Ezea, P., Swan, D., Shanley, D. and Hesketh, J. (2012) 'Glutathione peroxidase 4 has a major role in protecting mitochondria from oxidative damage and maintaining oxidative phosphorylation complexes in gut epithelial cells', *Free Radical Biology and Medicine*, 53(3), pp. 488-497.

Combs, G.F., Jr., Jackson, M.I., Watts, J.C., Johnson, L.K., Zeng, H., Idso, J., Schomburg, L., Hoeg, A., Hoefig, C.S., Chiang, E.C., Waters,

D.J., Davis, C.D. and Milner, J.A. (2012) 'Differential responses to selenomethionine supplementation by sex and genotype in healthy adults', *Br J Nutr*, 107(10), pp. 1514-25.

Connelly-Frost, A., Poole, C., Satia, J.A., Kupper, L.L., Millikan, R.C. and Sandler, R.S. (2009) 'Selenium, folate, and colon cancer', *Nutrition & Cancer*, 61(2), pp. 165-78.

Corcoran, D.L., Pandit, K.V., Gordon, B., Bhattacharjee, A., Kaminski, N. and Benos, P.V. (2009) 'Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data', *PLoS ONE [Electronic Resource]*, 4(4), p. e5279.

Crane, M.S., Howie, A.F., Arthur, J.R., Nicol, F., Crosley, L.K. and Beckett, G.J. (2009) 'Modulation of thioredoxin reductase-2 expression in EAhy926 cells: implications for endothelial selenoprotein hierarchy', *Biochim Biophys Acta*, 10(7), p. 10.

Cui, J., Zhong, R., Chu, E., Zhang, X.F., Zhang, W.G., Fang, C.F., Dong, Q., Li, F.L. and Li, H. (2012) 'Correlation between Oxidative Stress and L-type Calcium Channel Expression in the Ventricular Myocardia of Selenium-deficient Mice', *J Int Med Res*, 40(5), pp. 1677-87.

Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E., Raymond, C.K., Roberts, B.S., Juhl, H., Kinzler, K.W., Vogelstein, B. and Velculescu, V.E. (2006) 'The colorectal microRNAome', *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), pp. 3687-92.

De Felice, B., Garbi, C., Wilson, R.R., Santoriello, M. and Nacca, M. (2011) 'Effect of selenocystine on gene expression profiles in human keloid fibroblasts', *Genomics*, 97(5), pp. 265-76.

de Rosa, V., Erkekoglu, P., Forestier, A., Favier, A., Hincal, F., Diamond, A.M., Douki, T. and Rachidi, W. (2012) 'Low doses of selenium specifically stimulate the repair of oxidative DNA damage in LNCaP prostate cancer cells', *Free Radic Res*, 46(2), pp. 105-16.

Djuranovic, S., Nahvi, A. and Green, R. (2011) 'A Parsimonious Model for Gene Regulation by miRNAs', *Science*, 331(6017), pp. 550-553.

Downey, C.M., Horton, C.R., Carlson, B.A., Parsons, T.E., Hatfield, D.L., Hallgrimsson, B. and Jirik, F.R. (2009) 'Osteo-

chondroprogenitor-specific deletion of the selenocysteine tRNA gene, Trsp, leads to chondronecrosis and abnormal skeletal development: a putative model for Kashin-Beck disease', *PLoS Genet*, 5(8), p. 21.

Driscoll, D.M. and Copeland, P.R. (2003) 'Mechanism and regulation of selenoprotein synthesis', *Annu Rev Nutr*, 23, pp. 17-40.

Dweep, H., Sticht, C., Pandey, P. and Gretz, N. (2011) 'miRWalk -- Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes', *Journal of Biomedical Informatics*, 44(5), pp. 839-847.

Dykxhoorn, D.M. (2010) 'MicroRNAs and metastasis: little RNAs go a long way', *Cancer Research*, 70(16), pp. 6401-6.

Fairweather-Tait, S.J., Bao, Y., Broadley, M.R., Collings, R., Ford, D., Hesketh, J.E. and Hurst, R. (2011) 'Selenium in human health and disease', *Antioxid Redox Signal*, 14(7), pp. 1337-83.

Falcon, S. and Gentleman, R. (2007) 'Using GOstats to test gene lists for GO term association', *Bioinformatics*, 23(2), pp. 257-258.

Fawzi, W., Msamanga, G., Spiegelman, D. and Hunter, D.J. (2005) 'Studies of vitamins and minerals and HIV transmission and disease progression', *Journal of Nutrition*, 135(4), pp. 938-44.

Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M. and Danielsen, M. (1987) 'Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure', *Proceedings of the National Academy of Sciences*, 84(21), pp. 7413-7417.

Felice, F., Lucchesi, D., di Stefano, R., Barsotti, M.C., Storti, E., Penno, G., Balbarini, A., Del Prato, S. and Pucci, L. (2010) 'Oxidative stress in response to high glucose levels in endothelial cells and in endothelial progenitor cells: evidence for differential glutathione peroxidase-1 expression', *Microvascular Research*, 80(3), pp. 332-8.

Ferguson, L.R., Karunasinghe, N., Zhu, S. and Wang, A.H. (2012) 'Selenium and its' role in the maintenance of genomic stability', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 733(1–2), pp. 100-110.

Filipowicz, W., Bhattacharyya, S.N. and Sonenberg, N. (2008) 'Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?', *Nature Reviews Genetics*, 9(2), pp. 102-14.

Finley, J.W. (2006) 'Bioavailability of selenium from foods', *Nutrition Reviews*, 64(3), pp. 146-51.

Fletcher, J.E., Copeland, P.R., Driscoll, D.M. and Krol, A. (2001) 'The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2', *Rna-A Publication of the Rna Society*, 7(10), pp. 1442-53.

Flohe, L. (2007) 'Selenium in mammalian spermiogenesis', *Biological Chemistry*, 388(10), pp. 987-95.

Florian, S., Krehl, S., Loewinger, M., Kipp, A., Banning, A., Esworthy, S., Chu, F.-F. and Brigelius-Flohe, R. (2010) 'Loss of GPx2 increases apoptosis, mitosis, and GPx1 expression in the intestine of mice', *Free Radical Biology & Medicine*, 49(11), pp. 1694-702.

Flynt, A.S. and Lai, E.C. (2008) 'Biological principles of microRNA-mediated regulation: shared themes amid diversity', *Nature Reviews Genetics*, 9(11), pp. 831-42.

Fordyce, F.M., Brereton, N., Hughes, J., Luo, W. and Lewis, J. (2010) 'An initial study to assess the use of geological parent materials to predict the Se concentration in overlying soils and in five staple foodstuffs produced on them in Scotland', *Sci Total Environ*, 408(22), pp. 5295-305.

Forman, H.J., Zhang, H. and Rinna, A. (2009) 'Glutathione: overview of its protective roles, measurement, and biosynthesis', *Mol Aspects Med*, 30(1-2), pp. 1-12.

Fradejas, N., Serrano-Perez Mdel, C., Tranque, P. and Calvo, S. (2011) 'Selenoprotein S expression in reactive astrocytes following brain injury', *Glia*, 59(6), pp. 959-72.

Francis, N.J. (2011) 'Gene regulation: implications of histone dispersal patterns for epigenetics', *Current Biology*, 21(17), pp. R659-61.

Franco-Zorrilla, J.M., Del Toro, F.J., Godoy, M., Perez-Perez, J., Lopez-Vidriero, I., Oliveros, J.C., Garcia-Casado, G., Llave, C. and Solano, R. (2009) 'Genome-wide identification of small RNA targets based on target enrichment and microarray hybridizations', *Plant J*, 59(5), pp. 840-50.

Gabory, A., Attig, L. and Junien, C. (2011) 'Developmental programming and epigenetics', *American Journal of Clinical Nutrition*, 94(6 Suppl), pp. 1943S-1952S.

Gambari, R., Fabbri, E., Borgatti, M., Lampronti, I., Finotti, A., Brognara, E., Bianchi, N., Manicardi, A., Marchelli, R. and Corradini, R. (2011) 'Targeting microRNAs involved in human diseases: a novel approach for modification of gene expression and drug development', *Biochemical Pharmacology*, 82(10), pp. 1416-29.

Gammelgaard, B., Rasmussen, L.H., Gabel-Jensen, C. and Steffansen, B. (2012a) 'Estimating intestinal absorption of inorganic and organic selenium compounds by in vitro flux and biotransformation studies in Caco-2 cells and ICP-MS detection', *Biol Trace Elem Res*, 145(2), pp. 248-56.

Gammelgaard, B., Sturup, S. and Christensen, M.V. (2012b) 'Human urinary excretion and metabolism of (82)Se-enriched selenite and selenate determined by LC-ICP-MS', *Metallomics*, 4(2), pp. 149-55.

Gardaneh, M., Gholami, M. and Maghsoudi, N. (2011) 'Synergy between glutathione peroxidase-1 and astrocytic growth factors suppresses free radical generation and protects dopaminergic neurons against 6-hydroxydopamine', *Rejuvenation Research*, 14(2), pp. 195-204.

Gautrey, H., Nicol, F., Sneddon, A.A., Hall, J. and Hesketh, J. (2011) 'A T/C polymorphism in the GPX4 3'UTR affects the selenoprotein expression pattern and cell viability in transfected Caco-2 cells.[Erratum appears in Biochim Biophys Acta. 2011 Jun;1810(6):iii]', *Biochimica et Biophysica Acta*, 1810(6), pp. 584-91.

Gaziel-Sovran, A., Segura, M.F., Di Micco, R., Collins, M.K., Hanniford, D., Vega-Saenz de Miera, E., Rakus, J.F., Dankert, J.F., Shang, S., Kerbel, R.S., Bhardwaj, N., Shao, Y., Darvishian, F., Zavadil, J., Erlebacher, A., Mahal, L.K., Osman, I. and Hernando, E. (2011) 'miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis', *Cancer Cell*, 20(1), pp. 104-18.

