Downregulation of miRNA expression in malignant germ cell tumours: mechanism and functional significance

Marta Ferraresso

Department of Pathology

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

St. Edmund's College   September 2018

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except when specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except when specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except when specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Marta Ferraresso

September 2018

'Above all, don't fear difficult moments. The best comes from them'

Rita Levi Montalcini (1909-2012)

Acknowledgements

This thesis would not be possible without the contribution and encouragement of many people and I would like to send my warmest thank to them all.

Firstly, I would like to thank Professor Nicolas Coleman and Dr Matthew Murray for giving me the opportunity to undertake this PhD and for all their support and guidance.

I am also extremely grateful for and humbled by the support and friendship of all the Coleman laboratory past and present. In particular, Dr Cinzia Scarpini, who was always there during my PhD to offer me guidance on experimental approaches, PhD direction, thesis writing and most importantly moral support. I also need to say a huge thank to Dawn Ward and Danita Pearson, you were never too busy to offer guidance and give me your time – thank you for looking after me and being such wonderful influences. Thanks to all the Postdoc that were always available to help me during these years: Maria, Marta, Ian, Emma Bell, Lorena and Steve. In particular a huge thank to Steve to repeatedly explain me the importance of statistics. Thanks to other PhD student sharing this experience with me: Shivani, Emma, Katie, Jenny, Luz, Justina and Valtteri. In particularly Shivani, for going through the PhD experience with me and support each other's.

I would also like to say a huge thank you to my collaborators in Anton Enright group, in particular to Keaan Amin, which he spent a lot of time explaining me and helping me in the bionformatical work. I would also like to say thank you to my mentor Suzanne Turner for the words of encouragement throughout.

I have to say a huge thank you to my best friend Ala, which I met at the beginning of my PhD and she helped me going through these 4 years by smiling, laughing and dreaming of our future every day.

I also need to say a huge thank to all my Eddie's friends, which made my staying in Cambridge extremely fun and interesting, my home friends even if they haven't completely understood what I am doing and finally Jack that never stopped believing in me. Finally, I have to say a

Presentations

Oral presentations:

'*Understanding the role of under-expressed microRNAs in malignant germ cell tumours (mGCTs)'*

Presented at:

- St. Edmund's Student Conference, February 2016 – Cambridge, UK

Poster presentations:

'*Understanding the role of under-expressed microRNAs in malignant germ cell tumours (mGCTs)'*

Presented at:

- Department of Pathology Symposium, May 2016 – Cambridge, UK
- eSCAMPS Lives in Science Symposium, July 2016 – Cambridge, UK
- BACR Non-coding RNAs in Cancer and Development, May 2017 – Cambridge, UK

'Significance of miR-100-5p/99a-5p and miR-125b-5p under-expression in Malignant Germ Cell Tumours'

Presented at:

- SHON Conference: Accelerating Translation of Pediatric Cancer Research, February 2018 – London, UK
- EACR25 European Association for Cancer Research, July 2018 – Amsterdam, ND

Abstract

Germ cell tumours (GCTs) are clinically and pathologically heterogeneous neoplasms that arise at gonadal (testicular/ovarian) and extra-gonadal sites. The chemotherapy burden for patients with malignant germ cell tumours (mGCTs) that require treatment results in substantial long-term side-effects, and, furthermore, poor-risk patients have <50% survival. Consequently, identifying common molecular changes and novel therapeutic targets in mGCTs is of major clinical importance.

MicroRNAs are short, non-protein coding RNAs that regulate gene expression. We previously showed that miR-99a-5p/-100-5p and miR-125b-5p are among the most frequently under-expressed microRNAs in mGCTs, regardless of anatomical site, histological type or patient age. The present study investigates the upstream causes and downstream consequences of such under-expression.

The mature form of miR-125b-5p is the product of two genomic loci, which form a cluster with either miR-99a-5p (on chromosome 21q) or miR-100-5p (on chromosome 11q). MiR-99a-5p/-100-5p share identical 'seed' regions (at nucleotide positions 2-7), which determine their mRNA targets. Cross-reactivity experiment revealed that both miR-99a-5p and miR-100-5p probes were highly cross-reactive to each other's target (from 91% to 95%), indicating functional overlap. Linear regression analysis of qRT-PCR data reveals a strong positive correlation between miR-99a-5p/-100-5p and miR-125b-5p levels (R^2=0.989) in mGCTs, strongly suggesting co-regulation.

Primary microRNA transcripts (pri-miR-99a/-100 and pri-miR-125b), and other genes that co-localise to these miRNA clusters (e.g. *BLID* on chromosome 11), were quantified by RT-qPCR in four representative cell lines - TCam2, 1411H, 2102Ep, and GCT44 - which were derived from a range of common histological types of mGCTs. A significant down-regulation ($p<0.0001$) of all primary transcripts was observed, suggesting transcriptional repression of the entire cluster regions. Treatment of the cell lines with 5'-azacytidine resulted in significant up-regulation of all three miRNAs ($p<0.002$), as well as *BLID* ($p<0.02$). The methylation status of

potential CpG islands at the region of interest on chromosome 11 and chromosome 21 was therefore investigated by Pyrosequencing. Significant hyper methylation was found in 2102Ep, 1411H and GCT44 cell lines, suggesting that the miR-99a-5p/-100-5p and miR-125b-5p clusters are likely transcriptionally silenced by DNA methylation.

To assess the functional relevance of these microRNAs in GCT progression, co-transfection of microRNA mimics (8.3 nM miR-99a-5p/-100-5p + 8.3 nM miR-125b-5p) was performed. A significant decrease in cell growth was seen in 1411H (p<0.01) and TCam2 (p<0.03) cells. To identify the mimics' downstream mRNA targets, HumanHT-12 v4 Expression Bead Chip (Illumina) mRNA arrays were used and data analysed using Sylamer. This analysis showed that mimic-treated cells were enriched in downregulated genes involved in pro-proliferative mechanisms.

Among those, further functional characterisation focussed in particular on *TRIM71*, *FGFR3*, *E2F7* and *LIN28A*. Moreover, restoring miR-99a-5p/-100-5p and miR-125b-5p in TCam2 cells also resulted in G0-G1 accumulation, consistent with a cell cycle effect.

These data support a functionally important role for miR99a-5p/-100-5p and miR-125b-5p in GCT progression. They also raise the possibility of a therapeutic replenishment approach for treating these, and potentially other, tumours.

Contents

List of Figures

List of Tables

Nomenclature

5-aza - 5-azacytidine

ACTB - beta-actin

ARID 3A - AT-rich interaction domain 3A

ARID 3B - AT-rich Interaction domain 3B

Bp - base pairs

BEP - cisplatin, etoposide, bleomycin

BSA - bovine serum albumin

cDNA - complementary DNA

CDK - cyclin dependent kinase

CHC – choriocarcinoma

ChIP - chromatin immunoprecipitation

CNS - central nervous system

CO_2 - carbon dioxide

Ct - threshold cycle

CTDSPL - carboxy-terminal domain, small phosphatase like

dH_2O - distilled water

DMEM - Dulbecco's modified Eagle's medium

DMSO - dimethyl sulfoxide

DNMT – DNA methyltransferases

dNTP - deoxynucleotide triphosphate

E2F7 - E2F Transcription Factor 7

EC - embryonal carcinoma

ECL - enhanced standard chemiluminescence

EdU - 5-ethynyl-2'-deoxyuridine

EMBL-EBI - European Molecular Biology Laboratory-European Bioinformatics Institute

EtBr - ethidium bromide

FBS - fetal bovine serum

FGFR 3 - fibroblast growth factor receptor 3

FSC - forward scatter

Fw - forward primer

GCT - germ cell tumour

gDNA - genomic DNA

GAPDH - gliceraldeide-3-fosfato deidrogenasi

GEPIA - Gene Expression Profiling Interactive Analysis

GFP - green fluorescent protein

GMEM - Glasgow minimum essential medium

GO - Gene Ontology

GTEx - Genotype-Tissue Expression

GUSB - glucoronidase beta

HFF-1 - cell line derived from normal fibroblast

HK - housekeeping gene

HMBS - hydroxymethylbilane synthase

hr - hours

IT - immature teratoma

JEB - carboplatin, etoposide, bleomycin

Kb - Kilobase pairs

kDa - KiloDaltons

KD - knock down

KO - knockout

lincRNA - long intergenic noncoding RNA

LOH - loss of heterozygosity

miRNA - microRNA

MNC - mimic negative control

mRNA - messenger RNA

mTOR - mammalian target of rapamycin

MT - mature teratoma

NA - not available

NCX - normal cervix

n/s - not significant

nt - nucleotide

ORF - open reading frame

P-body - processing body

PBS - phosphate-buffered saline

PCR - polymerase chain reaction

Pi - propidium iodide

PIC - protease inhibitor cocktail

pre-miRNA - precursor microRNA

pri-miRNA - primary microRNA

PVB - cisplatin, vincristine, bleomycin

QC - quality control

qRT-PCR - quantitative reverse transcriptase polymerase chain reaction

RIN - RNA integrity score

RISC - RNA-induced silencing complex

RNAi - RNA interference

RPE - epithelial cell immortalized with hTERT derived from normal retina

RPL13A - ribosomal protein L13A

RT - reverse transcription

RTase - reverse transcriptase

Rw - reverse primer

SCR - seed complementary region

SDS PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEER - Surveillance, Epidemiology and End Results

Sem - seminoma

SEM - standard error of the mean

siRNA - short interfering RNA

snRNA - small nuclear RNA

SSC - side scatter

SSSS - single-sum significance score

TBS - tris-buffered saline

TCGA - The Cancer Genome Atlas

TF - transcription factor

TGCT - testicular GCT

TRIM 71 - tripartite motif containing 71

UTR - untranslated region

WHO - World Health Organisation

XPO5 - exportin 5

yrs - years

YST - yolk sac tumour

YWHAZ - tyrosine 3- / tryptophan 5- monooxygenase activation protein, zeta polypeptide

Chapter 1   Introduction

1.1   Preface

The work presented in this thesis focuses on better understanding the function that microRNAs (miRNAs) have in the development and progression of malignant germ cell tumours (mGCTs). Despite their extensive clinical and pathological heterogeneity, all mGCTs are believed to originate from primordial germ cells (PGCs) (Teilum, 1965). To date, no one has identified a common event, or biological abnormalities, that combines all the different subtypes of mGCTs. In 2010, members of the Coleman lab undertook a global miRNA study of mGCTs. Interestingly, they observed a remarkable separation between mGCTs and non-malignant controls (comprising the benign subtype teratomas and normal gonadal controls) by global profiling, regardless of patient age or histological subtype (Palmer *et al*., 2010). These findings offered the potential to be harnessed for prospective future therapies as, for the first time, a common biological in malignant GCTs has been found (Palmer *et al*., 2010).

GCTs can affect patients of all ages and the most common treatment options depend mainly on type and stage of tumour. Stage I GCT patients are treated with surgery, while for stage II-IV the only available treatment so far is chemotherapy. The usage of chemotherapy has improved cancer survival by almost 70% in the last 30 to 40 years (Einhorn and Williams, 1979). However, it has to be taken into consideration that the risks of having long-term and late side effects from chemotherapy treatment is almost 50% (Travis *et al*., 2005). One example of these effects comes from the usage of cisplatin as a chemotherapy agent. This chemical compound

impairs hearing abilities, especially in children, making the patient at risk of deafness. Other late effects due to chemotherapy agents include reproductive and sexual development problems, growth and hormone problems, lung and breathing problems. It has been reported from the National Health Interview Survey that approximately 1 in 4 cancer survivors reports a decrease in quality of life due to physical problems (Weaver *et al*., 2012). The need for more targeted and less toxic therapies is thus critical in maintaining a good quality of life to the patient, particularly given the young median age that GCTs affect patients (Murray MJ, Fern LA, Stark DP, Eden TO, 2009).

The discovery of dysregulated miRNAs expression as a potential common biological abnormality among GCTs (Palmer *et al*., 2010; Murray *et al*., 2013), encouraged us to further study them in detail at a molecular level. Needless to say, understanding and extrapolating their mechanistic framework could help in utilising them as plausible therapeutics in cancer treatment.

1.2   Overview of germ cell tumours (GCTs)

Germ cell tumours (GCTs) are a heterogeneous group of benign and malignant tumours that vary by clinical presentation, clinical course and tumour histology (Penn, Jenney and Nicholson, 2014). GCTs are believed to arise from totipotent PGCs (Teilum, 1965). During normal development, PGCs formation is observed in the embryo after implantation and gastrulation, when the epiblast, endoderm, mesoderm and ectoderm are first established (Matsui and Okamura, 2005). At this point, following molecular signalling received from the yolk sac, a small population of pluripotent stem cells are physically set apart from the yolk sac of the embryo and germ line specification occurs (Ying *et al*., 2000). The exact events regulating germ line specification remain unclear in humans because of the very limited access to early embryonic samples. In normal development, during the first trimester of human gestation, germ cells start to migrate through the hindgut and dorsal mesentery to the lateral mesoderm where they form gonocytes, oogonia in females or spermatogonia in males, within

the futures ovaries and testes, respectively (Fujimoto, Miyayama and Fuyuta, 1977). During this journey, modification or aberrant migration (including migration to the brain, mediastinum or other location) of these cells may occur, resulting in the development of GCTs (Figure 1.1).
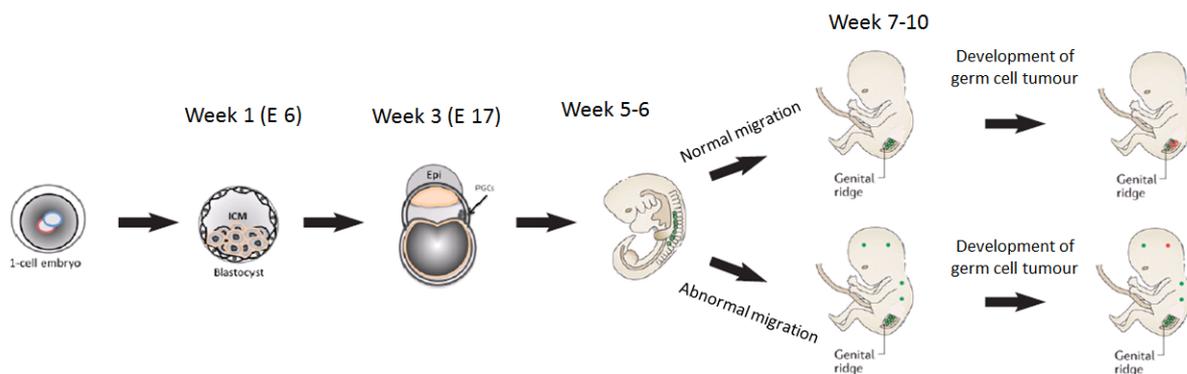
Figure 1.1: **Developmental cycle of mammalian germ cells**. Fertilisation of oocytes by sperm, promotes the formation of a zygote, which undergoes several events to form a blastocyst. Primordial germ cells (PGCs) derive from embryonic stem cells which are contained in the inner cell mass. Following normal migrations, oogonia and spermatogonia will develop in ovaries and testis, respectively. During this event formation of GCTs can occur. Following abnormal migration, cells that will survive to the hostile environment will develop into GCTs.

1.2.1 Histological heterogeneity of GCTs

Depending on the location and stage of PGCs differentiation, when the aforementioned modifications occur, different subtypes of tumour may arise (Figure 1.2). If a tumour develops during the somatic stages of differentiation, a teratoma is formed, which is considered a benign form of GCT, certainly in paediatric practice. If a tumour develops during embryonal-like differentiation, embryonal carcinoma (EC) arises, while extra-embryonic differentiation leads to yolk sac tumour (YST) or choriocarcinoma (CHC), respectively. The latter group of tumours, show embryonal differentiation, yolk sac differentiation and trophoblastic/placental-like differentiation, respectively, and collectively are named non-seminomatous tumour (Murray, Nicholson and Coleman, 2015a). Lastly, if PGCs remain undifferentiated when a

tumour arises, a germinoma results. The exact terminology of these tumours depends on tumour site - seminoma in the testis, dysgerminoma in the ovary and germinoma at other extragonadal sites, including the brain (Murray and Nicholson, 2010), and collectively are named seminoma tumours (Murray, Nicholson and Coleman, 2015b).
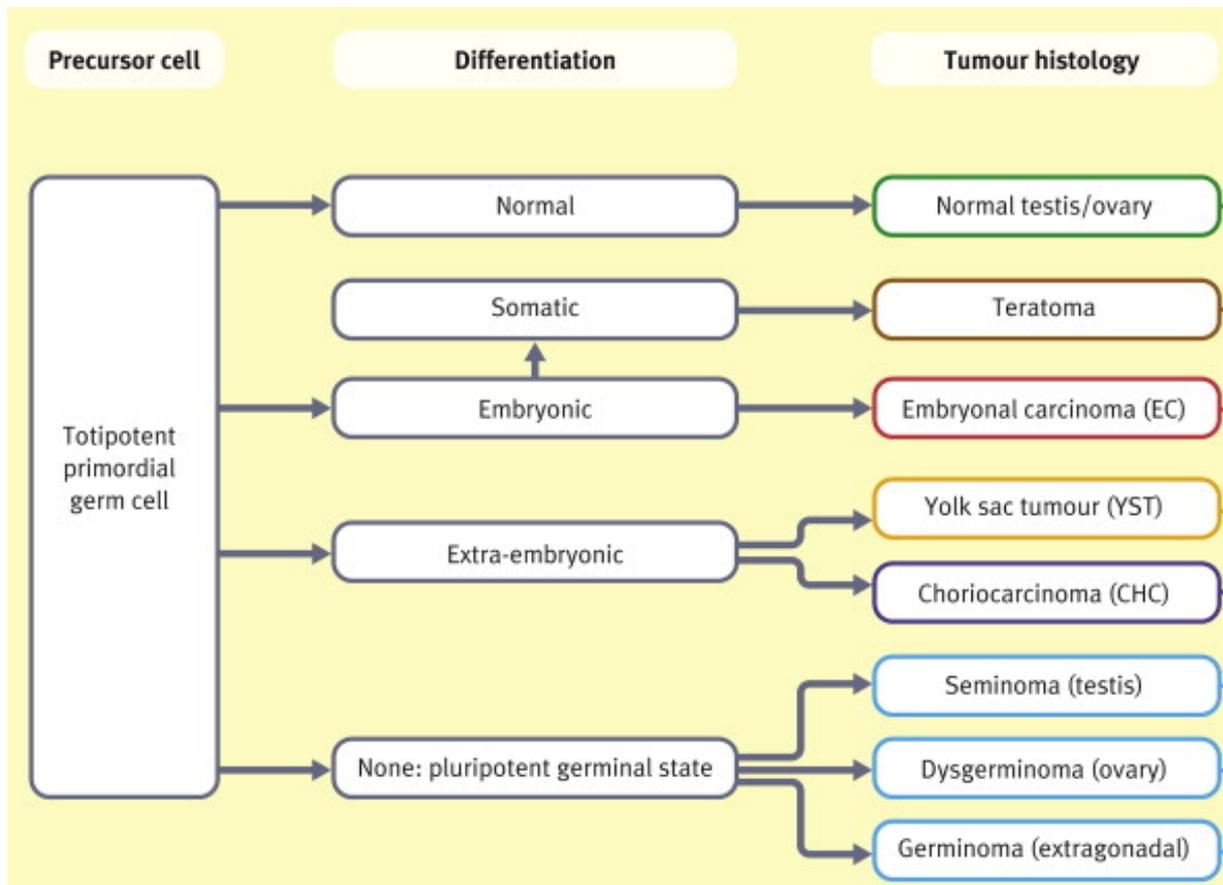
Figure 1.2: **Derivation of the various sub-categories of germ cell tumours.** Adapted from (Murray and Nicholson, 2010).

1.3   Epidemiology of GCTs

The heterogeneity of GCTs is also demonstrated in the age of onset in which the disease develops. GCTs are rare tumours that can present in all age groups - from newborns to late

adulthood. GCTs account for 4% of paediatric cancer (patients aged 0 to 14 years), 13.8% of early adulthood cancer (patients aged 15 to 24 years) (Murray and Nicholson, 2010; Arora *et al.*, 2012; Kaatsch *et al.*, 2015) and 95% of testicular cancer in adults and young adults males (M Dror Michaelson, MD, PhD William K Oh, 2018). Interestingly, around 95% of GCTs in adults are gonadal, whereas in children gonadal and extragonadal GCTs occur at similar frequencies (Murray and Nicholson, 2010).
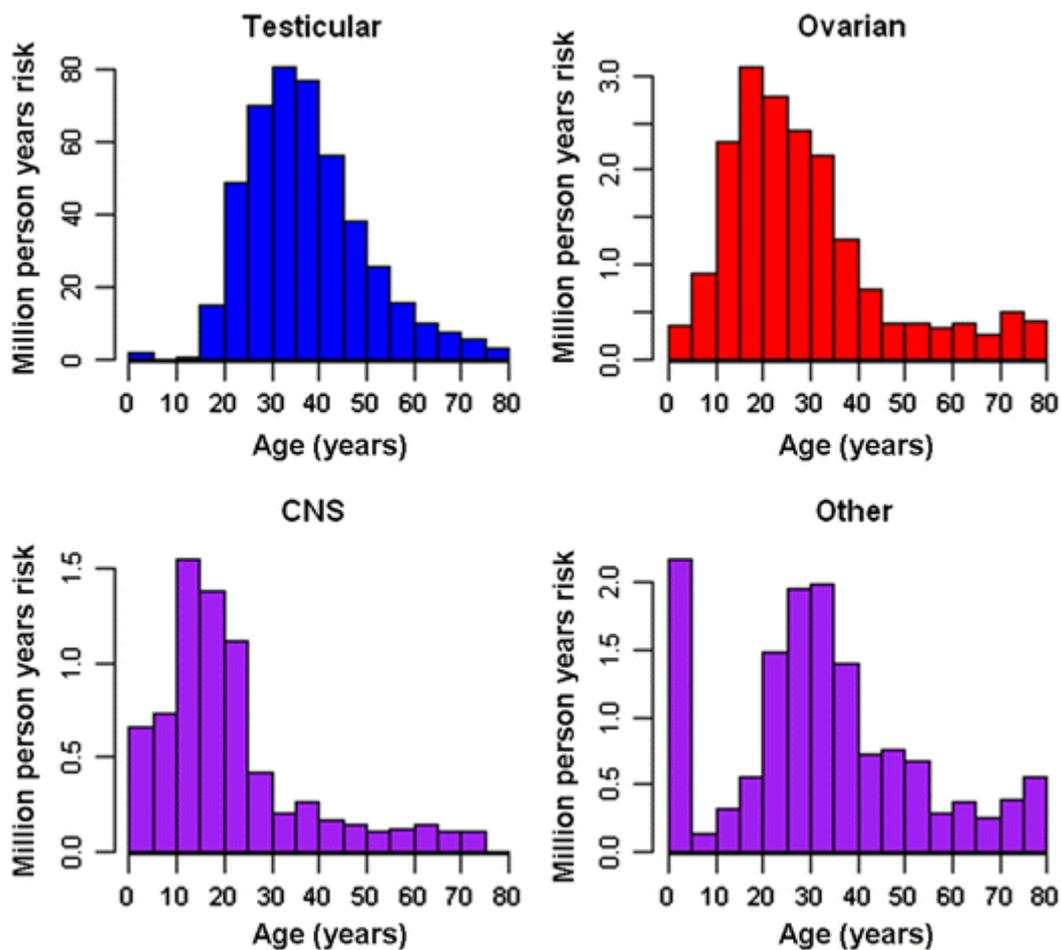
Figure 1.3: **Incidence rates of germ cell tumour by anatomical site.** Number of cases, per million-person year, aged 0-79 years in England and Wales, 1995-2003. Taken from (Murray *et al.*, 2009).

In children aged <15 years, gonadal GCTs are common in both female [with mGCTs developing in the ovaries (30%)] and male [with mGCTs developing in the testes (20%)] patients. Among the extragonadal GCTs, the most common location of tumour formation is the central nervous system (CNS), in both males and females (20% of extragonadal mGCTs), followed by mediastinum and thorax in males and abdomen and pelvis in females (Figure 1.4) (Murray and Nicholson, 2010).
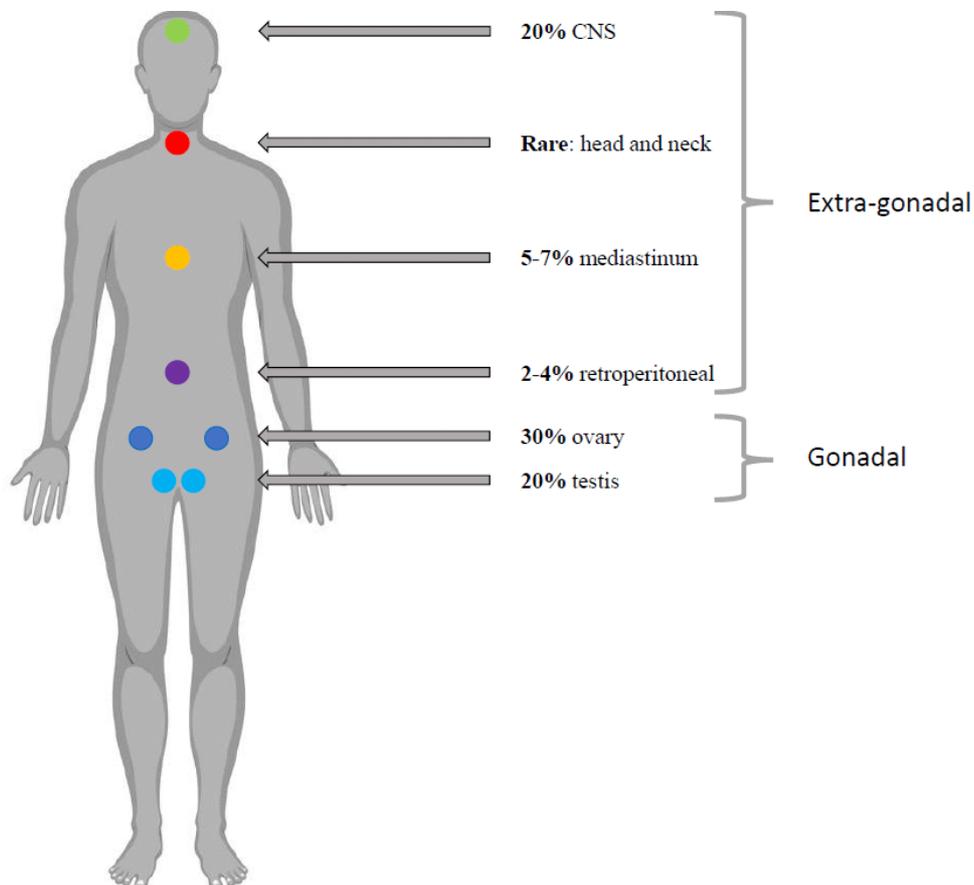
Figure 1.4: **Site distribution and frequency of germ cell tumours in children.** The graph shown the incidence rate of mGCTs in children depending on their origin, gonadal or extragonadal. Adapted from (Murray and Nicholson, 2010).