Gee, H.E., Buffa, F.M., Camps, C., Ramachandran, A., Leek, R., Taylor, M., Patil, M., Sheldon, H., Betts, G., Homer, J., West, C., Ragoussis, J. and Harris, A.L. (2011) 'The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis', *British Journal of Cancer*, 104(7), pp. 1168-77.

Ghadi, F.E., Ghara, A.R., Bhattacharyya, S. and Dhawan, D.K. (2009) 'Selenium as a chemopreventive agent in experimentally induced colon carcinogenesis', *World J Gastrointest Oncol.*, 1(1), pp. 74–81.

Ghildiyal, M. and Zamore, P.D. (2009) 'Small silencing RNAs: an expanding universe', *Nature Reviews Genetics*, 10(2), pp. 94-108.

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P. and Caldas, C. (2010) 'Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression', *Rna-A Publication of the Rna Society*, 16(5), pp. 991-1006.

Goldson, A.J., Fairweather-Tait, S.J., Armah, C.N., Bao, Y., Broadley, M.R., Dainty, J.R., Furniss, C., Hart, D.J., Teucher, B. and Hurst, R. (2011) 'Effects of Selenium Supplementation on Selenoprotein Gene Expression and Response to Influenza Vaccine Challenge: A Randomised Controlled Trial', *PLoS One*, 6(3), p. e14771.

Gong, H., Zuliani, P., Komuravelli, A., Faeder, J.R. and Clarke, E.M. (2010) 'Analysis and verification of the HMGB1 signaling pathway', *BMC Bioinformatics*, 11 Suppl 7, p. S10.

Griffiths-Jones, S., Saini, H.K., van Dongen, S. and Enright, A.J. (2008) 'miRBase: tools for microRNA genomics', *Nucl. Acids Res.*, 36(suppl_1), pp. D154-158.

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. (2007) 'MicroRNA targeting specificity in mammals: determinants beyond seed pairing', *Molecular Cell*, 27(1), pp. 91-105.

Grocock, R., van Dongen, S., Enright, A.J., Griffiths-Jones, S., Marks, D., Lab, S. and Giraldez, A. 'MicroCosm Targets (formerly miRBase Targets) is a web resource developed by the Enright Lab at the EMBL-EBI containing computationally predicted targets for microRNAs across many species. The miRNA sequences are obtained from the miRBase Sequence database and most genomic sequence from EnsEMBL. We aim to provide the most up-to-date and accurate predictions of miRNA targets and hence this resource will be updated regularly to incorporate new miRNAs or EnsEMBL sequences. ',

MicroCosm Targets Available at: http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/.

Gromer, S., Eubel, J.K., Lee, B.L. and Jacob, J. (2005) 'Human selenoproteins at a glance', *Cellular & Molecular Life Sciences*, 62(21), pp. 2414-37.

Gu, S. and Kay, M. (2010) 'How do miRNAs mediate translational repression?', *Silence*, 1(1), p. 11.

Guo, C.H., Liu, P.J., Hsia, S., Chuang, C.J. and Chen, P.C. (2011) 'Role of certain trace minerals in oxidative stress, inflammation, CD4/CD8 lymphocyte ratios and lung function in asthmatic patients', *Ann Clin Biochem*, 48(Pt 4), pp. 344-51.

Hackl, M., Brunner, S., Fortschegger, K., Schreiner, C., Micutkova, L., Muck, C., Laschober, G.T., Lepperdinger, G., Sampson, N., Berger, P., Herndler-Brandstetter, D., Wieser, M., Kuhnel, H., Strasser, A., Rinnerthaler, M., Breitenbach, M., Mildner, M., Eckhart, L., Tschachler, E., Trost, A., Bauer, J.W., Papak, C., Trajanoski, Z., Scheideler, M., Grillari-Voglauer, R., Grubeck-Loebenstein, B., Jansen-Durr, P. and Grillari, J. (2010) 'miR-17, miR-19b, miR-20a, and miR-106a are down-regulated in human aging', *Aging Cell*, 9(2), pp. 291-6.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano Jr, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M. and Tuschl, T. (2010) 'Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP', *Cell*, 141(1), pp. 129-141.

Haft, D.H. and Basu, M.K. (2011) 'Biological systems discovery in silico: radical S-adenosylmethionine protein families and their target peptides for posttranslational modification', *Journal of Bacteriology*, 193(11), pp. 2745-55.

Hamouda, H.E., Zakaria, S.S., Ismail, S.A., Khedr, M.A. and Mayah, W.W. (2011) 'p53 antibodies, metallothioneins, and oxidative stress markers in chronic ulcerative colitis with dysplasia', *World J Gastroenterol*, 17(19), pp. 2417-23.

Han, L., Witmer, P.D., Casey, E., Valle, D. and Sukumar, S. (2007) 'DNA methylation regulates MicroRNA expression', *Cancer Biology & Therapy*, 6(8), pp. 1284-8.

Harthill, M. (2011) 'Review: micronutrient selenium deficiency influences evolution of some viral infectious diseases', *Biol Trace Elem Res*, 143(3), pp. 1325-36.

Harvey, P.A. and Leinwand, L.A. (2011) 'The cell biology of disease: cellular mechanisms of cardiomyopathy', *J Cell Biol*, 194(3), pp. 355-65.

He, Y., Wang, Y., Li, P., Zhu, S., Wang, J. and Zhang, S. (2011) 'Identification of GPX3 epigenetically silenced by CpG methylation in human esophageal squamous cell carcinoma', *Digestive Diseases & Sciences*, 56(3), pp. 681-8.

Hebert, S.S. and De Strooper, B. (2009) 'Alterations of the microRNA network cause neurodegenerative disease', *Trends in Neurosciences*, 32(4), pp. 199-206.

Hesketh, J. (2008) 'Nutrigenomics and Selenium: Gene Expression Patterns, Physiological Targets, and Genetics', *The Annual Review of Nutrition*, 28, pp. 157-177.

Hino, K., Tsuchiya, K., Fukao, T., Kiga, K., Okamoto, R., Kanai, T. and Watanabe, M. (2008) 'Inducible expression of microRNA-194 is regulated by HNF-1alpha during intestinal epithelial cell differentiation', *Rna-A Publication of the Rna Society*, 14(7), pp. 1433-42.

Hock, J. and Meister, G. (2008) 'The Argonaute protein family', *Genome Biology*, 9(2), p. 210.

Hoefig, C.S., Renko, K., Köhrle, J., Birringer, M. and Schomburg, L. (2011) 'Comparison of different selenocompounds with respect to nutritional value vs. toxicity using liver cells in culture', *The Journal of Nutritional Biochemistry*, 22(10), pp. 945-955.

Hoffmann, P.R. and Berry, M.J. (2005) 'Selenoprotein synthesis: a unique translational mechanism used by a diverse family of proteins', *Thyroid*, 15(8), pp. 769-75.

Hu, Y., McIntosh, G.H., Le Leu, R.K., Upton, J.M., Woodman, R.J. and Young, G.P. (2011) 'The influence of selenium-enriched milk proteins and selenium yeast on plasma selenium levels and rectal selenoprotein

gene expression in human subjects', *British Journal of Nutrition*, 106(4), pp. 572-82.

Huang, J.Q., Li, D.L., Zhao, H., Sun, L.H., Xia, X.J., Wang, K.N., Luo, X. and Lei, X.G. (2011a) 'The selenium deficiency disease exudative diathesis in chicks is associated with downregulation of seven common selenoprotein genes in liver and muscle', *J Nutr*, 141(9), pp. 1605-10.

Huang, Y., Zou, Q., Song, H., Song, F., Wang, L., Zhang, G. and Shen, X. (2011b) 'A study of miRNAs targets prediction and experimental validation', *Protein & Cell*, 1(11), pp. 979-86.

Hubert, N., Walczak, R., Carbon, P. and Krol, A. (1996) 'A protein binds the selenocysteine insertion element in the 34-UTR of mammalian selenoprotein mRNAs', *Nucleic Acids Research,*, 24(3), pp. 464–469.

Huntzinger, E. and Izaurralde, E. (2011) 'Gene silencing by microRNAs: contributions of translational repression and mRNA decay', *Nature Reviews Genetics*, 12(2), pp. 99-110.

Hurst, R., Elliott, R.M., Goldson, A.J. and Fairweather-Tait, S.J. (2008) 'Se-methylselenocysteine alters collagen gene and protein expression in human prostate cells', *Cancer Letters*, 269(1), pp. 117-26.

Hurwitz, B.E., Klaus, J.R., Llabre, M.M., Gonzalez, A., Lawrence, P.J., Maher, K.J., Greeson, J.M., Baum, M.K., Shor-Posner, G., Skyler, J.S. and Schneiderman, N. (2007) 'Suppression of human immunodeficiency virus type 1 viral load with selenium supplementation: a randomized controlled trial', *Arch Intern Med*, 167(2), pp. 148-54.

Hwang, D.Y., Cho, J.S., Oh, J.H., Shim, S.B., Jee, S.W., Lee, S.H., Seo, S.J., Lee, S.K. and Kim, Y.K. (2005) 'Differentially expressed genes in transgenic mice carrying human mutant presenilin-2 (N141I): correlation of selenoprotein M with Alzheimer's disease', *Neurochem Res*, 30(8), pp. 1009-19.

Imam, J.S., Buddavarapu, K., Lee-Chang, J.S., Ganapathy, S., Camosy, C., Chen, Y. and Rao, M.K. (2010) 'MicroRNA-185 suppresses tumor growth and progression by targeting the Six1 oncogene in human cancers', *Oncogene*, 29(35), pp. 4971-9.

Ingenuity (2008) *Ingenuity Pathways Analysis*. Available at: http://www.ingenuity.com/products/pathways_analysis.html.

Integrated DNA Technologies, I. (2009) *Integrated DNA Tecnologies*. Available at: http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx (Accessed: 2009).

Irons, R., Carlson, B.A., Hatfield, D.L., Davis, C.D., Irons, R., Carlson, B.A., Hatfield, D.L. and Davis, C.D. (2006) 'Both selenoproteins and low molecular weight selenocompounds reduce colon cancer risk in mice with genetically impaired selenoprotein expression', *Journal of Nutrition*, 136(5), pp. 1311-7.