1.3.1 Demographic differences in the anatomical site and age of formation of GCTs.

The data presented in this section, is divided into two parts. The first one focuses on the demographic differences of mGCTs among paediatric patients, while the second part focuses on mGCTs patients in on all age groups.

Data from the National Cancer Institute's Surveillance, Epidemiology and End Results Program (SEER) evaluated trends for GCTs in boys and girls under 19 years old in the United States (Poynter, Amatruda and Ross, 2010). An incidence peak was observed before the age 1 and from the age of 15-19 years among boys and girls. Before the age of 4, extragonadal tumour was the only GCT detected in girls, while in boys' extragonadal and gonadal tumours were equally represented. After the age of 10, for both girls and boys, the majority of the tumours were localized in the ovaries and testes, respectively (Poynter, Amatruda and Ross, 2010). A similar study, conducted in Germany, analysed the incidence rates, time trends and survival for 1366 GCTs in children aged 0 to 14 years old. An incidence peak was observed in children aged <1 year for both girls and boys, confirming the previous data analysis (Kaatsch *et al.*, 2015). Although, when the range group 10 to 14 years was considered, the data showed a dominance in the extragonadal mGCT in males mostly in the central nervous system, not previously observed (Kaatsch *et al.*, 2015).

The former study also looked at mGCTs incidence between Caucasian Americans and Black Americans. Among the first group, boys and girls aged <9 had no difference in the incidence of GCTs, while between 10 and 19 years, girls had lower incidence than boys. Among the second group, black girls had higher incidence of GCT malignancies in both age range when compared with black boys. This variance was probably due to the lower incidence of gonadal tumours in black boys in the 10-19 age group compared with white boys (Poynter, Amatruda and Ross, 2010).

No differences were found in the overall GCTs incidence by tumour location or histology in non-Hispanic vs Hispanic boys' and girls' age 0-9. In the age group 10-19, higher incidence of

GCTs were noticed in the Hispanic population when compared with non-Hispanic population of the same sex. Hispanic girls resulted in higher gonadal incidence of GCTs, while Hispanic boys have higher overall GCTs incidence (Figure 1.5) (Poynter, Amatruda and Ross, 2010). Incidence rate is calculated as the number of new cases of disease divided by the time that each person was observed, totaled for all persons.
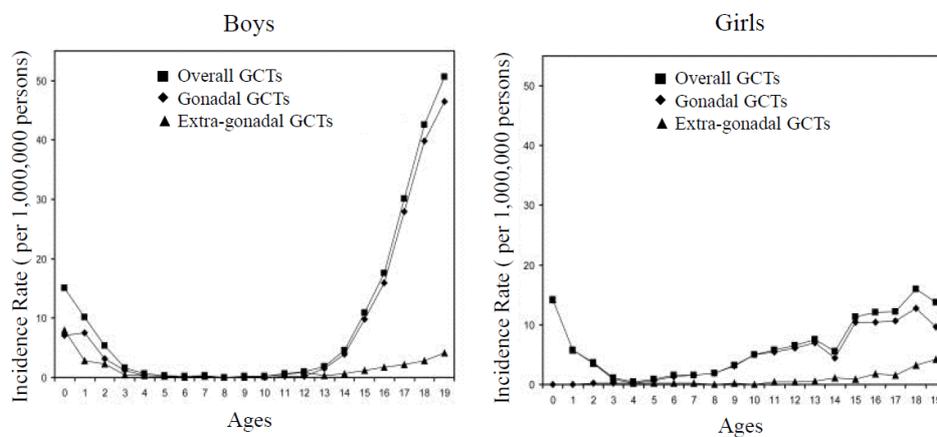
Figure 1.5: **Incidence of paediatric GCT by tumour location in boys and girls in the SEER registry**. Taken from (Poynter, Amatruda and Ross, 2010).

The most common GCTs malignancies in adults are gonadal GCTs (95%) -testicular cancer or ovarian cancer in males and females, respectively (Slap, Lehmann and Biro, 2008; Patricia M Baker, 2009). A recent study has analysed data collected from European population-based cancer registries (CRs) (Trama and Berrino, 2017). mGCTs incidence rate in Europe, was found to be equal to 34 cases every 1.000.000 people between 2000 to 2007, with marked differences between males (64/1.000.000) and females (4/1.000.000). Similarly, in the USA, between 1973 and 2007, the incidence rate for white males was 56/1.000.000 while for females was 3.5/1.000.000 (Stang *et al.*, 2012). This high discrepancy in female and male rate, is due to the fact that TGCTs are the most commonly occurring malignancy among white males between ages 15 and 44 years (McGlynn and Cook, 2009; Ghazarian *et al.*, 2015a). The incidence of testicular GCT (TGCT) was monitored between 1992 and 2011 from the SEER and racial/ethnicity-based comparison was assessed. Interestingly, the incidence of TGCT was

highest among non-Hispanic white men, (69.7 per 1.000.000 man-years) followed by American Indian/Alaska Native (46 per 1.000.000 man-years), Hispanic white (41 per 1.000.000 man-years), Asian/Pacific Islander (19 per 1.000.000 man-years), and black (12 per 1.000.000 man-years) men (Ghazarian *et al*., 2015b).

Understanding the differences observed among age, sex and ethnicity groups, needs further investigation.

1.3.2 Increasing trend of GCT incidence and possible explanations

Overall incidence of malignant GCTs, in both children and adults ranged between 0 and 84 years, underwent an increase from 1979 to 2003 (Arora *et al*., 2012). Across the patient aged 0 to 84 years, the increase was detected for germinoma (mainly testicular and central nervous system mGCTs) and non-germinoma (mainly ovarian mGCTs) (Arora *et al*., 2012). No difference was detected overall for mGCTs of mediastinum, thorax, abdomen or pelvis with only an increase seen for non-germinomas of the abdomen and pelvis in the age group 0 to 9 years (Arora *et al*., 2012). Most of this increase was observed for those aged between 10 to 49 years (Arora *et al*., 2012). A similar study conducted on mGCTs extragonadal and gonadal in Finland found a significant increase in testicular mGCTs in males between 15 and 44 years of age, while in females the incidence of ovarian mGCTs remained stable during the time from from 1969 and 2008 (Pauniaho *et al*., 2012), as shown from a previous study on ovarian cancer carried out in Osaka, Japan (1975-1998) (Ioka *et al*., 2003). In agreement with the previous study, no differences were noticed in the incidence of non-gonadal mGCTs between 1969 and 2008 (Pauniaho *et al*., 2012).

The highest testicular tumour incidence has been observed in many countries in the 21st century, such as western and northern Europe, Australia, New Zealand and USA, while a lower incidence has been noted in Asia and Africa (Manecksha and Fitzpatrick, 2009; Rosen *et al*., 2011; McGlynn and Trabert, 2012; Ghazarian *et al*., 2015b). No definitive answer has been found yet for this increase. Geographic and ethnic discrepancies suggest that environmental and genetic factors might contribute in causing testicular GCTs. Few studies have focused on

looking at differences in the environment in which males with testicular mGCTs are exposed. Interestingly, an interruption in the tumour incidences in the generation of men born during the World War II was clearly seen in Nordic countries in the 1940s, and in France and Switzerland during the 1930s. This was likely due to changes in the everyday habits mainly of food and tobacco consumption (Bergström *et al*., 1996; Richiardi *et al*., 2004). Sedentary behavior, consumption of fats, milk and dairy products, have all been shown to be associated with testicular cancer risk, all of which were not commonly present during World War II. Another environmental factor that can differ in time and in geographic locations is pollution. Blood screening were undertaken to look for the presence of chemical compounds that could have a putative effect on the observed increased incidence (Hardell *et al*., 2003). No differences were found among the patients' blood. However, when the blood derived from the patients' mothers was tested, they found increased traces of polychlorinated biphenyls (PCBs), hexachlorobenzene, and some chlordanes in the blood of cancer patient's mothers compared with the blood of healthy patients' mothers. Some of these chemicals are endocrine disrupters and have been found in the animal kingdom to reverse gonadal sex or create abnormalities of reproductive development (Bergeron, Crews and McLachlan, 1994; Guillette *et al*., 1994). The exposure to these chemicals could plausibly be involved in testicular carcinogenesis (Hardell *et al*., 2003).

1.4   Classification of GCTs

The classification of GCTs has been a contentious issue. There are two main classification systems used, namely, the World Health Organisation (WHO) and the British Testicular Tumor Panel (BTTP). To date, the most prominent classification used is derived from the WHO (97%), with only a small minority using the BTTP alone (<1%) or the BTTP in tandem with the WHO classification (2%) (Figure 6) (Williamson *et al.*, 2017).
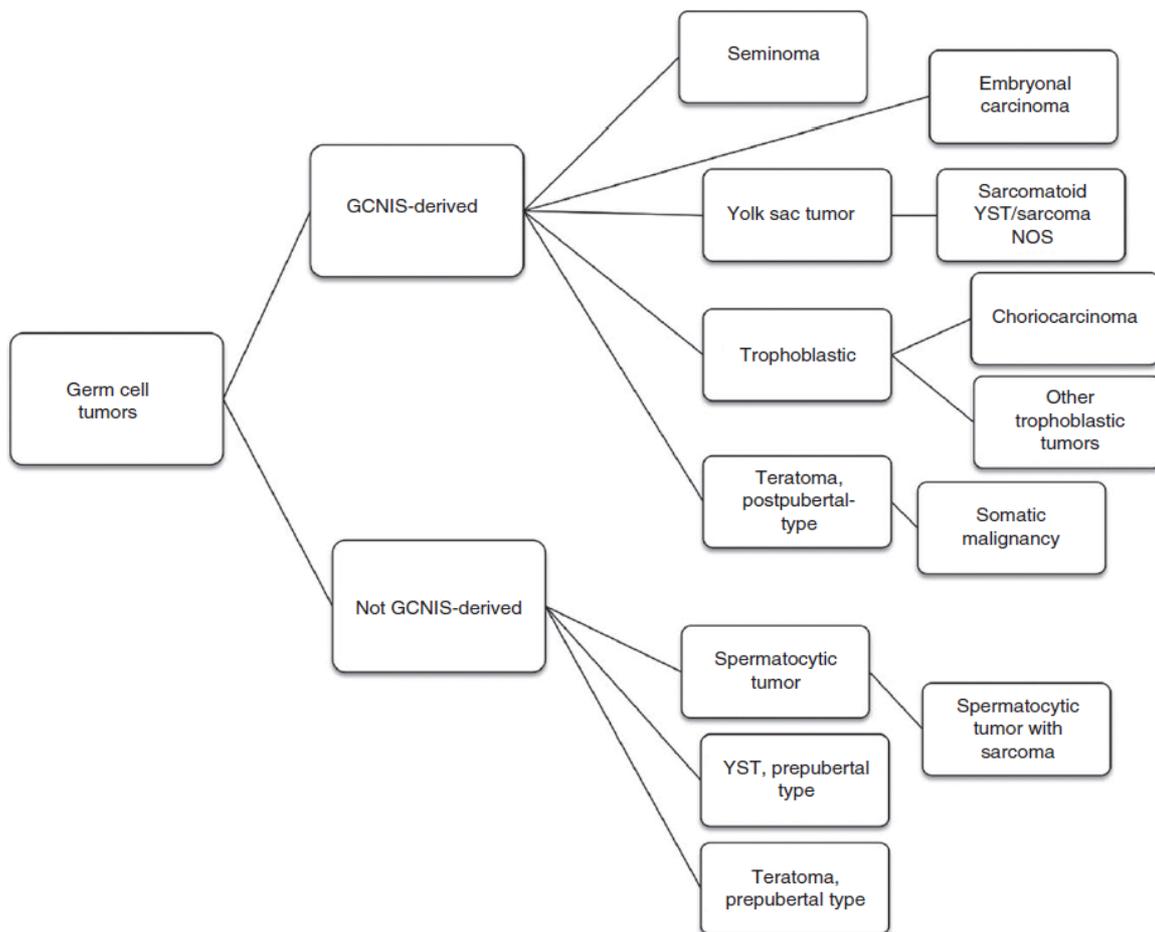
Figure 1.6: **2016 edition of World Health Organisation classification.**

Unfortunately, these classifications don't take into consideration the class of GCTs arising in childhood, which differs from those arising in adulthood. Currently, the paediatric histopathological classification in international clinical use is modified from the original description by Dehner (Table 1.1) (Dehner, 1983). The following part of this section will further describe, in brief, the general definition, occurrence and diagnosis patterns of the tumours classified under Dehner's system.

Type	Histology	Comment
1	Germinoma	a. Intratubular b. Invasive
2	Teratoma	a. Mature (MTD) b. Immature 　　Grade 1-immature tissue < 1 foci per LPF per slide 　　Grade 2-immature tissue 1-3 foci per LPF slide 　　Grade 3-immaure tissue ≥ 4 foci per LPF slide
3	Embryonal carcinoma	Malignant teratoma undifferentiated (MTU) or teratocarcinoma
4	Yolk sac tumour	Also termed endodermal sinus tumour
5	Choriocarcinoma	Malignant teratoma trophoblastic (MTT)
6	Gonadoblastoma	Pre-malignant lesion
7	Mixed malignant GCT	Malignant teratoma intermediate (MTI)

Table 1.1: **Histopathological classification of paediatric germ cell tumours**. In immature teratoma, microfoci of yolk sac tumour, termed Heifetz lesions, are permitted without the tumour being classified as malignant (Heifetz *et al*., 1998). Malignant teratoma refers to a teratoma containing non-germ cell malignant component, e.g. sarcoma. LPF = low power field (x4). MTD = malignant teratoma

differentiated. MTD, MTU, MTT and MTI are terms used in adult histopathological classifications. Adapted and modified from Dehner (Dehner, 1983).

1.4.1 Germinoma/Seminoma

Germinoma or seminoma (in females or males, respectively) are undifferentiated tumours that maintain features of pluripotency. Seminoma is the most common testicular germ cell tumour, representing 50% of the cases and germinoma the second most common ovarian germ cell tumour (Hochstetter and Hedinger, 1982; Herrington, 1999). Immunohistochemically, they can be distinguished from other subtypes of GCTs by negatively staining for cytokeratin (AE1/AE3) and alpha-fetoprotein (AFP) and positively for POU5F1 (OCT3/4) (Looijenga *et al*., 2003). In addition, these tumours are positive for SOX17 and negative for SOX2 staining, which assists in distinguishing them from EC cases (Nonaka, 2009).

1.4.2 Teratoma

Teratoma is a GCT that predominantly occurs in the gonads: the testis and the ovaries. They account for 95% of ovarian GCT while only 4% represents testicular germ cell tumour (Katsube, Berg and Silverberg, 1982). In paediatric patients, teratomas represent 50% of GCTs and they are considered benign (Murray and Nicholson, 2010). They may arise in the ovary (37% of teratomas), sacrococcyx (SCT; 28%), testis (15%), thorax (7%) and other sites (13%). Teratomas are usually composed of multiple cells types derived from at least two germ layers (endoderm, ectoderm and/or mesoderm). The cell type in this type of GCT may be well-differentiated or undifferentiated (Wetherell *et al*., 2014). They are graded as immature (IT), when the teratoma elements are incompletely differentiated and more similar to fetal or embryonic tissue, or mature (MT), when the teratoma elements are well differentiated relative to the germ cell layers (Murray and Nicholson, 2010; McDougal, 2012).

1.4.3 Embryonal carcinoma (EC)

EC display early embryonal differentiation and constitute a rare GCT subtype in childhood. In adolescence and adulthood, EC may present as a component of mixed malignant GCTs. Immunohistochemically, they express POU5F1 (OCT3/4) like germinoma (Poulos *et al.*, 2006). In addition, they express CD30 (Pallesen and Hamilton-Dutoit, 1988) and SOX2, but not SOX17 (de Jong *et al.*, 2008; Nonaka, 2009), which allows them to be distinguished from germinoma.

1.4.4 Yolk sac tumour (YST)

YST is the commonest malignant histological subtype of GCT in childhood, secreting AFP and displaying extra-embryonic yolk sac differentiation (Murray and Nicholson, 2010). However, many different microscopic types are recognised that may cause difficulty in establishing an accurate histopathological diagnosis. Glandular, or endometrioid-like, tumours may mimic certain carcinomas- hepatoid-like tumours may mimic hepatoblastoma (HB) or hepatocellular carcinoma (HCC) and solid patterns of YST may be misinterpreted as germinoma (Ulbright, 2005). Immunohistochemically, yolk sac tumour can be distinguished from germinoma by staining positively for cytokeratin (AE1/AE3) and AFP, and staining negative for POU5F1 (OCT3/4) (Looijenga *et al.*, 2003; Kao *et al.*, 2012).

1.4.5 Choriocarcinoma (CHC)

CHC is extremely rare in childhood and displays extra-embryonic trophoblastic (i.e. placental-like) differentiation. Exceptionally, as a consequence of gestational trophoblastic disease due to abnormally sited placental tissue, they may arise as pure tumours in infants. CHC are highly invasive tumours, often associated with either extensive spontaneous haemorrhage, or haemorrhage on initiation of chemotherapy, termed 'choriocarcinoma syndrome' and often necessitating provision of intensive care to affected patients (Kandori *et al.*, 2010). They are

distinguished immunohistochemically by strong HCG staining (Looijenga, 2009) and are typically characterised by very high levels of serum HCG (Murray and Nicholson, 2010).

1.5 Current clinical management of germ cell tumour

1.5.1 Initial management

The management for mature (MT) and immature teratomas (IT) is surgical. Upon recurrence, a second surgery might be required (Mann *et al*., 2008). Chemotherapy is almost never adopted for teratoma treatment, unless mGCT arises, usually a YST. In the latter case chemotherapy is adopted (Mann *et al*., 2008).

Ovarian GCTs can present either as mature or immature forms of teratoma (80% of all ovarian GCTs) or as a malignant form, with increasing incidence during adolescence (20% of all ovarian GCTs). The presentation of the disease is similar in the two forms, with abdominal pain and lower abdominal mass (Cecchetto, 2014). The surgical approach consists of ovarian tumour resection with sampling of any peritoneal modules (Cecchetto, 2014; Zhao *et al*., 2016). If nodules are composed purely of mature glial tissue ('gliomatosis peritonei', a recognised phenomenon associated with ovarian immature teratoma), no active management is required (Murray and Nicholson, 2010).

Testicular GCTs usually present as a painless scrotal mass. Surgical approach consists of orchiectomy via inguinal excision (Horwich, Nicol and Huddart, 2013). Recently, partial orchiectomy is considered as an initial noninvasive step to preserve the fertility and hormonal function of the patient (Saltzman and Cost, 2018). Although, if the pathology from the partial orchiectomy returns malignant, a radical orchiectomy is completed (Caldwell, Wilcox and Cost, 2017). Pathology is either of a mature teratoma or a malignant GCT - seminoma or non-seminoma (often with 'mixed' malignant components present).

1.5.2 **Chemotherapy**

Chemotherapy is reserved for the stages 2-4 of cancer, with stage 1 patients entering follow-up with an expectant 'watch-and-wait' policy, as only ~25% will have a recurrence. The use of chemotherapy as treatment for GCTs commenced from 1974, led by Einhorn (Einhorn and Williams, 1979). Following the treatment with cisplatin, vinblastine and bleomycin (named PVB regime), the rate of the cure of these diseases improved from 5 to 60% (Einhorn and Williams, 1979). Unfortunately, in paediatric patients, the use of PVB was shown not to be tolerated (Exelby, 1980). Subsequently, this regime was modified to accommodate both adult patients who did not respond to the PVB treatment and children. Initially, etoposide was used in pre-clinical models instead of vinblastine in a new combinatorial treatment named 'BEP' or 'PEB' regime (Schabel *et al*., 1979). This work demonstrated a remarkable synergism between etoposide and cisplatin, showing for the first time that a solid tumour could be cured with second-line chemotherapy (Schabel *et al*., 1979). BEP regime (cisplatin, etoposide, bleomycin) is still currently used by many US paediatric groups, with a reduced dose and frequency of bleomycin compared with adult BEP schedules.

Unfortunately, the BEP regimen was found to be accompanied by underlying toxic side-effects, which remains a serious problem amongst treated patients. Rabinowits et al. showed that the combination of bleomycin and cisplatin gave rise to pulmonary toxicity and renal damage (Rabinowits *et al*., 1990; Kollmannsberger *et al*., 1999). Studies to diminish the toxicity were done by treating GCT patients with four cycles of carboplatin (CBDCA) + etoposide + bleomycin (EBC). Lower levels of renal and gastrointestinal toxicity were observed but myelosuppression was observed (Motzer *et al*., 1990). Consequently, adolescents with EC, CHC or mixed malignant GCTs should be treated with cisplatin-based chemotherapy, as insufficient numbers of patients were treated with carboplatin in the UK CCLG GCII study to be confident that outcomes were equivalent to adult schedules (Mann *et al*., 2000). BEP is, to date, the first line of chemotherapy most used in clinical practice (Honecker *et al*., 2018).

Although, studies have shown that the overall survival after chemotherapy treatment in GCTs treated patients is highly successful (up to 80% survival rate), at least half of the survivors develop toxicity such as ototoxicity (caused by cisplatin and carboplatin) (Haugnes *et al.*, 2010, 2012), nephrotoxicity (caused by cisplatin) or pulmonary fibrosis (caused by inflammation of the lung via the drug) later in their lives (Murray, Nicholson and Coleman, 2015b). The remaining 20% of patients have cisplatin-resistant tumours. These patients have very poor outcomes and significantly increased risk of solid cancers of the pleura and esophagus, and half of them usually succumb to it (Figure 1.7) (Travis *et al.*, 2005; Lorch *et al.*, 2012). Needless to say, it is crucial to focus on the therapies and regimens that could minimise these toxic effects whilst further improving the cure rates for patients with poor-outcomes. There is an urgent requirement to investigate molecular changes that may represent biomarkers and targets for novel therapy, in order to achieve this goal (Murray, Nicholson and Coleman, 2015c).
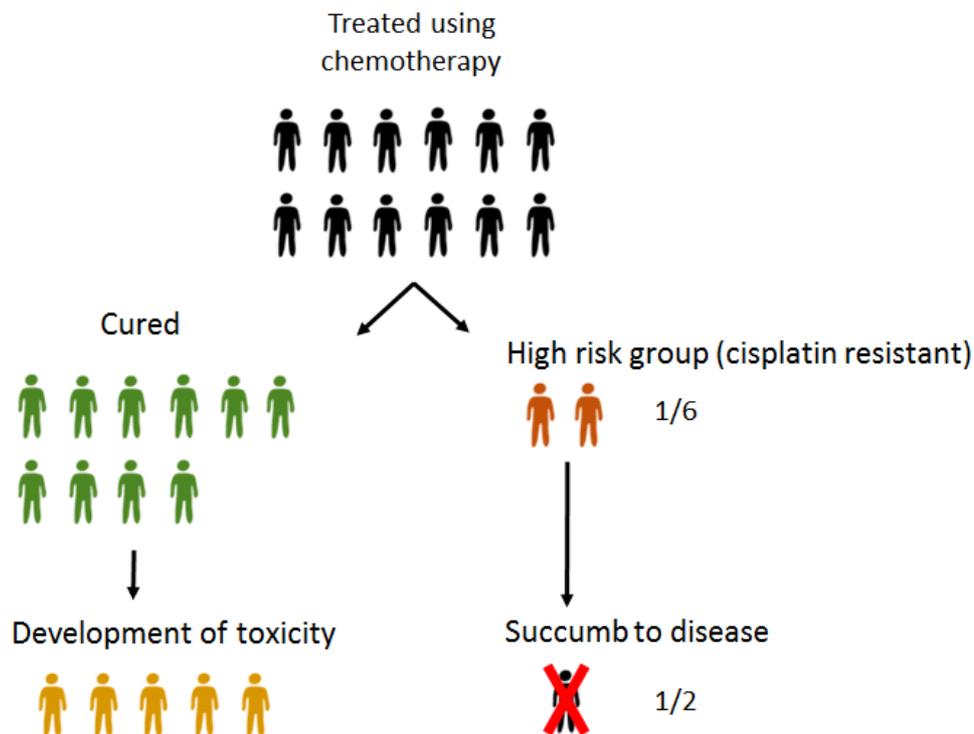
Figure 1.7: **Current treatments and outcome of GCT patients.**

1.6   Molecular biology of germ cell tumours

The earlier section of this chapter introduced an overview of GCTs and the following section will cover the current level of molecular understanding.

1.6.1 Genomic abnormalities

To understand the causes of GCT pathogenesis, genomic studies have previously been undertaken to elucidate specific chromosomal regions of gain or loss. The majority of the studies were performed on adult malignant testicular GCTs (TGCTs) and identified the presence of isochromosome 12p (i12p) (Figure 1.8) (Atkin and Baker, 1982). i12p material is found in 80% of adults with TGCTs and more limited i12p is seen invariably in the other 20% of the cases (Looijenga *et al*., 1993; Collinson *et al*., 2014). Interestingly, in few earlier studies, this gain was not found in tumours in patients <9 years of age (Bussey *et al*., 2001). However, more recent studies done on a larger scale of paediatric samples revealed 12p gain in 44% of cases, with an increasing incidence with age (Palmer *et al*., 2007).
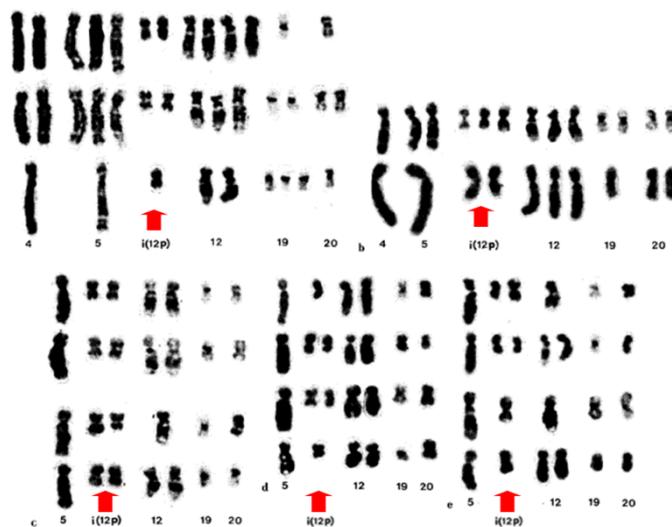
Figure 1.8: **Presence of isochromosome 12p [i(12p)] in adult malignant germ cell tumours.** Karyograms are shown for 5 patients and the i(12p) formation is indicated by red arrows. Taken from (Atkin and Baker, 1982).