Jackson, R.J. and Standart, N. (2007) 'How do microRNAs regulate gene expression?', *Science's Stke [Electronic Resource]: Signal Transduction Knowledge Environment*, 2007(367), p. re1.

Jakymiw, A., Patel, R.S., Deming, N., Bhattacharyya, I., Shah, P., Lamont, R.J., Stewart, C.M., Cohen, D.M. and Chan, E.K.L. (2010) 'Overexpression of dicer as a result of reduced let-7 MicroRNA levels contributes to increased cell proliferation of oral cancer cells', *Genes, Chromosomes & Cancer*, 49(6), pp. 549-59.

Jia, S.S., Xi, G.P., Zhang, M., Chen, Y.B., Lei, B., Dong, X.S. and Yang, Y.M. (2013) 'Induction of apoptosis by D-limonene is mediated by inactivation of Akt in LS174T human colon cancer cells', *Oncol Rep*, 29(1), pp. 349-54.

Jin, H., Tuo, W., Lian, H., Liu, Q., Zhu, X.-Q. and Gao, H. (2010) 'Strategies to identify microRNA targets: new advances', *New Biotechnology*, 27(6), pp. 734-8.

Jin, R.C., Mahoney, C.E., Coleman Anderson, L., Ottaviano, F., Croce, K., Leopold, J.A., Zhang, Y.-Y., Tang, S.-S., Handy, D.E. and Loscalzo, J. (2011) 'Glutathione peroxidase-3 deficiency promotes platelet-dependent thrombosis in vivo', *Circulation*, 123(18), pp. 1963-73.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C. and Marks, D.S. (2004) 'Human MicroRNA targets', *Plos Biology*, 2(11), p. e363.

Ju, X., Li, D., Shi, Q., Hou, H., Sun, N. and Shen, B. (2009) 'Differential microRNA expression in childhood B-cell precursor acute lymphoblastic leukemia', *Pediatr Hematol Oncol*, 26(1), pp. 1-10.

Juhila, J., Sipila, T., Icay, K., Nicorici, D., Ellonen, P., Kallio, A., Korpelainen, E., Greco, D. and Hovatta, I. (2011) 'MicroRNA expression profiling reveals miRNA families regulating specific biological pathways in mouse frontal cortex and hippocampus', *PLoS One*, 6(6), p. 22.

Kahlert, C., Klupp, F., Brand, K., Lasitschka, F., Diederichs, S., Kirchberg, J., Rahbari, N., Dutta, S., Bork, U., Fritzmann, J., Reissfelder, C., Koch, M. and Weitz, J. (2011) 'Invasion front-specific expression and prognostic significance of microRNA in colorectal liver metastases', *Cancer Science*, 102(10), pp. 1799-807.

Karra, D. and Dahm, R. (2010) 'Transfection Techniques for Neuronal Cells', *The Journal of Neuroscience*, 30(18), pp. 6171-6177.

Kasaikina, M.V., Fomenko, D.E., Labunskyy, V.M., Lachke, S.A., Qiu, W., Moncaster, J.A., Zhang, J., Wojnarowicz, M.W., Jr., Natarajan, S.K., Malinouski, M., Schweizer, U., Tsuji, P.A., Carlson, B.A., Maas, R.L., Lou, M.F., Goldstein, L.E., Hatfield, D.L. and Gladyshev, V.N. (2011a) 'Roles of the 15-kDa selenoprotein (Sep15) in redox homeostasis and cataract development revealed by the analysis of Sep 15 knockout mice', *Journal of Biological Chemistry*, 286(38), pp. 33203-12.

Kasaikina, M.V., Kravtsova, M.A., Lee, B.C., Seravalli, J., Peterson, D.A., Walter, J., Legge, R., Benson, A.K., Hatfield, D.L. and Gladyshev, V.N. (2011b) 'Dietary selenium affects host selenoproteome expression by influencing the gut microbiota', *FASEB Journal*, 25(7), pp. 2492-9.

Kato, M.A., Finley, D.J., Lubitz, C.C., Zhu, B., Moo, T.A., Loeven, M.R., Ricci, J.A., Zarnegar, R., Katdare, M. and Fahey, T.J., 3rd (2010) 'Selenium decreases thyroid cancer cell growth by increasing expression of GADD153 and GADD34', *Nutrition & Cancer*, 62(1), pp. 66-73.

Kawamata, T. and Tomari, Y. (2010) 'Making RISC', *Trends in Biochemical Sciences*, 35(7), pp. 368-76.

Khan, A.A., Betel, D., Miller, M.L., Sander, C., Leslie, C.S. and Marks, D.S. (2009) 'Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs.[Erratum appears in Nat Biotechnol. 2009 Jul;27(7):671]', *Nature Biotechnology*, 27(6), pp. 549-55.

Kiewe, P., Gueller, S., Komor, M., Stroux, A., Thiel, E. and Hofmann, W.-K. (2009) 'Prediction of qualitative outcome of

oligonucleotide microarray hybridization by measurement of RNA integrity using the 2100 Bioanalyzer™ capillary electrophoresis system', *Annals of Hematology*, 88(12), pp. 1177-1183.

Kim, B.S., Jung, J.S., Jang, J.H., Kang, K.S. and Kang, S.K. (2011a) 'Nuclear Argonaute 2 regulates adipose tissue-derived stem cell survival through direct control of miR10b and selenoprotein N1 expression', *Aging Cell*, 10(2), pp. 277-91.

Kim, J.-S., Choi, Y.Y., Jin, G., Kang, H.-G., Choi, J.-E., Jeon, H.-S., Lee, W.-K., Kim, D.-S., Kim, C.H., Kim, Y.J., Son, J.-W., Jung, T.H. and Park, J.Y. (2010) 'Association of a common AGO1 variant with lung cancer risk: a two-stage case-control study', *Molecular Carcinogenesis*, 49(10), pp. 913-21.

Kim, K., Lee, H.C., Park, J.L., Kim, M., Kim, S.Y., Noh, S.M., Song, K.S., Kim, J.C. and Kim, Y.S. (2011b) 'Epigenetic regulation of microRNA-10b and targeting of oncogenic MAPRE1 in gastric cancer', *Epigenetics*, 6(6), pp. 740-51.

Kim, V.N., Han, J. and Siomi, M.C. (2009a) 'Biogenesis of small RNAs in animals', *Nature Reviews Molecular Cell Biology*, 10(2), pp. 126-39.

Kim, W.K., Park, M., Kim, Y.-K., Tae, Y.K., Yang, H.-K., Lee, J.M. and Kim, H. (2011c) 'MicroRNA-494 downregulates KIT and inhibits gastrointestinal stromal tumor cell proliferation', *Clinical Cancer Research*, 17(24), pp. 7584-94.

Kim, Y.-K., Yu, J., Han, T.S., Park, S.-Y., Namkoong, B., Kim, D.H., Hur, K., Yoo, M.-W., Lee, H.-J., Yang, H.-K. and Kim, V.N. (2009b) 'Functional links between clustered microRNAs: suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer', *Nucleic Acids Research*, 37(5), pp. 1672-81.

Kipp, A., Banning, A., van Schothorst, E.M., Méplan, C., Schomburg, L., Evelo, C., Coort, S., Gaj, S., Keijer, J., Hesketh, J.E. and Brigelius-Flohé, R. (2009) 'Four selenoproteins, protein biosynthesis, and Wnt signalling are particularly sensitive to limited selenium intake in mouse colon', *Molecular Nutrition & Food Research*, 53(12), pp. 1561-1572.

Kipp, A.P., Muller, M.F., Goken, E.M., Deubel, S. and Brigelius-Flohe, R. (2012) 'The selenoproteins GPx2, TrxR2 and TrxR3 are

regulated by Wnt signalling in the intestinal epithelium', *Biochim Biophys Acta*, 7(10), pp. 1588-1596.

Kohler, A. and Hurt, E. (2007) 'Exporting RNA from the nucleus to the cytoplasm', *Nat Rev Mol Cell Biol*, 8(10), pp. 761-773.

Kong, W., Zhao, J.-J., He, L. and Cheng, J.Q. (2009) 'Strategies for profiling microRNA expression', *Journal of Cellular Physiology*, 218(1), pp. 22-5.

Kong, Y.W., Cannell, I.G., de Moor, C.H., Hill, K., Garside, P.G., Hamilton, T.L., Meijer, H.A., Dobbyn, H.C., Stoneley, M., Spriggs, K.A., Willis, A.E. and Bushell, M. (2008) 'The mechanism of micro-RNA-mediated translation repression is determined by the promoter of the target gene', *Proceedings of the National Academy of Sciences of the United States of America*, 105(26), pp. 8866-71.

Krehl, S., Loewinger, M., Florian, S., Kipp, A.P., Banning, A., Wessjohann, L.A., Brauer, M.N., Iori, R., Esworthy, R.S., Chu, F.-F. and Brigelius-Flohe, R. (2012) 'Glutathione peroxidase-2 and selenium decreased inflammation and tumors in a mouse model of inflammation-associated carcinogenesis whereas sulforaphane effects differed with selenium supply', *Carcinogenesis*, 33(3), pp. 620-8.

Krol, J., Loedige, I. and Filipowicz, W. (2010) 'The widespread regulation of microRNA biogenesis, function and decay', *Nat Rev Genet*, 11(9), pp. 597-610.

Kryukov, G.V., Castellano, S., Novoselov, S.V., Lobanov, A.V., Zehtab, O., Guigó, R. and Gladyshev, V.N. (2003) 'Characterization of Mammalian Selenoproteomes', *Science*, 300(5624), pp. 1439-1443.

Kuhn, D.E., Martin, M.M., Feldman, D.S., Terry, A.V., Jr., Nuovo, G.J. and Elton, T.S. (2008) 'Experimental validation of miRNA targets', *Methods*, 44(1), pp. 47-54.

Kuhn, R.M., Karolchik, D., Zweig, A.S., Trumbower, H., Thomas, D.J., Thakkapallayil, A., Sugnet, C.W., Stanke, M., Smith, K.E., Siepel, A., Rosenbloom, K.R., Rhead, B., Raney, B.J., Pohl, A., Pedersen, J.S., Hsu, F., Hinrichs, A.S., Harte, R.A., Diekhans, M., Clawson, H., Bejerano, G., Barber, G.P., Baertsch, R., Haussler, D. and Kent, W.J. (2007) 'The UCSC genome browser database: update 2007', *Nucleic Acids Research*, 35(Database issue), pp. D668-73.

Kurosawa, T., Nakamura, H., Yamaura, E., Fujino, H., Matsuzawa, Y., Kawashima, T. and Murayama, T. (2009) 'Cytotoxicity induced by inhibition of thioredoxin reductases via multiple signaling pathways: role of cytosolic phospholipase A(2)alpha-dependent and -independent release of arachidonic acid', *J Cell Physiol*, 219(3), pp. 606-16.

Lai, E.C. (2002) 'Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation', *Nature Genetics*, 30(4), pp. 363-4.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., Lin, C., Socci, N.D., Hermida, L., Fulci, V., Chiaretti, S., Foa, R., Schliwka, J., Fuchs, U., Novosel, A., Muller, R.-U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D.B., Choksi, R., De Vita, G., Frezzetti, D., Trompeter, H.-I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Di Lauro, R., Wernet, P., Macino, G., Rogler, C.E., Nagle, J.W., Ju, J., Papavasiliou, F.N., Benzing, T., Lichter, P., Tam, W., Brownstein, M.J., Bosio, A., Borkhardt, A., Russo, J.J., Sander, C., Zavolan, M. and Tuschl, T. (2007) 'A mammalian microRNA expression atlas based on small RNA library sequencing', *Cell*, 129(7), pp. 1401-14.

Langenberger, D., Bartschat, S., Hertel, J., Hoffmann, S., Tafer, H. and Stadler, P. (2011) 'MicroRNA or Not MicroRNA?', in Norberto de Souza, O., Telles, G. and Palakal, M. (eds.) *Advances in Bioinformatics and Computational Biology*. Springer Berlin Heidelberg, pp. 1-9.

Lee, I., Ajay, S.S., Yook, J.I., Kim, H.S., Hong, S.H., Kim, N.H., Dhanasekaran, S.M., Chinnaiyan, A.M. and Athey, B.D. (2009) 'New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites', *Genome Research*, 19(7), pp. 1175-83.

Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) 'The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14', *Cell*, 75(5), pp. 843-854.

Leist, M., Raab, B., Maurer, S., Rösick, U. and Brigelius-Flohé, R. (1996) 'Conventional cell culture media do not adequately supply cells with antioxidants and thus facilitate peroxide-induced genotoxicity', *Free Radical Biology and Medicine*, 21(3), pp. 297-306.

Lenaz, G. (2012) 'Mitochondria and reactive oxygen species. Which role in physiology and pathology?', *Adv Exp Med Biol*, 942, pp. 93-136.

Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) 'Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets', *Cell*, 120(1), pp. 15-20.

Li, C.L., Nan, K.J., Tian, T., Sui, C.G. and Liu, Y.F. (2007) 'Selenoprotein P mRNA expression in human hepatic tissues', *World J Gastroenterol*, 13(16), pp. 2363-8.

Li, J., Chen, Y., Zhao, J., Kong, F. and Zhang, Y. (2011) 'miR-203 reverses chemoresistance in p53-mutated colon cancer cells through downregulation of Akt2 expression', *Cancer Letters*, 304(1), pp. 52-9.

Li, L.-C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H. and Dahiya, R. (2006) 'Small dsRNAs induce transcriptional activation in human cells', *Proceedings of the National Academy of Sciences of the United States of America*, 103(46), pp. 17337-42.

Li, S.Y., Cao, J.L., Shi, Z.L., Chen, J.H., Zhang, Z.T., Hughes, C.E. and Caterson, B. (2008) 'Promotion of the articular cartilage proteoglycan degradation by T-2 toxin and selenium protective effect', *J Zhejiang Univ Sci B*, 9(1), pp. 22-33.

Li, Z., Su, Z., Wen, Z., Shi, L. and Chen, T. (2009) 'Microarray platform consistency is revealed by biologically functional analysis of gene expression profiles', *BMC Bioinformatics*, 10 Suppl 11, p. S12.

Liao, S.F., Brown, K.R., Stromberg, A.J., Burris, W.R., Boling, J.A. and Matthews, J.C. (2011) 'Dietary supplementation of selenium in inorganic and organic forms differentially and commonly alters blood and liver selenium concentrations and liver gene expression profiles of growing beef heifers', *Biological Trace Element Research*, 140(2), pp. 151-69.

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. and Johnson, J.M. (2005) 'Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs', *Nature*, 433(7027), pp. 769-773.

Lin, S.-L., Chang, D.C., Chang-Lin, S., Lin, C.-H., Wu, D.T.S., Chen, D.T. and Ying, S.-Y. (2008) 'Mir-302 reprograms human skin cancer

cells into a pluripotent ES-cell-like state', *Rna-A Publication of the Rna Society*, 14(10), pp. 2115-24.

Linsley, P.S., Schelter, J., Burchard, J., Kibukawa, M., Martin, M.M., Bartz, S.R., Johnson, J.M., Cummins, J.M., Raymond, C.K., Dai, H., Chau, N., Cleary, M., Jackson, A.L., Carleton, M. and Lim, L. (2007) 'Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression', *Molecular & Cellular Biology*, 27(6), pp. 2240-52.

Liu, J. (2008) 'Control of protein synthesis and mRNA degradation by microRNAs', *Current Opinion in Cell Biology*, 20(2), pp. 214-221.

Liu, J., Uematsu, H., Tsuchida, N. and Ikeda, M.-A. (2011a) 'Essential role of caspase-8 in p53/p73-dependent apoptosis induced by etoposide in head and neck carcinoma cells', *Molecular Cancer*, 10(1), p. 95.

Liu, M., Lang, N., Chen, X., Tang, Q., Liu, S., Huang, J., Zheng, Y. and Bi, F. (2011b) 'miR-185 targets RhoA and Cdc42 expression and inhibits the proliferation potential of human colorectal cells', *Cancer Letters*, 301(2), pp. 151-60.

Liu, X., Jin, D.-Y., McManus, Michael T. and Mourelatos, Z. (2012) 'Precursor MicroRNA-Programmed Silencing Complex Assembly Pathways in Mammals', *Molecular Cell*, (0).

Liu, Y., Chiba, M., Inaba, Y. and Kondo, M. (2002) '[Keshan disease--a review from the aspect of history and etiology]', *Nippon Eiseigaku Zasshi - Japanese Journal of Hygiene*, 56(4), pp. 641-8.

Llave, C., Franco-Zorrilla, J.M., Solano, R. and Barajas, D. (2011) 'Target validation of plant microRNAs', *Methods Mol Biol*, 732, pp. 187-208.

Long, D., Chan, C.Y. and Ding, Y. (2008) 'Analysis of microRNA-target interactions by a target structure based hybridization model', *Pacific Symposium on Biocomputing*, pp. 64-74.

Lu, J. and Holmgren, A. (2009) 'Selenoproteins', *Journal of Biological Chemistry*, 284(2), pp. 723-7.

Lu, Y., Ryan, S.L., Elliott, D.J., Bignell, G.R., Futreal, P.A., Ellison, D.W., Bailey, S. and Clifford, S.C. (2009) 'Amplification and overexpression of Hsa-miR-30b, Hsa-miR-30d and KHDRBS3 at 8q24.22-q24.23 in medulloblastoma', *PLoS One*, 4(7).

Lubos, E., Kelly, N.J., Oldebeken, S.R., Leopold, J.A., Zhang, Y.-Y., Loscalzo, J. and Handy, D.E. (2011) 'Glutathione peroxidase-1 deficiency augments proinflammatory cytokine-induced redox signaling and human endothelial cell activation', *Journal of Biological Chemistry*, 286(41), pp. 35407-17.

Lubos, E., Mahoney, C.E., Leopold, J.A., Zhang, Y.-Y., Loscalzo, J. and Handy, D.E. (2010) 'Glutathione peroxidase-1 modulates lipopolysaccharide-induced adhesion molecule expression in endothelial cells by altering CD14 expression', *FASEB Journal*, 24(7), pp. 2525-32.

Luo, H., Yang, Y., Huang, F., Li, F., Jiang, Q., Shi, K. and Xu, C. (2012) 'Selenite induces apoptosis in colorectal cancer cells via AKT-mediated inhibition of beta-catenin survival axis', *Cancer Letters*, 315(1), pp. 78-85.

Lv, J., Wang, W., Krafft, T., Li, Y., Zhang, F. and Yuan, F. (2011) 'Effects of several environmental factors on longevity and health of the human population of Zhongxiang, Hubei, China', *Biol Trace Elem Res*, 143(2), pp. 702-16.

Lytle, J.R., Yario, T.A. and Steitz, J.A. (2007) 'Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR', *Proceedings of the National Academy of Sciences of the United States of America*, 104(23), pp. 9667-72.

Mao, H., Lebrun, D.G., Yang, J., Zhu, V.F. and Li, M. (2012) 'Deregulated signaling pathways in glioblastoma multiforme: molecular mechanisms and therapeutic targets', *Cancer Invest*, 30(1), pp. 48-56.

Maraldi, T., Riccio, M., Zambonin, L., Vinceti, M., De Pol, A. and Hakim, G. (2011) 'Low levels of selenium compounds are selectively toxic for a human neuron cell line through ROS/RNS increase and apoptotic process activation', *Neurotoxicology*, 32(2), pp. 180-7.

Mariotti, M. and Guigo, R. (2010) 'Selenoprofiles: profile-based scanning of eukaryotic genome sequences for selenoprotein genes', *Bioinformatics*, 26(21), pp. 2656-63.

Mariotti, M., Ridge, P.G., Zhang, Y., Lobanov, A.V., Pringle, T.H., Guigo, R., Hatfield, D.L. and Gladyshev, V.N. (2012) 'Composition and Evolution of the Vertebrate and Mammalian Selenoproteomes', *PLoS ONE*, 7(3), p. e33066.