These findings suggest that genomic copy number imbalances can distinguish GCT subgroups primarily by patient age (Palmer *et al.*, 2007) (Murray, Nicholson and Coleman, 2015a). This discovery led to further studies to identify the main gene(s) involved in GCT progression (Rodriguez *et al.*, 2003). On chromosome 12, stem cells genes such as *NANOG* are transcribed. *NANOG* overexpression leads to uncontrolled cell proliferation and consequently tumour formation. Interestingly, the TGCT subtypes, EC and seminoma showed *NANOG* overexpression, but other malignant subtypes such as YST, did not (Korkola *et al.*, 2006). Other genomic copy number imbalances have been described in GCTs in adult patients. Most commonly, gain of chromosomes 1, 7, 8, 12, 21 and X and a loss of chromosomes 4, 5, 11, 13 are observed (Terashima, Yu, Wing-Yuk T Chow, *et al.*, 2014). Recent genome wide association studies (GWAS) have identified 18 main susceptibility gene loci in the development of TGCTs in adults, including: activating transcription factor 7 interacting protein (*ATF7IP*), BCL-2 antagonist/killer 1 (*BAK1*), doublesex and mab-3 related transcription factor 1 (*DMRT1*), KIT ligand (*KITLG*), sprouty homolog 4 (*SPRY4*), and telomerase reverse transcriptase (*TERT*) (Koster *et al.*, 2014).

1.6.2 Epigenetic abnormalities

It is also important to take into account the role that epigenetic mechanisms have in the initiation and protection of pluripotency, and therefore their potential role in GCT pathogenesis. From the early stage of germline development, an intensive epigenetic remodelling occurs (Kimmins and Sassone-Corsi, 2005). Deregulation of these tightly regulated processes is known to be involved in the formation and progression of many cancers (Portela and Esteller, 2010). During normal physiological development, PGC formation and maturation is dependent upon epigenetic processes such as histone modification and DNA methylation, to alter DNA accessibility and transcription, respectively. The 'epigenome' is highly dynamic and there is growing evidence that GCTs profoundly differ in their gene expression profiles and epigenetic regulation from normal germ cells (Dina G Kristensen *et al.*, 2013).

1.6.2.1 **Histone modifications**

Looijenga *et al*. studied genome-wide profiles for histone marks (H3K4me3 and H3K27ac) in malignant GCT cell lines [TCam2 (seminoma) and NCCIT (EC)]. H3K4me3 and H3K27ac, are associated with promoter activation and enhancer activation (Rada-Iglesias *et al*., 2011; Clouaire *et al*., 2012). Depending on the cell line, *SOX17* and *SOX2* were differentially enriched for both H3K4me3 and H3K27ac in TCam2 and NCCIT, respectively (van der Zwan *et al*., 2014). Another study has found a substantial difference in the levels of histone modification between mGCT cell lines, with high levels of H3K9me2, H3K27me3, H4/H2A R3me2 and low levels of H3K4me1in seminoma compared with EC (Dina G. Kristensen *et al*., 2013). Differences in the epigenetic landscape, between seminoma and non-seminomatous GCTs, suggests that epigenetic phenomena may alter protein-coding and non-protein-coding genes in these tumours.

1.6.2.2 **Methylation modifications**

Methylation mechanisms are highly dynamic and crucial during germ cell development. Recent studies have analysed human DNA methylation from the PGC migration stage to the gonadal stage. They found that the methylation was decreased to 7.8% in the 11-week male PGCs and 6% in the 10-week female PGCs (Figure 1.9) (Guo *et al*., 2015).
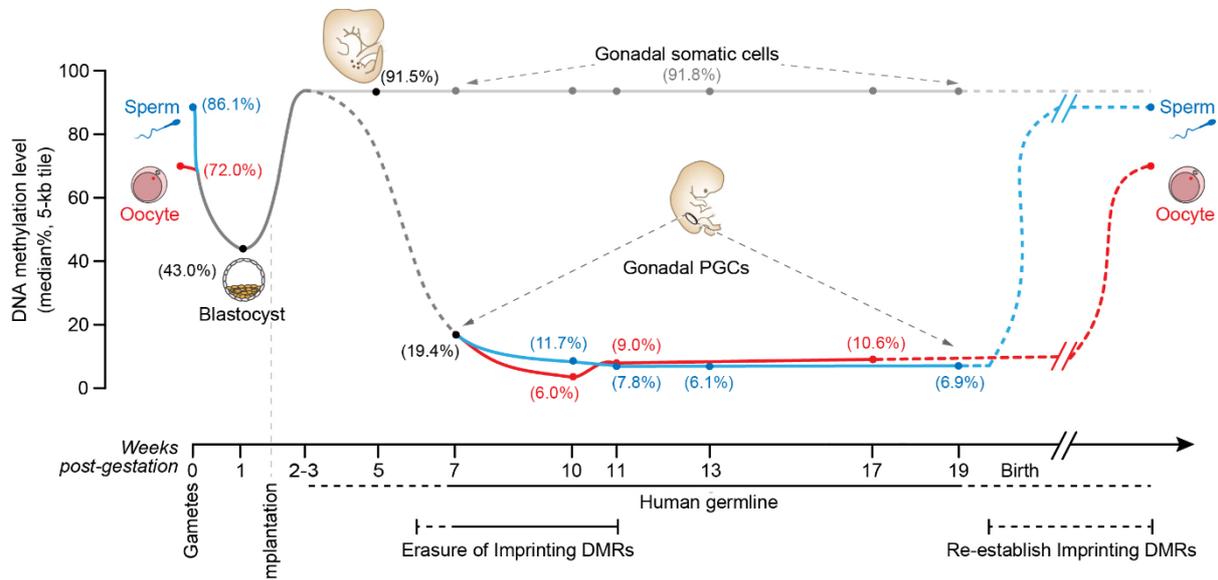
Figure 1.9: **DNA methylation dynamics.** Methylation of human sperm, oocyte, blastocyst, post/implantation embryos, primordial germ cell and gonadal somatic cells. Taken from Fuchou Tang's paper (Guo *et al.*, 2015).

Aberrant DNA methylation is believed to contribute to malignancy in GCTs due to potential errors resulting from irregular methylation during their development (Smiraglia *et al.*, 2002). Methylation mechanisms might also have a role in explaining the heterogeneity of mGCTs. The first global methylation array analysis of gene regulatory regions in mGCTs revealed that YSTs have increased methylation at many loci, including silencing of potential tumour-suppressor genes (TSGs) and those genes associated with apoptosis, such as APC (Jeyapalan *et al.*, 2011) and SFRP2 (Schonberger *et al.*, 2010), are methylated in paediatric YSTs but not in germinomas, resulting in activation of the Wnt signalling pathway.

A second study performed an Illumina GoldenGate Cancer Methylation Panel platform to compare DNA methylation in three histological subtypes of paediatric GCTs (germinoma, teratoma and YST). Methylation patterns showed hypermethylation in a large number of cancer-related genes in YST such as *APC*, *RUNX3* and *HIC1* (Amatruda *et al.*, 2013a). Further analyses were done using Infinium HumanMethylome450 arrays in cell lines of the major

subtypes of GCTs. They studied 7,244 genes and discovered that 147 of them showed correlating differential methylation and expression. This study further confirmed that non-seminoma cells exhibit very high levels of gene methylation as compared with seminoma cells. The highest correlation between methylated areas and gene silencing was found in CpG islands along the promoter region of the silenced gene (Noor *et al*., 2016). More recent studies have found non-canonical CpG sites (CpA, CpT and CpC, collectively named CpH site) to exhibit an extensive methylation in embryonal carcinoma. This pattern was associated with high levels of DNMT3A/3B, two well-known *de novo* DNA methyltransferases (Shen *et al*., 2018). All these studies suggest a model where seminomas, which are the most similar to PGCs based on histological appearance, gene expression and methylation profile, are a precursor of embryonal carcinoma, which is in turn the precursor of yolk sac tumours (Honecker *et al*., 2006).

1.6.3 **Protein coding transcriptome**

The protein-coding transcriptome has predominantly been studied in GCTs from adult patients and a distinct segregation has been shown between histological mGCT subtypes. For seminoma, this included *LZTS1*; for EC, *DNMT3B, GAL* and *GPC4*; for CHC *CGA*; and for YSTs the genes *AFP, APOA2, BMP2, VTN* and *OTX2* (Alagaratnam *et al*., 2011). Similarly, global mRNA gene expression profiles in paediatric malignant GCTs completely segregate the two main histological subtypes, YSTs and germinomas (Palmer et al 20008). Higher enrichment was seen for genes associated with pluripotency such as *NANOG, POU5F1, TFAP2C* and *UTF* for paediatric germinomas while higher enrichment in genes involved in differentiation (*KRT8, KRT19*), lipid metabolism (*APOA1, APOA2*) and proliferation was observed for paediatric YST (Palmer *et al*., 2008). No difference was detected in tumours with the same histology but arising in different anatomical sites (Palmer *et al*., 2008). Furthermore, when adult and paediatric samples were compared, global mRNA profiles robustly segregated these two groups, leading to the conclusions that transcriptome profiles segregate both histological subtype and then age of tumour onset.

1.7   MicroRNAs (miRNAs)

1.7.1 MiRNA overview

1.7.1.1 History and importance of miRNAs

MicroRNAs (miRNAs) are a class of short, non-protein coding RNAs, typically 21-23 nucleotide (nt) long, that post-transcriptionally regulate the expression of protein-coding genes (Farazi *et al*., 2011; Takahashi, Miyazaki and Ochiya, 2014). The importance of miRNAs is underlined by an evolutionary conservation across different species, suggesting a critical role during development and growth (Warnefors *et al*., 2014). After binding to the 3' Untranslated region (3'UTR) region, the miRNA reduces protein levels of its target mRNA via either direct degradation (by processing via P-bodies in the cytoplasm), or through translational repression (Bartel, 2004a). Novel microRNAs are continuously being identified through next generation RNA sequencing and other techniques and therefore miRBase, the online microRNA registry (www.miRBase.org/) (Griffiths-Jones *et al*., 2008), is regularly updated. There are now 48,885 confirmed human mature microRNAs in the most recent miRBase release (v22; March 2018).

They were first discovered by Lee, Feinbaum and Ambros, in 1993, by isolating *Lin-4* from the nematode *Caenorhabditis elegans*. These studies showed that a specific miRNA, *lin-4*, was essential for the normal temporal regulation of development. They discovered that *lin-4* was acting by negatively regulating the level of LIN-14 protein. In these organisms, the downregulation of LIN-14 protein was found to be essential for the progression from the first larval stage (L1) to the second stage (L2) (Lee, Feinbaum and Ambros, 1993). Seven years after the finding of the first miRNAs, *let-7* was discovered. Equally important, *let-7* was needed in *C.elegans* to control the L4-to-adult transition of larval development (Reinhart *et al*., 2000). Soon after the discovery of *let-7*, a new exciting study looked at the sequence conservation of miRNAs among different species. They detected the same miRNA in vertebrate, ascidian, hemichordate, mollusc, annelid and arthropod. In addition to having its sequence conserved,

its function was also conserved between species, suggesting a critical role in the control of the development (Pasquinelli *et al*., 2000).

Since then, there has been a dramatic increase in reports identifying roles for miRNAs in normal development in humans and other species, almost quadrupling between 2005 and 2008 and increasing eighteen-fold by 2013 (Casey *et al*., 2015), highlighting the importance of this new emerging field.

1.7.1.2 **Current knowledge of miRNAs biogenesis and function**

MiRNAs are transcribed by RNA polymerase II (RNAse II) and III (RNAse III) from either 'independent' miRNA genes or from intronic portions of protein-coding genes, to form primary transcripts from DNA, termed pri-microRNAs. Pri-miRNAs, which may be up to a few kilo bases (Kb) in length can also be polycistronic, i.e. code for more than one miRNA (Lee *et al*., 2002). The pri-miRNA transcript fold into hairpins, which act as substrates for an RNAse III family enzyme, DROSHA (Lee *et al*., 2003). It has been recently estimated that a hairpin stem length of 36 nt +/- 3 is optimal for pri-miRNA processing (Roden *et al*., 2017).

The product of DROSHA cleavage is a ~ 70 nt pre-miRNA, which contain a 2-nt overhang at its 3' end (Lee *et al*., 2003).

In addition to the DROSHA-dependent mechanism of miRNA biogenesis, a 'non-canonical' DROSHA-independent pathway has been identified. Some pre-miRNAs are produced from very short introns (mirtrons). Mirtrons are non-canonical miRNAs encoded in introns and their biogenesis starts with a splicing event. They are not processed by DROSHA and enter the canonical pathway at the Exportin-5 level (Figure 1.10) (Westholm and Lai, 2011). Mirtrons are much less evolutionary conserved than canonical miRNAs, and it was hypothesised that their non-canonical pathway could have provided an early avenue for the emergence of miRNA, prior to the advent of *DROSHA* (Ruby, Jan and Bartel, 2007). Recently, awareness of the significance of mirtron miRNAs is increasing, and their role in cancer progression has been established in spinocerebellar ataxia 7 (SCA7), myelodysplastic

syndrome (MDS) and in the promotion of new vessel formation in the angiogenesis event (Aslan *et al.*, 2016; Curtis *et al.*, 2017; Sakai *et al.*, 2017).