Mazière, P. and Enright, A.J. (2007) 'Prediction of microRNA targets', *Drug Discovery Today*, 12(11-12), pp. 452-8.

McCann, J.C. and Ames, B.N. (2011) 'Adaptive dysfunction of selenoproteins from the perspective of the triage theory: why modest selenium deficiency may increase risk of diseases of aging', *FASEB Journal*, 25(6), pp. 1793-814.

McCleary-Wheeler, A.L., McWilliams, R. and Fernandez-Zapico, M.E. (2012) 'Aberrant signaling pathways in pancreatic cancer: a two compartment view', *Molecular Carcinogenesis*, 51(1), pp. 25-39.

McKay, J.A. and Mathers, J.C. (2011) 'Diet induced epigenetic changes and their implications for health', *Acta Physiologica*, 202(2), pp. 103-118.

McKay, M.M. and Morrison, D.K. (2007) 'Integrating signals from RTKs to ERK//MAPK', *Oncogene*, 26(22), pp. 3113-3121.

Mendelev, N., Mehta, S.L., Witherspoon, S., He, Q., Sexton, J.Z. and Li, P.A. (2011) 'Upregulation of human selenoprotein H in murine hippocampal neuronal cells promotes mitochondrial biogenesis and functional performance', *Mitochondrion*, 11(1), pp. 76-82.

Mendelev, N., Witherspoon, S. and Li, P.A. (2009) 'Overexpression of human selenoprotein H in neuronal cells ameliorates ultraviolet irradiation-induced damage by modulating cell signaling pathways', *Experimental Neurology*, 220(2), pp. 328-34.

Meng, Y., Shao, C. and Chen, M. (2011) 'Toward microRNA-mediated gene regulatory networks in plants', *Briefings in Bioinformatics*, 12(6), pp. 645-59.

Meplan, C. and Hesketh, J. (2012) 'The influence of selenium and selenoprotein gene variants on colorectal cancer risk', *Mutagenesis*, 27(2), pp. 177-86.

Méplan C., P.V., Hesketh J.E. (2006) 'Advances in Selenoprotein Expression: Patterns and Individual Variations', in Brigelius-Flohé, R., Joost, H. (ed.) *Nutritional Genomics. Impact on Health and Disease*. First edn. Nuthetal: Wiley-VCH, Weinheim, pp. 132-153.

Mestdagh, P., Van Vlierberghe, P., De Weer, A., Muth, D., Westermann, F., Speleman, F. and Vandesompele, J. (2009) 'A novel and

universal method for microRNA RT-qPCR data normalization', *Genome Biology*, 10(6), p. R64.

Mihaylova, M.M. and Shaw, R.J. (2011) 'The AMPK signalling pathway coordinates cell growth, autophagy and metabolism', *Nat Cell Biol*, 13(9), pp. 1016-1023.

Mizui, M., Kumanogoh, A. and Kikutani, H. (2009) 'Immune semaphorins: novel features of neural guidance molecules', *Journal of Clinical Immunology*, 29(1), pp. 1-11.

Moghadaszadeh, B. and Beggs, A.H. (2006) 'Selenoproteins and their impact on human health through diverse physiological pathways', *Physiology*, 21, pp. 307-15.

Monzo, M., Navarro, A., Bandres, E., Artells, R., Moreno, I., Gel, B., Ibeas, R., Moreno, J., Martinez, F., Diaz, T., Martinez, A., Balague, O. and Garcia-Foncillas, J. (2008) 'Overlapping expression of microRNAs in human embryonic colon and colorectal cancer', *Cell Research*, 18(8), pp. 823-33.

Moon, H.-J., Ko, W.-K., Han, S.W., Kim, D.-S., Hwang, Y.-S., Park, H.-K. and Kwon, I.K. (2012) 'Antioxidants, like coenzyme Q10, selenite, and curcumin, inhibited osteoclast differentiation by suppressing reactive oxygen species generation', *Biochemical and Biophysical Research Communications*, 418(2), pp. 247-253.

Moreno-Reyes, R., Mathieu, F., Boelaert, M., Begaux, F., Suetens, C., Rivera, M.T., Neve, J., Perlmutter, N. and Vanderpas, J. (2003) 'Selenium and iodine supplementation of rural Tibetan children affected by Kashin-Beck osteoarthropathy', *American Journal of Clinical Nutrition*, 78(1), pp. 137-44.

Morisseau, C., Schebb, N.H., Dong, H., Ulu, A., Aronov, P.A. and Hammock, B.D. (2012) 'Role of soluble epoxide hydrolase phosphatase activity in the metabolism of lysophosphatidic acids', *Biochem Biophys Res Commun*, 419(4), pp. 796-800.

Mueller, A.C., Sun, D. and Dutta, A. (2012) 'The miR-99 family regulates the DNA damage response through its target SNF2H', *Oncogene*.

Naderi, A., Ahmed, A., Barbosa-Morais, N., Aparicio, S., Brenton, J. and Caldas, C. (2004) 'Expression microarray reproducibility is

improved by optimising purification steps in RNA amplification and labelling', *BMC Genomics*, 5(1), p. 9.

Nafisi, S., Montazeri, M. and Manouchehri, F. (2012) 'The effect of Se salts on DNA structure', *Journal of Photochemistry and Photobiology B: Biology*, 113(0), pp. 36-41.

Navarro-Alarcon, M. and Cabrera-Vique, C. (2008) 'Selenium in food and the human body: a review', *Science of the Total Environment*, 400(1-3), pp. 115-41.

NCBI *Entrez Nucleotide database*. Available at: http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore (Accessed: 2009).

Nishida, N., Nagahara, M., Sato, T., Mimori, K., Sudo, T., Tanaka, F., Shibata, K., Ishii, H., Sugihara, K., Doki, Y. and Mori, M. (2012) 'Microarray Analysis of Colorectal Cancer Stromal Tissue Reveals Upregulation of Two Oncogenic miRNA Clusters', *Clinical Cancer Research*, 18(11), pp. 3054-70.

Norden-Krichmar, T.M., Holtz, J., Pasquinelli, A.E. and Gaasterland, T. (2007) 'Computational prediction and experimental validation of Ciona intestinalis microRNA genes', *BMC Genomics*, 8, p. 445.

Nordgren, K.K., Peng, Y., Pelleymounter, L.L., Moon, I., Abo, R., Feng, Q., Eckloff, B., Yee, V.C., Wieben, E. and Weinshilboum, R.M. (2011) 'Methionine adenosyltransferase 2A/2B and methylation: gene sequence variation and functional genomics', *Drug Metab Dispos*, 39(11), pp. 2135-47.

Nugent, M., Miller, N. and Kerin, M.J. (2011) 'MicroRNAs in colorectal cancer: Function, dysregulation and potential as novel biomarkers', *European Journal of Surgical Oncology (EJSO)*, 37(8), pp. 649-654.

Ogier-Denis, E., Fasseu, M., Vandewalle, A. and Laburthe, M. (2007) 'MicroRNAs and intestinal pathophysiology', *M S-Medecine Sciences*, 23(5), pp. 509-14.

Okamura, K., Phillips, M.D., Tyler, D.M., Duan, H., Chou, Y.-t. and Lai, E.C. (2008) 'The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution', *Nature Structural & Molecular Biology*, 15(4), pp. 354-63.

Olive, V., Jiang, I. and He, L. (2010) 'mir-17-92, a cluster of miRNAs in the midst of the cancer network', *International Journal of Biochemistry & Cell Biology*, 42(8), pp. 1348-54.

Olsson, M., Olsson, B., Jacobson, P., Thelle, D.S., Bjorkegren, J., Walley, A., Froguel, P., Carlsson, L.M. and Sjoholm, K. (2011) 'Expression of the selenoprotein S (SELS) gene in subcutaneous adipose tissue and SELS genotype are associated with metabolic risk factors', *Metabolism*, 60(1), pp. 114-20.

Orom, U.A., Nielsen, F.C. and Lund, A.H. (2008) 'MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation', *Molecular Cell*, 30(4), pp. 460-71.

Ottaviano, F.G., Tang, S.-S., Handy, D.E. and Loscalzo, J. (2009) 'Regulation of the extracellular antioxidant selenoprotein plasma glutathione peroxidase (GPx-3) in mammalian cells', *Molecular & Cellular Biochemistry*, 327(1-2), pp. 111-26.

Pagmantidis, V., Bermano, G., Villette, S., Broom, I., Arthur, J. and Hesketh, J. (2005) 'Effects of Se-depletion on glutathione peroxidase and selenoprotein W gene expression in the colon', *FEBS Letters*, 579(3), pp. 792-796.

Pagmantidis, V., Méplan, C., van Schothorst, E.M., Keijer, J. and Hesketh, J.E. (2008) 'Supplementation of healthy volunteers with nutritionally relevant amounts of selenium increases the expression of lymphocyte protein biosynthesis genes', *Am J Clin Nutr*, 87, pp. 181-189.

Papp, L.V., Lu, J., Holmgren, A. and Khanna, K.K. (2007) 'From selenium to selenoproteins: synthesis, identity, and their role in human health', *Antioxidants & Redox Signaling*, 9(7), pp. 775-806.

Park, L., Min, D., Kim, H., Park, J., Choi, S. and Park, Y. (2011) 'The combination of metallothionein and superoxide dismutase protects pancreatic beta cells from oxidative damage', *Diabetes Metab Res Rev*, 27(8), pp. 802-8.

Park, S.-M., Gaur, A.B., Lengyel, E. and Peter, M.E. (2008) 'The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2', *Genes & Development*, 22(7), pp. 894-907.

Pasquinelli, A.E. (2012) 'MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship', *Nat Rev Genet*, 13(4), pp. 271-282.

Pasterkamp, R.J. (2012) 'Getting neural circuits into shape with semaphorins', *Nat Rev Neurosci*, 13(9), pp. 605-618.

Pei, Z., Li, H., Guo, Y., Jin, Y. and Lin, D. (2010) 'Sodium selenite inhibits the expression of VEGF, TGFbeta(1) and IL-6 induced by LPS in human PC3 cells via TLR4-NF-(K)B signaling blockage', *Int Immunopharmacol*, 10(1), pp. 50-6.