Once formed in the nucleus, pre-microRNAs or mirtrons are transported into the cytoplasm by Exportin 5 (XPO5) (Yi *et al.*, 2003; Bohnsack, Czaplinski and Gorlich, 2004), which is also important for the stability of pre-microRNAs as loss of expression results in decreased cytoplasmic microRNA levels without an increase in nuclear pre-microRNAs (Melo *et al.*, 2010) and downregulation of mature miRNA in the cytoplasm (Wu *et al.*, 2018). Once in the cytoplasm, the 2 nt 3' overhang present in the pre-microRNAs is recognised by another RNAse III family enzyme (DICER1) which cleaves the molecule to a~ 20 nt miRNA/miRNA* duplex (the asterisk denotes the passenger strand) (Bernstein *et al.*, 2001). One strand of this duplex is then incorporated into the RNA induced silencing complex (miRISC) (Rivas *et al.*, 2005). Finally, the mature miRNA strand incorporated into the RISC complex results in mRNA degradation or translational repression (Bartel, 2004a; Lim *et al.*, 2005; Krol, Loedige and Filipowicz, 2010).The key components of miRISC complex are argonaute (AGO), which directly binds miRNA, and a glycine-tryptophan protein of 182 kDa (GW182), which act as downstream effector in the repression (Figure 1.10) (Krol, Loedige and Filipowicz, 2010).

The remaining strand (passenger strand) from the original microRNA duplex, which is not incorporated into RISC, is believed to undergo rapid degradation.
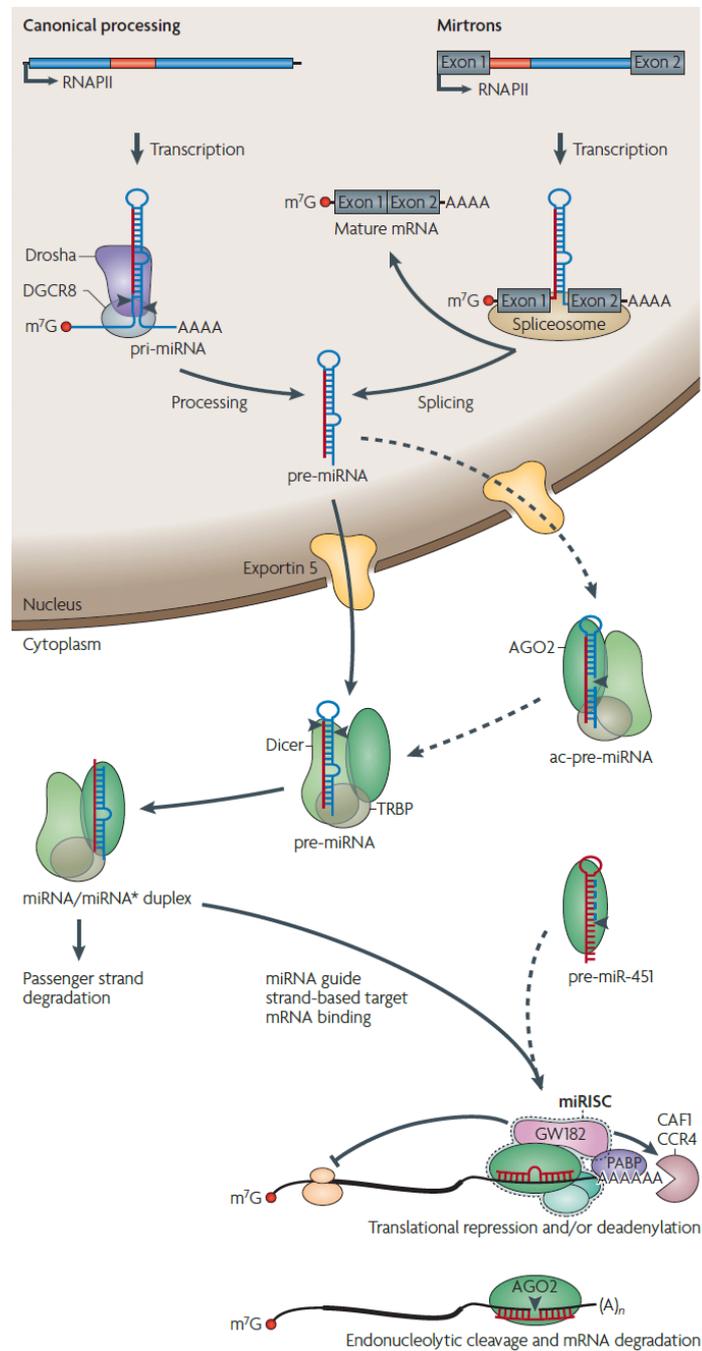
Figure 1.10: **Schematic diagram of miRNA biogenesis.** Both the canonical (DROSHA-dependent) and non-canonical (DROSHA-independent, or 'mirtron') biogenesis pathways are shown. Taken from (Krol, Loedige and Filipowicz, 2010).

1.7.1.3 **Mechanism of miRNA regulation of gene expression**

Upon assembly of the mature miRNA into the miRISC complex (miRNA assembled with Argonaute proteins), 3'UTR target recognition occurs through conventional Watson-Crick miRNA:mRNA base pairing (Ambros, 2000; Esquela-Kerscher and Slack, 2006). The miRNA binding at the 3'UTR of its target mRNA, is predominantly determined by the 'seed' region of the microRNA, located at nt positions 2-8 at the 5' end of the microRNA. The 2-7 nt sequence represents the 'core' component of the seed and this is critical in determining binding specificity to mRNA targets (Lewis, Burge and Bartel, 2005a). Consequently, as each seed region is short and 3'UTRs of mRNA transcripts are much longer, each microRNA is predicted to have hundreds of mRNA targets (Lim *et al*., 2005).

The interaction occurs between the 5' end (seed region) of the miRNA, and the 3' end untranslated region (3'UTR) of the target mRNAs. Thought not only to happen when there is absolute complementarity in the seed region but less frequently also when few unperfected seed matches occur. It has been found that some interaction can occur with a nucleation bulge at position 5-6 (Tay *et al*., 2008; Chi, Hannon and Darnell, 2012). After binding to the 3'UTR region, the miRISC complex reduces protein levels of miRNA target mRNA via either direct degradation (by processing via P-bodies in the cytoplasm), or through translational repression (Bartel, 2004b; Iwakawa and Tomari, 2015). In mammalian cells, the majority of this negative regulation has been shown to occur through mRNA degradation (Baek *et al*., 2008; Selbach *et al*., 2008; Guo *et al*., 2010a; Eichhorn *et al*., 2014). For the degragation to happen, the mRNAs are firstly deadenylated by the deadenylase complex (PAN2-PAN3 and CCR4-NOT) (Wahle and Winkler, 2013). Once deadenylated, they are decapped by decapping protein 2 (DCP2), which in tandem with other cofactors degradate the mRNA (Jonas and Izaurralde, 2013). The deadenylase complex, decapping complex, Ago and GW182 all localise in P-bodies, which are cytoplasmatic ribonucleoprotein granules primarly composed of translationally repressed mRNA and protein related to mRNA decay (Luo, Na and Slavoff, 2018). In addition to mRNA degragation, miRNAs can promote translational repression. The question of how miRNAs repress translation is controversial. With the introduction of ribosome profiling method, the

latter hypotesis could be elucidated. Bartel *et al*. monitored both mRNA levels and the corresponding ribosome density after the introduction of exogenous mimics miRNA (Guo *et al*., 2010b; Kuersten *et al*., 2013). Reduction in mRNA levels due to mimic miRNAs also decreased the levels of ribosome density, giving less credibility to the theory of translational repression.

More recently, new studies have started to re-evaluate the miRNA:mRNA interactions. Unconventional miRNA binding sites in the 5'UTRs or coding genes have been successfully reported (Lytle, Yario and Steitz, 2007; Cloonan, 2015). No differece in mRNA repression was detected when the miRNA binding site was located at the 5′ or 3′ to the luciferase open reading frame (ORF). Further functional studies have been done of miR-103a-3p, miR-122 and miR-10a, whose target sites are within 5′UTR regions of their mRNA target (Ørom, Nielsen and Lund, 2008; Roberts, Lewis and Jopling, 2011; Zhou and Rigoutsos, 2014). Coding DNA Sequences (CDS) have also been shown to conserve some target sequences. Similar to the experiment above, miRNA target sequences have been cloned into the CDS of the luciferase, showing a decrease in its activity (Hausser *et al*., 2013).

These new ways of miRNA:mRNA interaction opens new avenues in the study of miRNAs binding and function. For example, currently most of the target predictions databases are tailored in order to predict binding sites in the 3'UTR of a gene, which means that we are missing out on other sets of mRNAs that could be equally important in the regulation of our cells. Further discoveries and validations on this argument need to be undertaken.

1.7.2 MiRNAs role in tumorigenesis

To experimentally determine the role of miRNAs in tumorigenesis, previous studies have looked at the effects of disruption of the miRNA machinery in normal cells. This was done by testing short hairpin RNAs (sh-RNAs) against DROSHA, DGC8 and DICER1. As expected, an increase in cell proliferation was observed, with a particular increase in the cell percentage in S phase and a reduce percentage of cells in G0/G1 phase (Kumar *et al*., 2007). This *in vitro* study was followed by an *in vivo* study which confirmed the importance of miRNAs in cancer

progression; when Dicer1 was inactivated in murine models the same phenotypic effect was observed (Kumar *et al*., 2007).

Since then, many studies have aimed to find a role for miRNAs in the control of oncogenes or tumour suppressor genes. Mutation or deregulation of these genes are often combined with tumour progression (Huebner and Todaro, 1969).

Together with an enhancement in cell growth, levels of C-MYC, K-RAS and P53 are usually observed (Johnson *et al*., 2005; Hoffman and Liebermann, 2008; Vazquez *et al*., 2008; Campaner and Amati, 2012; Li *et al*., 2012; Drosten *et al*., 2014).

It is important to mention that miRNAs function is cancer-specific and that the same miRNA can act differently depending of the context (Banzhaf-Strathmann and Edbauer, 2014). Thus, the same miRNA can act as oncogene or tumour suppressor in their own right. For example, miR-125b-5p is broadly studied in a variety of tumours such as breast cancer and glioma (Akhavantabasi *et al*., 2012; Li *et al*., 2018; Yuan *et al*., 2018), where it acts as tumour suppressor, or colorectal cancer or renal cell carcinoma (Mori *et al*., 2011; Jin *et al*., 2017; Yu *et al*., 2017), where it acts as oncogene.

1.7.3 **MiRNAs in the regulation of cell cycle**

The cell cycle is under the strict control of several mechanisms that enforce correct cell division. The basic cell cycle in divided into four phases namely, G1, S, G2 and M (Norbury and Nurse, 1992). During S and M phases, the cell initially generates a faithful copy of its genetic material (S phase) and then it divides it together with all the other cellular components into two identical daughter cells (M phase). G1 and G2 phases are a 'gap period' in which the cell prepares itself for a successful completion of S and M phases. Adult cells are almost all in a quiescent state known as G0 phase, meaning that after proliferation, cells exit the cell cycle. Quiescent cells may re-enter the cell cycle or remain in their non-proliferative state. Re-entering the cell cycle requires a regulated process of activation and inactivation of specific proteins at specific times (Malumbres and Barbacid, 2001).