Peng, S., Zeng, X., Li, X., Peng, X. and Chen, L. (2009) 'Multi-class cancer classification through gene expression profiles: microRNA versus mRNA', *Journal Of Genetics & Genomics = Yi Chuan Xue Bao*, 36(7), pp. 409-16.

Peters, U. and Takata, Y. (2008) 'Selenium and the prevention of prostate and colorectal cancer', *Molecular Nutrition & Food Research*, 52(11), pp. 1261-72.

Piana, C., Wirth, M., Gerbes, S., Viernstein, H., Gabor, F. and Toegel, S. (2008) 'Validation of reference genes for qPCR studies on Caco-2 cell differentiation', *European Journal of Pharmaceutics and Biopharmaceutics*, 69(3), pp. 1187-1192.

Pitto, L., Ripoli, A., Cremisi, F., Simili, M. and Rainaldi, G. (2008) 'microRNA(interference) networks are embedded in the gene regulatory networks', *Cell Cycle*, 7(16), pp. 2458-61.

Poerschke, R.L. and Moos, P.J. (2011) 'Thioredoxin reductase 1 knockdown enhances selenazolidine cytotoxicity in human lung cancer cells via mitochondrial dysfunction', *Biochem Pharmacol*, 81(2), pp. 211-21.

Ramakrishnan, K., Shenbagarathai, R., Kavitha, K., Thirumalaikolundusubramanian, P. and Rathinasabapati, R. (2012) 'Selenium levels in persons with HIV/tuberculosis in India, Madurai City', *Clin Lab*, 58(1-2), pp. 165-8.

Rayman, M.P. (2012) 'Selenium and human health', *Lancet*, 379(9822), pp. 1256-68.

Rayman, M.P., Thompson, A.J., Bekaert, B., Catterick, J., Galassini, R., Hall, E., Warren-Perry, M. and Beckett, G.J. (2008)

'Randomized controlled trial of the effect of selenium supplementation on thyroid function in the elderly in the United Kingdom', *Am J Clin Nutr*, 87(2), pp. 370-8.

Raymond, C.K., Roberts, B.S., Garrett-Engele, P., Lim, L.P. and Johnson, J.M. (2005) 'Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs', *Rna-A Publication of the Rna Society*, 11(11), pp. 1737-44.

Reeves, M.A., Bellinger, F.P. and Berry, M.J. (2010) 'The neuroprotective functions of selenoprotein M and its role in cytosolic calcium regulation', *Antioxidants & Redox Signaling*, 12(7), pp. 809-18.

Reeves, M.A. and Hoffmann, P.R. (2009) 'The human selenoproteome: recent insights into functions and regulation', *Cellular & Molecular Life Sciences*, 66(15), pp. 2457-78.

Reinhardt, H.C. and Schumacher, B. (2012) 'The p53 network: cellular and systemic DNA damage responses in aging and cancer', *Trends in Genetics*, 28(3), pp. 128-136.

Reszka, E., Jablonska, E., Gromadzinska, J. and Wasowicz, W. (2012) 'Relevance of selenoprotein transcripts for selenium status in humans', *Genes Nutr*, 7(2), pp. 127-37.

Ro, S., Park, C., Jin, J., Sanders, K.M. and Yan, W. (2006) 'A PCR-based method for detection and quantification of small RNAs', *Biochemical & Biophysical Research Communications*, 351(3), pp. 756-63.

Rock, C. and Moos, P.J. (2010) 'Selenoprotein P protects cells from lipid hydroperoxides generated by 15-LOX-1', *Prostaglandins Leukot Essent Fatty Acids*, 83(4-6), pp. 203-10.

Röther, S. and Meister, G. (2011) 'Small RNAs derived from longer non-coding RNAs', *Biochimie*, 93(11), pp. 1905-1915.

Rump, L., Asamoah, B. and Gonzalez-Escalona, N. (2010) 'Comparison of commercial RNA extraction kits for preparation of DNA-free total RNA from Salmonella cells', *BMC Research Notes*, 3(1), p. 211.

Saetrom, P., Heale, B.S.E., Snove, O., Jr., Aagaard, L., Alluin, J. and Rossi, J.J. (2007) 'Distance constraints between microRNA target sites dictate efficacy and cooperativity', *Nucleic Acids Research*, 35(7), pp. 2333-42.

Saito, Y., Suzuki, H. and Hibi, T. (2009) 'The role of microRNAs in gastrointestinal cancers', *Journal of Gastroenterology*, 44 Suppl 19, pp. 18-22.

Sales, G., Coppe, A., Bicciato, S., Bortoluzzi, S. and Romualdi, C. (2010) 'Impact of probe annotation on the integration of miRNA–mRNA expression profiles for miRNA target detection', *Nucleic Acids Research*, 38(7), p. e97.

Sanderson, P., Elsom, R.L., Kirkpatrick, V., Calder, P.C., Woodside, J.V., Williams, E.A., Rink, L., Fairweather-Tait, S., Ivory, K., Cantorna, M., Watzl, B. and Stone, E.M. (2010) 'UK food standards agency workshop report: diet and immune function', *Br J Nutr*, 103(11), pp. 1684-7.

Sarver, A., French, A., Borralho, P., Thayanithy, V., Oberg, A., Silverstein, K., Morlan, B., Riska, S., Boardman, L., Cunningham, J., Subramanian, S., Wang, L., Smyrk, T., Rodrigues, C., Thibodeau, S. and Steer, C. (2009) 'Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states', *BMC Cancer*, 9(1), p. 401.

Sarveswaran, S., Liroff, J., Zhou, Z., Nikitin, A.Y. and Ghosh, J. (2010) 'Selenite triggers rapid transcriptional activation of p53, and p53-mediated apoptosis in prostate cancer cells: Implication for the treatment of early-stage prostate cancer', *Int J Oncol*, 36(6), pp. 1419-28.

Sato, F., Tsuchiya, S., Terasawa, K. and Tsujimoto, G. (2009) 'Intra-platform repeatability and inter-platform comparability of microRNA microarray technology', *PLoS ONE [Electronic Resource]*, 4(5), p. e5540.

Satzger, I., Mattern, A., Kuettler, U., Weinspach, D., Voelker, B., Kapp, A. and Gutzmer, R. (2010) 'MicroRNA-15b represents an independent prognostic parameter and is correlated with tumor cell proliferation and apoptosis in malignant melanoma', *Int J Cancer*, 126(11), pp. 2553-62.

Schmidt, H. and Rathjen, F.G. (2010) 'Signalling mechanisms regulating axonal branching in vivo', *Bioessays*, 32(11), pp. 977-85.

Schnall-Levin, M., Rissland, O.S., Johnston, W.K., Perrimon, N., Bartel, D.P. and Berger, B. (2011) 'Unusually effective microRNA targeting

within repeat-rich coding regions of mammalian mRNAs', *Genome Research*, 21(9), pp. 1395-403.

Schneider, D.A. (2012) 'RNA polymerase I activity is regulated at multiple steps in the transcription cycle: recent insights into factors that influence transcription elongation', *Gene*, 493(2), pp. 176-84.

Schomburg, L. and Kohrle, J. (2008) 'On the importance of selenium and iodine metabolism for thyroid hormone biosynthesis and human health', *Molecular Nutrition & Food Research*, 52(11), pp. 1235-46.

Schopman, N.C., Heynen, S., Haasnoot, J. and Berkhout, B. (2010) 'A miRNA-tRNA mix-up: tRNA origin of proposed miRNA', *RNA Biol*, 7(5), pp. 573-6.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M. and Ragg, T. (2006) 'The RIN: an RNA integrity number for assigning integrity values to RNA measurements', *BMC Molecular Biology*, 7(1), p. 3.

Schweizer, U., Brauer, A.U., Kohrle, J., Nitsch, R. and Savaskan, N.E. (2004) 'Selenium and brain function: a poorly recognized liaison', *Brain Research - Brain Research Reviews*, 45(3), pp. 164-78.

Sharbati-Tehrani, S., Kutz-Lohroff, B., Bergbauer, R., Scholven, J. and Einspanier, R. (2008) 'miR-Q: a novel quantitative RT-PCR approach for the expression profiling of small RNA molecules such as miRNAs in a complex sample', *BMC Molecular Biology*, 9, p. 34.

Shimono, Y., Zabala, M., Cho, R.W., Lobo, N., Dalerba, P., Qian, D., Diehn, M., Liu, H., Panula, S.P., Chiao, E., Dirbas, F.M., Somlo, G., Pera, R.A.R., Lao, K. and Clarke, M.F. (2009) 'Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells', *Cell*, 138(3), pp. 592-603.

Siegfried, N., Irlam, J.H., Visser, M.E. and Rollins, N.N. (2012) 'Micronutrient supplementation in pregnant women with HIV infection', *Cochrane Database Syst Rev*, 14(3).

Slattery, M.L., Lundgreen, A., Welbourn, B., Corcoran, C. and Wolff, R.K. (2012) 'Genetic Variation in Selenoprotein Genes, Lifestyle, and Risk of Colon and Rectal Cancer', *PLoS One*, 7(5), p. e37312.

Slezak-Prochazka, I., Durmus, S., Kroesen, B.-J. and van den Berg, A. (2010) 'MicroRNAs, macrocontrol: regulation of miRNA processing', *Rna-A Publication of the Rna Society*, 16(6), pp. 1087-95.

Smith, C.M., Watson, D.I., Leong, M.P., Mayne, G.C., Michael, M.Z., Wijnhoven, B.P.L. and Hussey, D.J. (2011) 'miR-200 family expression is downregulated upon neoplastic progression of Barrett's esophagus', *World Journal of Gastroenterology*, 17(8), pp. 1036-44.

Song, B., Wang, Y., Xi, Y., Kudo, K., Bruheim, S., Botchkina, G.I., Gavin, E., Wan, Y., Formentini, A., Kornmann, M., Fodstad, O. and Ju, J. (2009) 'Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells', *Oncogene*, 28(46), pp. 4065-4074.

Speckmann, B., Bidmon, H.-J., Pinto, A., Anlauf, M., Sies, H. and Steinbrenner, H. (2011) 'Induction of glutathione peroxidase 4 expression during enterocytic cell differentiation', *Journal of Biological Chemistry*, 286(12), pp. 10764-72.

Squires, J.E., Stoytchev, I., Forry, E.P. and Berry, M.J. (2007) 'SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay', *Mol Cell Biol*, 27(22), pp. 7848-55.

Standart, N. and Jackson, R.J. (2007) 'MicroRNAs repress translation of m7Gppp-capped target mRNAs in vitro by inhibiting initiation and promoting deadenylation.[comment]', *Genes & Development*, 21(16), pp. 1975-82.

Stark, A., Lin, M.F., Kheradpour, P., Pedersen, J.S., Parts, L., Carlson, J.W., Crosby, M.A., Rasmussen, M.D., Roy, S., Deoras, A.N., Ruby, J.G., Brennecke, J., Hodges, E., Hinrichs, A.S., Caspi, A., Paten, B., Park, S.-W., Han, M.V., Maeder, M.L., Polansky, B.J., Robson, B.E., Aerts, S., van Helden, J., Hassan, B., Gilbert, D.G., Eastman, D.A., Rice, M., Weir, M., Hahn, M.W., Park, Y., Dewey, C.N., Pachter, L., Kent, W.J., Haussler, D., Lai, E.C., Bartel, D.P., Hannon, G.J., Kaufman, T.C., Eisen, M.B., Clark, A.G., Smith, D., Celniker, S.E., Gelbart, W.M. and Kellis, M. (2007) 'Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures', *Nature*, 450(7167), pp. 219-232.

Stefanovsky, V., Langlois, F., Gagnon-Kugler, T., Rothblum, L.I. and Moss, T. (2006) 'Growth Factor Signaling Regulates Elongation of

RNA Polymerase I Transcription in Mammals via UBF Phosphorylation and r-Chromatin Remodeling', *Molecular Cell*, 21(5), pp. 629-639.

Stegh, A.H. (2012) 'Targeting the p53 signaling pathway in cancer therapy – the promises, challenges and perils', *Expert Opinion on Therapeutic Targets*, 16(1), pp. 67-83.

Strachan T and P, R.A. (1999) 'DNA structure and gene expression', in *Human Molecular Genetics*. Wiley-Liss.

Stratmann, J., Wang, C.-J., Gnosa, S., Wallin, A., Hinselwood, D., Sun, X.-F. and Zhang, H. (2011) 'Dicer and miRNA in relation to clinicopathological variables in colorectal cancer patients', *BMC Cancer*, 11, p. 345.

Su, N., Wang, Y., Qian, M. and Deng, M. (2010) 'Combinatorial regulation of transcription factors and microRNAs', *BMC Systems Biology*, 4(1), p. 150.

Subramanyam, D. and Blelloch, R. (2011) 'From microRNAs to targets: pathway discovery in cell fate transitions', *Current Opinion in Genetics & Development*, 21(4), pp. 498-503.

Sun, G., Li, H. and Rossi, J.J. (2010) 'Sequence context outside the target region influences the effectiveness of miR-223 target sites in the RhoB 3'UTR', *Nucleic Acids Research*, 38(1), pp. 239-52.

Sun, J., Sun, Q. and Lu, S. (2011a) 'From selenoprotein to endochondral ossification: A novel mechanism with microRNAs potential in bone related diseases?', *Medical Hypotheses*, 77(5), pp. 807-811.

Sun, T., Wang, C., Xing, J. and Wu, D. (2011b) 'miR-429 modulates the expression of c-myc in human gastric carcinoma cells', *European Journal of Cancer*, 47(17), pp. 2552-9.

Sunde, R.A., Paterson, E., Evenson, J.K., Barnes, K.M., Lovegrove, J.A. and Gordon, M.H. (2008) 'Longitudinal selenium status in healthy British adults: assessment using biochemical and molecular biomarkers', *British Journal of Nutrition*, 99 Suppl 3, pp. S37-47.

Sunde, R.A. and Raines, A.M. (2011) 'Selenium regulation of the selenoprotein and nonselenoprotein transcriptomes in rodents', *Advances in nutrition (Bethesda, Md.)*, 2(2), pp. 138-150.

Sutherland, A., Kim, D.H., Relton, C., Ahn, Y.O. and Hesketh, J. (2010) 'Polymorphisms in the selenoprotein S and 15-kDa selenoprotein

genes are associated with altered susceptibility to colorectal cancer', *Genes Nutr*, 5(3), pp. 215-23.

Suzuki, H.I., Yamagata, K., Sugimoto, K., Iwamoto, T., Kato, S. and Miyazono, K. (2009) 'Modulation of microRNA processing by p53', *Nature*, 460(7254), pp. 529-533.

Tchernitsa, O., Kasajima, A., Schafer, R., Kuban, R.-J., Ungethum, U., Gyorffy, B., Neumann, U., Simon, E., Weichert, W., Ebert, M.P.A. and Rocken, C. (2010a) 'Systematic evaluation of the miRNA-ome and its downstream effects on mRNA expression identifies gastric cancer progression', *Journal of Pathology*, 222(3), pp. 310-9.

Tchernitsa, O., Kasajima, A., Schafer, R., Kuban, R.J., Ungethum, U., Gyorffy, B., Neumann, U., Simon, E., Weichert, W., Ebert, M.P. and Rocken, C. (2010b) 'Systematic evaluation of the miRNA-ome and its downstream effects on mRNA expression identifies gastric cancer progression', *J Pathol*, 222(3), pp. 310-9.

The University of Minnesota and Houston, T.U.o. (2008) *miRecords*. Available at: http://mirecords.umn.edu/miRecords/.

The UniProt, C. (2012) 'Reorganizing the protein space at the Universal Protein Resource (UniProt)', *Nucleic Acids Research*, 40(D1), pp. D71-D75.

Thirunavukkarasu, C., Premkumar, K., Sheriff, A.K. and Sakthisekaran, D. (2008) 'Sodium selenite enhances glutathione peroxidase activity and DNA strand breaks in hepatoma induced by N-nitrosodiethylamine and promoted by phenobarbital', *Molecular & Cellular Biochemistry*, 310(1-2), pp. 129-39.

Thomson, C.D., Campbell, J.M., Miller, J. and Skeaff, S.A. (2011) 'Minimal impact of excess iodate intake on thyroid hormones and selenium status in older New Zealanders', *Eur J Endocrinol*, 165(5), pp. 745-52.

Thomson, C.D., Wickens, K., Miller, J., Ingham, T., Lampshire, P., Epton, M.J., Town, G.I., Pattemore, P. and Crane, J. (2012) 'Selenium status and allergic disease in a cohort of New Zealand children', *Clinical & Experimental Allergy*, pp. n/a-n/a.

Thorne, N., Inglese, J. and Auld, D.S. (2010) 'Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology', *Chem Biol*, 17(6), pp. 646-57.

Tirosh, O., Levy, E. and Reifen, R. (2007) 'High selenium diet protects against TNBS-induced acute inflammation, mitochondrial dysfunction, and secondary necrosis in rat colon', *Nutrition*, 23(11–12), pp. 878-886.

Tsang, W.P., Ng, E.K.O., Ng, S.S.M., Jin, H., Yu, J., Sung, J.J.Y. and Kwok, T.T. (2010) 'Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer', *Carcinogenesis*, 31(3), pp. 350-358.

Tsuchida, A., Ohno, S., Wu, W., Borjigin, N., Fujita, K., Aoki, T., Ueda, S., Takanashi, M. and Kuroda, M. (2011) 'miR-92 is a key oncogenic component of the miR-17-92 cluster in colon cancer', *Cancer Science*, 102(12), pp. 2264-71.

Tsujiura, M., Ichikawa, D., Komatsu, S., Shiozaki, A., Takeshita, H., Kosuga, T., Konishi, H., Morimura, R., Deguchi, K., Fujiwara, H., Okamoto, K. and Otsuji, E. (2010) 'Circulating microRNAs in plasma of patients with gastric cancers', *Br J Cancer*, 102(7), pp. 1174-9.

Ueta, T., Inoue, T., Furukawa, T., Tamaki, Y., Nakagawa, Y., Imai, H. and Yanagi, Y. (2012) 'Glutathione peroxidase 4 is required for maturation of photoreceptor cells', *Journal of Biological Chemistry*, 287(10), pp. 7675-82.

Uthus, E., Begaye, A., Ross, S. and Zeng, H. (2010) 'The von Hippel-Lindau (VHL) Tumor-suppressor Gene is Down-regulated by Selenium Deficiency in Caco-2 Cells and Rat Colon Mucosa', *Biological Trace Element Research*, pp. 1-9.

Vainio, P., Gupta, S., Ketola, K., Mirtti, T., Mpindi, J.P., Kohonen, P., Fey, V., Perala, M., Smit, F., Verhaegh, G., Schalken, J., Alanen, K.A., Kallioniemi, O. and Iljin, K. (2011) 'Arachidonic acid pathway members PLA2G7, HPGD, EPHX2, and CYP4F8 identified as putative novel therapeutic targets in prostate cancer', *Am J Pathol*, 178(2), pp. 525-36.

van den Berg, A., Mols, J. and Han, J. (2008) 'RISC-target interaction: cleavage and translational suppression', *Biochimica et Biophysica Acta*, 1779(11), pp. 668-77.

Várallyay, É., Burgyán, J. and Havelda, Z. (2008) 'MicroRNA detection by northern blotting using locked nucleic acid probes', *Nature Protocols*, 3(2), pp. 190-6.

Vasudevan, S. (2012) 'Posttranscriptional Upregulation by MicroRNAs', *Wiley Interdisciplinary Reviews: RNA*, 3(3), pp. 311-330.

Wahid, F., Shehzad, A., Khan, T. and Kim, Y.Y. (2010) 'MicroRNAs: synthesis, mechanism, function, and recent clinical trials', *Biochimica et Biophysica Acta*, 1803(11), pp. 1231-43.