The transition of the cell from one phase to another is orchestrated by a series of different regulators. A major group of regulators are the Cyclin-Dependent Kinases (CDKs). Activity of CDKs, as the name suggests, is dependent on the activity of cyclins – a family of regulatory subunits. Cyclins are synthesized and degraded at different stages of the cell cycle providing temporal regulation of CDKs (Vermeulen, Van Bockstaele and Berneman, 2003). One of the first links between miRNAs and the cell cycle was the discovery of the anti-proliferative potential of the miR-15a-16-1 cluster, due to targeting of CDKs. This cluster induced arrest in G1 phase by targeting CDK1, CDK2 and CDK6 (Linsley *et al*., 2007; Wang *et al*., 2009; Takeshita *et al*., 2010). Other regulatory miRNAs were soon discovered - miR-124, miR-34a, miR-125b, miR-195 and found to play a role in cell cycle regulation by targeting CDK4 and CDK6 (Pierson *et al*., 2008; Sun *et al*., 2008; Xu *et al*., 2009; Shi *et al*., 2010). CDKs, in turn, phosphorylate target proteins and drive the cell sequentially through the different phases of growth and division. One example of this regulation is the entry into G1 phase. Entry into G1 is triggered by the D-type cyclins (CycD1, CycD2, CycD3) that bind CDK4 and CDK6 eliciting their activity during this phase. Consequently, negative regulators of the cell cycle such as the retinoblastoma proteins RB, RBL1 and EBL2 are down-regulated. Its inactivation is essential for starting a new cell cycle and to release E2F-dependent transcription of cell cycle genes. RB has been shown to be targeted from miR-106a in gastric, prostate and lung tumour cancer (Volinia *et al*., 2006). Levels of E2F were firstly found to be regulated by miR-17-5p and miR-20a cluster, together with C-MYC expression. C-MYC regulates the expression of miR-17-5p and miR-20a cluster, which, in turn, regulate the expression of E2F1, supporting a model in which miR-17-5p and miR-20a limit C-MYC-mediated induction of E2F1 expression, preventing uncontrolled reciprocal activation of these gene products (O'Donnell *et al*., 2005). Since then, many more miRNAs have been found to control the cell cycle progression through direct or indirect binding of the main effectors of cell cycle progression (Figure 1.11) (Bueno and Malumbres, 2011).
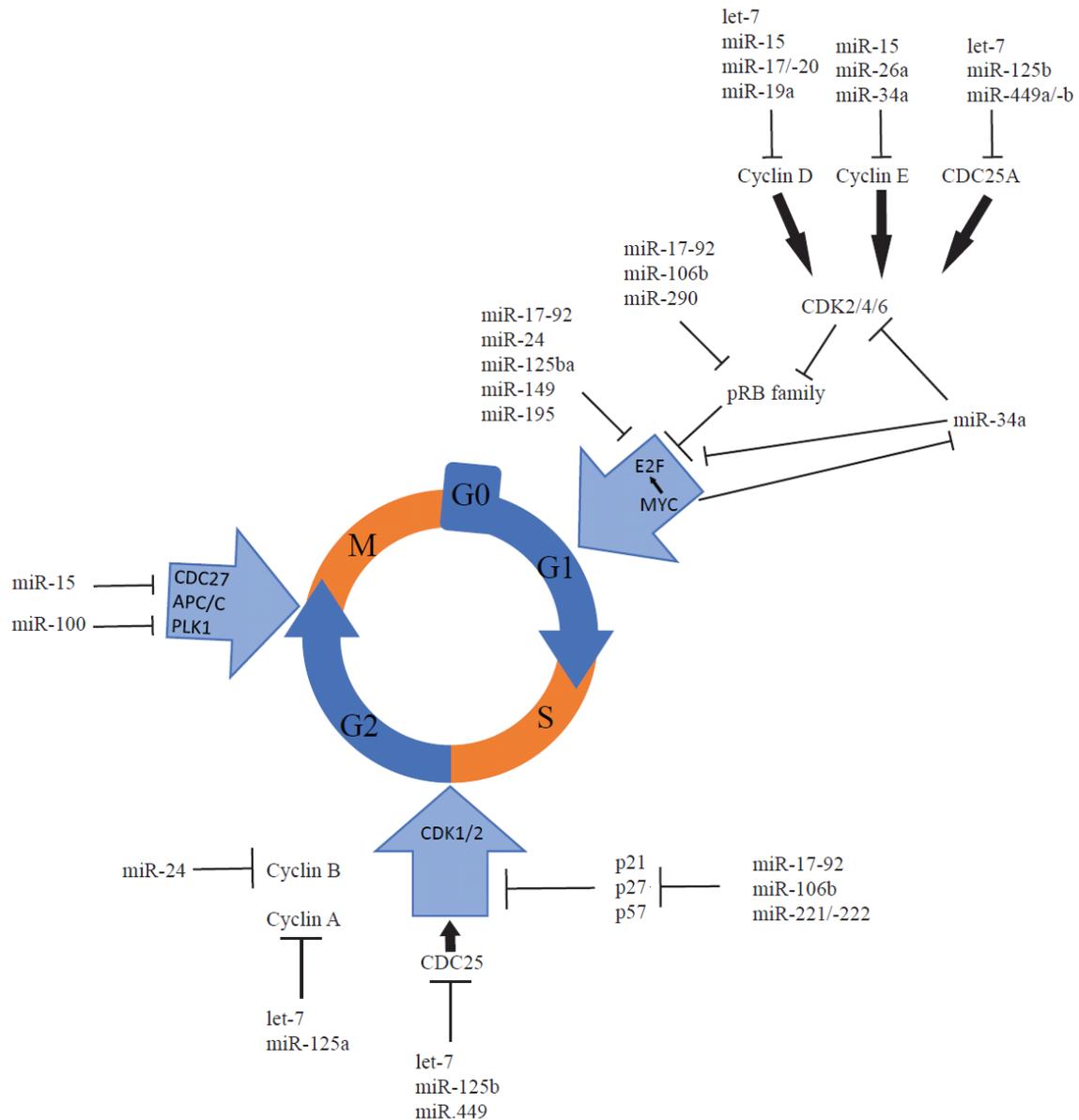
Figure 1.11: **Overview of miRNAs role in the cell cycle progression.** MiRNAs with proliferative potential are shown in red whereas anti-proliferative miRNAs are in blue. S, S-phase; M, Mitosis; G1 and G2 indicate transition phases of the cell cycle whereas G0 indicate quiescent cells. Adapted from (Bueno and Malumbres, 2011).

1.7.4 MiRNAs in the regulation of apoptosis, autophagy and necroptosis

Multicellular organisms have tightly controlled cell division and also cell death (Tyson and Novak, 2014). Billions of cells in a healthy human body die from programmed cell death every hour in organs such as the bone marrow and the intestine. This programmed cell death is named apoptosis, necroptosis (programmed necrosis) and autophagy. Programmed cell death is strictly regulated by a multitude of factors of which miRNA seems to have a role (Su *et al.*, 2015). Apoptosis is the most common form of programmed cell death (Danial and Korsmeyer, 2004). There are two main apoptotic pathways: the extrinsic or death receptor pathway and the intrinsic or mitochondrial pathway.

The intrinsic pathway relies on the formation of a complex termed the apoptosome, composed of procaspase-9, apoptotic protease activating factor 1 (APAF-1), and cytochrome c. A series of Bcl-2 family members control the release of cytochrome c by regulating mitochondrial membrane permeabilisation. When cytochrome c leaks out of the membrane, a series of pro-apoptotic protein are activated, which terminates with the activation of caspase-3, 6 and 7. Many studies have found a connection between deregulated miRNAs expression and uncontrolled cell death by the activation of the intrinsic apoptosis pathway such as miR-125b, which was found to target Bcl-2 antagonist killer 1 (BAK1) that is essential to induce Taxol-induced apoptosis. Furthermore, its downregulation confers resistance of breast cancer cells to paclitaxel promoting cancer survival (Zhou *et al.*, 2010). BCL-2 has been found to be the target of other miRNAs such as miR-204, miR-148a and miR-24-2c (H. Zhang *et al.*, 2011; Srivastava *et al.*, 2011; Sacconi *et al.*, 2012).

The extrinsic pathway is initiated by the binding of death ligand to the death receptors in the TNF receptors superfamily. This interaction is then followed by a cascade of events that terminates with the activation of caspase-3, 6 and 7 and the induction of apoptosis. MiRNAs involved in the extrinsic pathway have been discovered such as miR-20a, miR-146a and miR-196b, which by targeting FASL death ligand promote cancer cell survival (Z. Li *et al.*, 2011; Huang *et al.*, 2012; Guo *et al.*, 2013).

In the last few years, autophagy studies have increased our understanding, allowing researchers to speculate that this regulatory mechanism is one of the most important pathways in cancer progression. In 2016, Yomiuri Shimbun received a Nobel Prize for his contributions in understanding the hidden cellular mechanisms of autophagy (Takeshige *et al.*, 1992). Autophagy is usually initiated by starvation stimuli (Kroemer, Mariño and Levine, 2010; Anding and Baehrecke, 2015), hypoxic and endoplasmic reticulum (ER) stress (Rashid *et al.*, 2015; Niu *et al.*, 2017) and reactive oxygen species (ROS) (Chen *et al.*, 2017). AMPK-mTORC1 has been characterized as being at the centre of autophagy regulation in which it integrates multiple stimuli to initiate the UNC-51-like kinase (ULK) complex, which is situated at the beginning of the autophagy cascade (Kim and Lee, 2014). MiRNA have been shown to regulate this complex. For example, miR-20a and miR-106b have been shown to regulate the UNC-51-like kinase (ULK) complex (H. Wu *et al.*, 2012), while miR-148b was reported to target AMPKα1 to inhibit cell proliferation and invasion and to enhance cancer cell chemo-sensitivity (H. Wu *et al.*, 2012). Many other miRNAs have been shown to participate in the correct autophagy function, giving them a central role in this regulatory pathway as intensely reviewed from Gozuacik (Gozuacik *et al.*, 2017).

Necroptosis is the third mechanism of programmed cell death. Few studies to date have linked necroptosis and miRNA expression, as for many years necrosis was considered an unregulated form of cell death. Only recently a form of regulated necrosis has been discovered, hence the term 'necroptosis' (Galluzzi and Kroemer, 2008). The formation of the 'necrosome' by receptor-interacting protein kinase 1 (RIP1, also known as RIPK1) and RIP3 is the most critical event in necroptosis (Holler *et al.*, 2000). Recently, two studies were published on the role that miRNAs (miR-223-5p/-3p and miR-21) have in the regulation of necroptosis. MiR-223-5p/-3p has been found to promote necroptosis in myocardial ischemia/reperfusion (I/R) through suppression of RIP1/RIP3/MLKL pathway (Qin *et al.*, 2016), while miR-21 promotes necroptosis in the pancreas through repression of CDK2AP1 (Afonso *et al.*, 2018). Clearly, the quest to find the contributions of miRNAs in cellular death is on-going and requires further attention.

1.7.5 **Deregulation of miRNAs**

Due to the central role of miRNAs in the regulation of normal cellular activities, and their tumorigenic role when deregulated, it is now vital to understand how they are controlled. Their regulatory pathway has to be well understood for further research to be able to identify potential new therapies to control and combat cancer progression. This section will shed some light on the deregulation of miRNAs and its consequences.

1.7.5.1 **Amplification or deletion of miRNA genes**

Abnormal miRNA expression can be due to alteration in the genomic locus, where the miRNA transcript is located. It has been estimated that half of the miRNAs that exist are located in regions of high instability due to partial chromosome deletion, loss of heterozygosity (LOH), minimal regions of amplification, common breakpoint regions or SNPs (Calin *et al*., 2004a). This was confirmed by interrogating 283 known human miRNA genes by high-resolution array-based comparative genomic hybridisation in 227 specimens from human ovarian cancer, breast cancer and melanoma (Zhang *et al*., 2006a).

The first discovery of aberrant miRNA expression due to genomic instability was the loss of miR-15a/-16-1 cluster at chromosome 13q14 observed in B-cell chronic lymphocytic leukaemias (B-CLL) (Calin *et al*., 2002). On the other hand, in lung cancer, the same cluster gene was amplified, resulting in an increase in the production of miRNAs, and enhanced cellular proliferation (Hayashita *et al*., 2005). Gain of genetic material was also observed in eight cases of diffuse large B-cell lymphoma (DLBCL) (Tagawa and Seto, 2005).

Another form of genomic alteration can derive from chromosomal translocation (Sax and Enzmann, 1939). Chromosomal translocations have also been shown to have a role in the overexpression of miRNAs such as miR-125b-5p in acute lymphoblastic leukemia (ALL) in children. A juxtaposition of a immunoglobulin heavy chain (IGH) enhancer on chromosome 11 in front of miR-125b-1 and BLID was found to be the cause of the upregulation of both genes (Tassano *et al*., 2010).

1.7.5.2 Transcriptional control of miRNAs

Altered levels of transcription factors (TFs) that bind to promoter or enhancer regions, upstream of miRNA loci, may account for dysregulated miRNA expression in malignancy. Examples of this deregulation are proteins that are often deregulated in cancer such as C-MYC and P53. The proto-oncogene C-MYC encodes for a TF that regulates cell proliferation, growth and apoptosis. C-MYC overexpression causes aberrant binding to the promoter region of miRNAs, resulting in upregulation of miR-17-92 cluster and repression of miR-26, miR-29, miR-30 and let-7 family members (O'Donnell *et al*., 2005; Chang *et al*., 2008). Feedback-loops have been discovered which may act as possible TF-miRNA regulation mechanisms. An interesting example has been shown in c-Myc-miR-122 regulation. C-MYC was shown to repress miR-122 expression through binding its promoter region. Mir-122 in turn indirectly inhibits the transcription of C-MYC itself by targeting TFDP2 and E2F1. The dysregulation of this feedback loop leads to the onset of hepatocellular cancer (Wang *et al*., 2014). Interestingly, in hepatocellular carcinogenesis, they found another feedback-loop mechanism that also involves C-MYC and miR-148a-5p/miR-363-3p (Han *et al*., 2013).

P53 causes overexpression of miR-34a in CLL through binding its promoter and in turn, miR-34a promotes P53 expression by targeting SIRT1, a negative regulator of p53 (Yamakuchi and Lowenstein, 2009; Hermeking, 2010). In addition to C-MYC and P53, there are other examples of TFs that regulate miRNA expression. For instance, *CDX2* is a homeobox TF that regulates the expression of miR-125b by binding to its promoter region. *CDX2* overexpression induces miR-125b which inhibits core binding factor β (*CBFβ*) translation, which is essential for myeloid cell differentiation. Its repression promotes leukaemogenesis (Lin *et al*., 2011).

1.7.5.3 Dysregulated epigenetic changes

Another well-known feature of cancer development and progression is epigenetic alterations. These alterations can be presented as DNA hypomethylation, aberrant hypermethylation of tumour suppressor genes and disruption of histone modifications (Gronbaek, Kirsten Hother, Christoffer Jones, 2007).