Wan, G., Mathur, R., Hu, X., Zhang, X. and Lu, X. (2011) 'miRNA response to DNA damage', *Trends in biochemical sciences*, 36(9), pp. 478-484.

Wang, B.-D., Kline, C., Pastor, D., Olson, T., Frank, B., Luu, T., Sharma, A., Robertson, G., Weirauch, M., Patierno, S., Stuart, J., Irby, R. and Lee, N. (2010a) 'Prostate apoptosis response protein 4 sensitizes human colon cancer cells to chemotherapeutic 5-FU through mediation of an NFkappaB and microRNA network', *Molecular Cancer*, 9(1), p. 98.

Wang, Q., Liu, Q.Y., Liu, Z.S., Qian, Q., Sun, Q. and Pan, D.Y. (2008) 'Lentivirus mediated shRNA interference targeting MAT2B induces growth-inhibition and apoptosis in hepatocelluar carcinoma', *World J Gastroenterol*, 14(29), pp. 4633-42.

Wang, Y.X., Zhang, X.Y., Zhang, B.F., Yang, C.Q., Chen, X.M. and Gao, H.J. (2010b) 'Initial study of microRNA expression profiles of colonic cancer without lymph node metastasis', *Journal of Digestive Diseases*, 11(1), pp. 50-4.

Wastney, M.E., Combs, G.F., Jr., Canfield, W.K., Taylor, P.R., Patterson, K.Y., Hill, A.D., Moler, J.E. and Patterson, B.H. (2011) 'A human model of selenium that integrates metabolism from selenite and selenomethionine', *J Nutr*, 141(4), pp. 708-17.

Weake, V.M. and Workman, J.L. (2010) 'Inducible gene expression: diverse regulatory mechanisms', *Nature Reviews Genetics*, 11(6), pp. 426-37.

Weekley, C.M., Aitken, J.B., Vogt, S., Finney, L.A., Paterson, D.J., de Jonge, M.D., Howard, D.L., Musgrave, I.F. and Harris, H.H. (2011) 'Uptake, distribution, and speciation of selenoamino acids by human cancer cells: X-ray absorption and fluorescence methods', *Biochemistry*, 50(10), pp. 1641-50.

Wiedmeier, J.E., Joss-Moore, L.A., Lane, R.H. and Neu, J. (2011) 'Early postnatal nutrition and programming of the preterm neonate', *Nutrition Reviews*, 69(2), pp. 76-82.

Willenbrock, H., Salomon, J., Sokilde, R., Barken, K.B., Hansen, T.N., Nielsen, F.C., Moller, S. and Litman, T. (2009) 'Quantitative miRNA expression analysis: comparing microarrays with next-generation sequencing', *Rna-A Publication of the Rna Society*, 15(11), pp. 2028-34.

Winter, J., Jung, S., Keller, S., Gregory, R.I. and Diederichs, S. (2009) 'Many roads to maturity: microRNA biogenesis pathways and their regulation', *Nature Cell Biology*, 11(3), pp. 228-34.

Wojewoda, M., Duszynski, J., Wieckowski, M. and Szczepanowska, J. (2012) 'Effect of selenite on basic mitochondrial function in human osteosarcoma cells with chronic mitochondrial stress', *Mitochondrion*, 12(1), pp. 149-55.

Wood, S.M., Beckham, C., Yosioka, A., Darban, H. and Watson, R.R. (2000) 'β-Carotene and selenium supplementation enhances immune response in aged humans', *Integrative Medicine*, 2(2–3), pp. 85-92.

Wszolek, M.F., Rieger-Christ, K.M., Kenney, P.A., Gould, J.J., Silva Neto, B., Lavoie, A.K., Logvinenko, T., Libertino, J.A. and Summerhayes, I.C. (2011) 'A MicroRNA expression profile defining the invasive bladder tumor phenotype', *Urologic Oncology*, 29(6), pp. 794-801.e1.

Wu, M., Kang, M.M., Schoene, N.W. and Cheng, W.-H. (2010) 'Selenium compounds activate early barriers of tumorigenesis', *Journal of Biological Chemistry*, 285(16), pp. 12055-62.

Wu, M., Wu, R.T., Wang, T.T. and Cheng, W.H. (2011a) 'Role for p53 in selenium-induced senescence', *J Agric Food Chem*, 59(21), pp. 11882-7.

Wu, Y., Wang, X.-f., Mo, X.-a., Li, J.-m., Yuan, J., Zheng, J.-o., Feng, Y. and Tang, M. (2011b) 'Expression of laminin beta1 and integrin alpha2 in the anterior temporal neocortex tissue of patients with intractable epilepsy', *International Journal of Neuroscience*, 121(6), pp. 323-8.

Wu, Y., Wang, X.-f., Mo, X.-a., Sun, H.-b., Li, J.-m., Zeng, Y., Lin, T., Yuan, J., Xi, Z.-q., Zhu, X. and Zheng, J.-o. (2008) 'Expression of

laminin beta1 in hippocampi of patients with intractable epilepsy', *Neuroscience Letters*, 443(3), pp. 160-4.

Xia, Y., Hill, K.E., Byrne, D.W., Xu, J. and Burk, R.F. (2005) 'Effectiveness of selenium supplements in a low-selenium area of China', *Am J Clin Nutr*, 81(4), pp. 829-34.

Xu, Y.-W., Wang, B., Ding, C.-H., Li, T., Gu, F. and Zhou, C. (2011) 'Differentially expressed micoRNAs in human oocytes', *Journal of Assisted Reproduction & Genetics*, 28(6), pp. 559-66.

Xun, P., Liu, K., Steven Morris, J., Daviglus, M.L., Stevens, J., Jacobs, D.R. and He, K. (2010) 'Associations of Toenail Selenium Levels With Inflammatory Biomarkers of Fibrinogen, High-Sensitivity C-Reactive Protein, and Interleukin-6', *American Journal of Epidemiology*, 171(7), pp. 793-800.

Yagublu, V., Arthur, J.R., Babayeva, S.N., Nicol, F., Post, S. and Keese, M. (2011) 'Expression of selenium-containing proteins in human colon carcinoma tissue', *Anticancer Research*, 31(9), pp. 2693-8.

Yamakuchi, M., Lotterman, C.D., Bao, C., Hruban, R.H., Karim, B., Mendell, J.T., Huso, D. and Lowenstein, C.J. (2010) 'P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), pp. 6334-9.

Yamakuchi, M., Yagi, S., Ito, T. and Lowenstein, C.J. (2011) 'MicroRNA-22 regulates hypoxia signaling in colon cancer cells', *PLoS ONE [Electronic Resource]*, 6(5), p. e20291.

Yeung, M.L., Bennasser, Y., Le, S.Y. and Jeang, K.T. (2005) 'siRNA, miRNA and HIV: promises and challenges', *Cell Res*, 15(11-12), pp. 935-946.

Yin, J.Q., Zhao, R.C. and Morris, K.V. (2008) 'Profiling microRNA expression with microarrays', *Trends in Biotechnology*, 26(2), pp. 70-6.

Yokoi, T. and Nakajima, M. (2011) 'Toxicological Implications of Modulation of Gene Expression by MicroRNAs', *Toxicological Sciences*, 123(1), pp. 1-14.

Yu, D.C., Waby, J.S., Chirakkal, H., Staton, C.A. and Corfe, B.M. (2010) 'Butyrate suppresses expression of neuropilin I in colorectal cell lines through inhibition of Sp1 transactivation', *Mol Cancer*, 9, p. 276.

Yu, J., Li, A., Hong, S.-M., Hruban, R.H. and Goggins, M. (2012) 'MicroRNA alterations of pancreatic intraepithelial neoplasias', *Clinical Cancer Research*, 18(4), pp. 981-92.

Zeng, H., Jackson, M.I., Cheng, W.-H. and Combs, G.F., Jr. (2011) 'Chemical form of selenium affects its uptake, transport, and glutathione peroxidase activity in the human intestinal Caco-2 cell model', *Biological Trace Element Research*, 143(2), pp. 1209-18.

Zhang, H., Li, W., Nan, F., Ren, F., Wang, H., Xu, Y. and Zhang, F. (2011a) 'MicroRNA expression profile of colon cancer stem-like cells in HT29 adenocarcinoma cell line', *Biochemical and Biophysical Research Communications*, 404(1), pp. 273-278.

Zhang, H., Li, W., Nan, F., Ren, F., Wang, H., Xu, Y. and Zhang, F. (2011b) 'MicroRNA expression profile of colon cancer stem-like cells in HT29 adenocarcinoma cell line', *Biochem Biophys Res Commun*, 404(1), pp. 273-8.

Zhang, J., Zhang, H., Liu, J., Tu, X., Zang, Y., Zhu, J., Chen, J., Dong, L. and Zhang, J. (2012) 'miR-30 inhibits TGF-β1-induced epithelial-to-mesenchymal transition in hepatocyte by targeting Snail1', *Biochemical and Biophysical Research Communications*, 417(3), pp. 1100-1105.

Zhang, X.-d., Qin, Z.-h. and Wang, J. (2010) 'The role of p53 in cell metabolism', *Acta Pharmacol Sin*, 31(9), pp. 1208-1212.

Zhang, Y., Zhang, R. and Su, B. (2009) 'Diversity and evolution of MicroRNA gene clusters', *Science in China*, Series C, Life Sciences. 52(3), pp. 261-6.

Zhao, J.-J., Yang, J., Lin, J., Yao, N., Zhu, Y., Zheng, J., Xu, J., Cheng, J.Q., Lin, J.-Y. and Ma, X. (2009) 'Identification of miRNAs associated with tumorigenesis of retinoblastoma by miRNA microarray analysis', *Childs Nervous System*, 25(1), pp. 13-20.

Zhao, W., Blagev, D., Pollack, J.L. and Erle, D.J. (2011) 'Toward a systematic understanding of mRNA 3' untranslated regions', *Proceedings of the American Thoracic Society*, 8(2), pp. 163-6.

Zimmermann, M.B. and Kohrle, J. (2002) 'The impact of iron and selenium deficiencies on iodine and thyroid metabolism: biochemistry and relevance to public health', *Thyroid*, 12(10), pp. 867-78.