MiRNA expression was found to be regulated by epigenetic changes in cancer, for the first time, in 2006. Saito and co-workers found that in bladder cancer almost 6% of miRNAs were upregulated following treatment with DNA methylation and histone acetylation inhibitors (Saito *et al*., 2006a). Since then, more studies on the importance of methylation have been published. For example, the expression of miR-223 has been shown to be epigenetically silenced by CpG methylation in primary leukaemic blasts (Fazi *et al*., 2007) and miR-125b was found to be highly downregulated in invasive breast cancer because of hypermethylation of its promoter (Y. Zhang *et al*., 2011a). Other miRNAs more recently have been discovered to be modulated by epigenetic regulation in a variety of cancers (Suzuki *et al*., 2012; Ramassone *et al*., 2018). MiR-34a promoter hypermethylation is responsible for its reduced expression in a multitude of cancers, including acute myeloid leukemia (AML), lung cancer and breast cancers (Watanabe *et al*., 2012; Pigazzi *et al*., 2013; Vrba *et al*., 2013). Epigenetic regulation of miRNAs has been combined into a database called EpimiR, which collated evidence of 1,974 such regulations, with 19 discrete types of epigenetic modification involving 617 miRNAs across seven species, all experimentally validated (Dai *et al*., 2014). These evidences have highlighted the important role of epigenetic in miRNA regulation. Aberrant methylation of miRNA could be used as potential prognostic markers or as target for therapeutic intervention in cancer.

1.7.5.4 **Defects in miRNA biogenesis machinery**

As previously described, miRNAs are strictly regulated during their biogenesis and numerous proteins act in tandem to promote their maturation. Therefore, mutations or aberrant expression of protein involved in the miRNAs biogenesis machinery can lead to miRNAs dysregulation (Hata and Kashima, 2016). For example, *DROSHA* and *DICER1*, two of the key component of miRNA maturation, have been found to be dysregulated in many tumours (Lu *et al*., 2005; Murray *et al*., 2014). To demonstrate the functional significance of *DROSHA* and *DICER1*, knockdown using RNA interference (RNAi) an RNAi approach was done in neuroblastoma cell line and an increase of cell growth *in vitro* was observed (Lin *et al*., 2010). On the other hand, increased levels of *DICER1* and *DROSHA*, appear important in the development of

certain cancers such as serious ovarian carcinoma (Vaksman *et al*., 2012). Recently, a single-nucleotide substitution/deletion mutation was observed in *DGCR8* and *DROSHA* in patient affected from Wilms' tumours leading to decrease of miR-200 and let-7a (Walz *et al*., 2015).

Proteins such as AGO and exportin-5 (XPO5) also have been found to be affected in certain tumours. For example, a defect in XPO5 traps pre-miRNA in the nucleus due to a genetic mutation of XPO5, which generates a protein without the C-terminal, necessary for the formation of the pre-miRNA/XPO5/Ran-GTP ternary complex (Melo *et al*., 2010). These mutations lead to abnormal mature miRNA expression levels and consequently to tumorigenesis.

1.7.6 MiRNAs in Germ Cell Tumours (GCTs)

The Coleman group previously investigated whether dysregulation of miRNAs was present across the clinical spectrum of GCTs. They profiled 615 miRNAs in paediatric malignant GCTs, control and cell lines (48 samples in total) (Palmer *et al*., 2010). They found that the most significant differentially expressed miRNAs in mGCTs were derived from miR-371-373 and miR-302/367 clusters, which were over-expressed, all with a *p* value<0.00005. However, they found that most of the differentially-expressed miRNAs were down-regulated in mGCTs compared with controls and benign tumours (Figure 1.12), consistent with previous findings in other epithelial cancers (Lu *et al*., 2005).
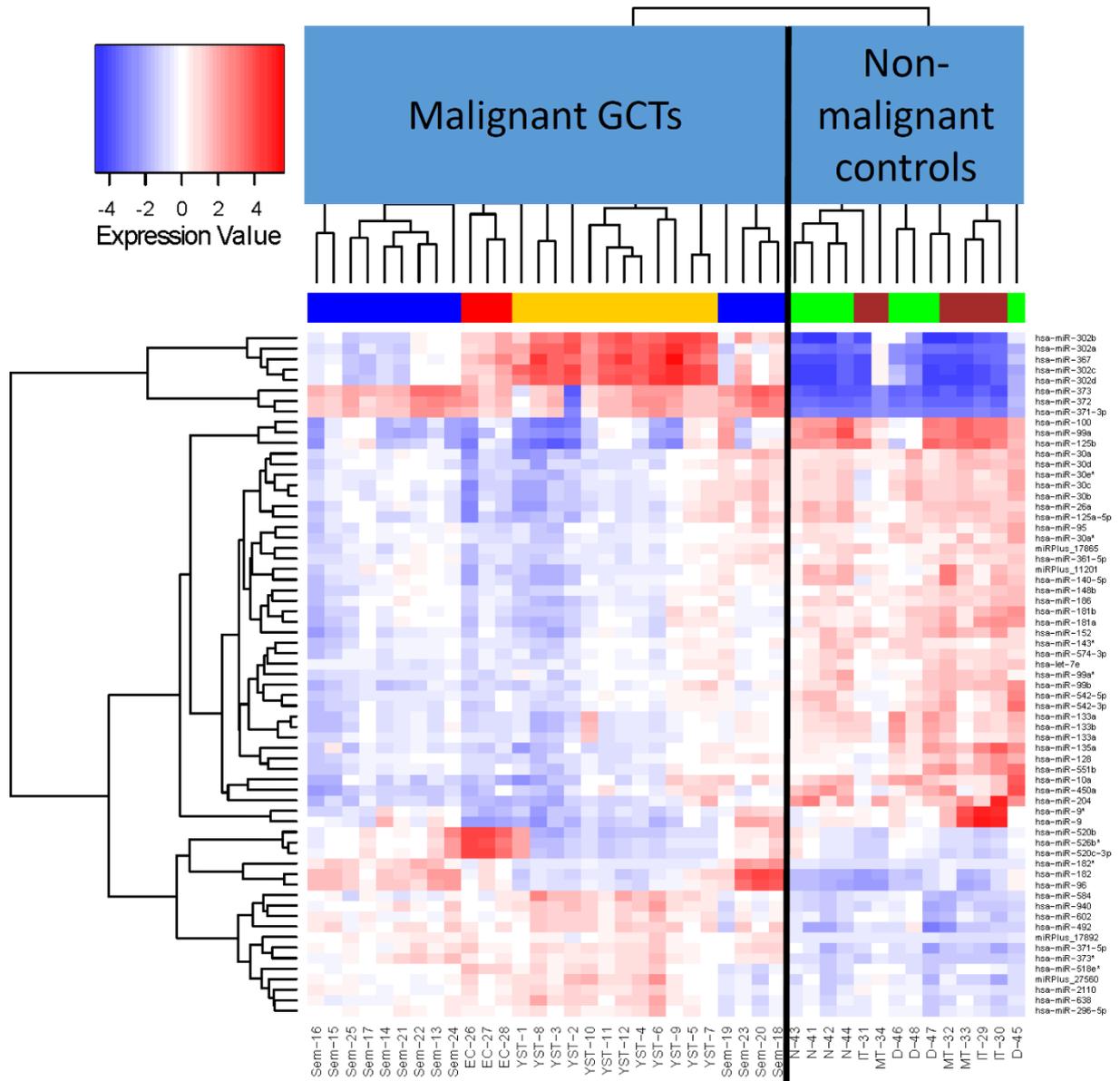
Figure 1.12: **Heatmap of miRNAs expression in 42 clinical tissue samples.** The heatmap is based on the most significantly differentially expressed miRNAs (in red we can see an overexpression and in blue an underexpression of the different miRNAs). The dendogram completely segregates paediatric malignant germ cell tumours (seen in blue for germinoma/seminoma, yellow for yolk sac tumour and red for embryonal carcinoma) from benign tumours (teratoma) in brown and normal controls in green. Supplementary Figure S2 in Palmer et al. CR 2010 (Palmer *et al.*, 2010).

1.7.6.1 **Downregulated miRNAs**

In view of the aforementioned findings, the Coleman group further investigated the role of certain downregulated miRNAs in mGCTS, which represent putative tumour-suppressor genes. Initially, they performed an in-depth study to find the potential significance of downregulation of the entire *let-7* miRNA family in GCTs (Murray *et al*., 2013), which is a known tumour suppressor gene in other cancers that regulates cell proliferation (Takamizawa *et al*., 2004; Johnson *et al*., 2007). From microarray data, they discovered LIN28 (an RNA binding protein and known negative regulator of *let-7* biogenesis) was particularly abundant in mGCTs. LIN28 acts by blocking the downstream processing of precursor *let-7* molecules (Viswanathan, Daley and Gregory, 2008). Once LIN28 binds *pri-let-7* and *pre-let-7*, it prevents their processing by DROSHA and DICER1 (Viswanathan and Daley, 2010a). The binding occurs through a stem loop motif that includes the nucleotide sequence 'GGAG' leading to the recruitment of a terminal-uridyl-transferase (TUTase), resulting in *pre-let-7* uridylation and degradation (Heo *et al*., 2009). In GCTs, *LIN28* knockdown (KD) resulted in specific upregulation of *let-7* family members, confirming a negative regulation of *let-7* by LIN28. Furthermore, replenishment of *let-7* levels in mGCT cell lines using a mimic *let-7* molecule (namely *let-7e*) resulted in reduced cell numbers associated with reductions in *let-7* targets including *MYCN*. In addition, reductions in LIN28 itself were also observed, demonstrating a feedback loop, through a let-7 binding site in the LIN28 3'UTR (Figure 1.13) (Murray *et al*., 2013). Moreover, other miRNAs known to be downregulated in GCTs have an SCR in the 3'UTR of *LIN28* (e.g. miR-125b-5p).
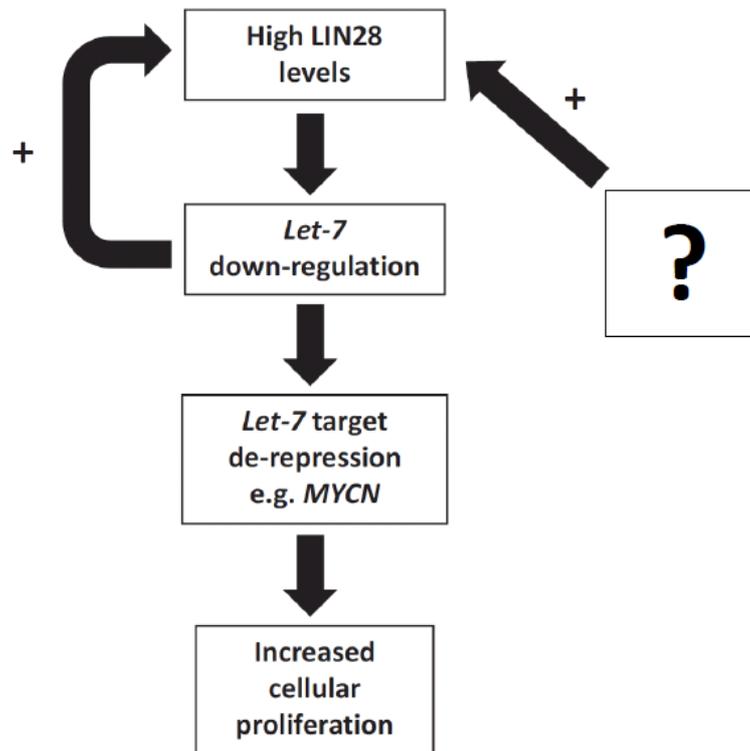
Figure 1.13: **Schematic representation of the LIN28/let-7 pathway.** This model has been described in GCTs. Adapted from (Murray, Nicholson and Coleman, 2015b).

1.7.6.2 **Other important miRNAs downregulated in mGCTs**

Subsequently, the Coleman group commenced the study of other important downregulated miRNAs. The miRNAs with expression levels shown in deep blue in the heat map (Figure 1.12) are the most significantly downregulated miRNAs (adjusted p-values $<1\text{x}10^5$). From these data, five top-ranking putative tumour suppressor miRNAs were chosen for further investigation. These are the five most downregulated miRNAs, ranked by adjusted p-value, with at least a mean 4-fold decrease in paediatric mGCTs compared with non-malignant samples. These miRNAs are: miR-152-3p, -204-5p, -99a-5p, -100-5p and -125b-5p (in descending order of p-value significance) (Table 1.2) (Palmer *et al*., 2010).

Putative tumour suppressor miRNAs	FC	Adjusted p-value	Sequence
miR-152-3p	-4.17	5.09×10^{-9}	UCAGUGCAUGACAGAACUUGG
miR-204-5p	-7.78	2.32×10^{-8}	UUCCCUUUGUCAUCCUAUGCCU
miR-99a-5p	-11.1	9.25×10^{-8}	AACCCGUAGAUCCGAUCUUGUG
miR-100-5p	-10.5	1.23×10^{-7}	AACCCGUAGAUCCGAACUUGUG
miR-125b-5p	-10.9	9.59×10^{-7}	UCCCUGAGACCCUAACUUGUGA

Table 1.2: **The most downregulated miRNAs in paediatric malignant GCTs.** The mean level of this downregulation in the tumour subtypes is expressed as a negative log2 fold-change (log2FC) based on previous global miRNA profiling in paediatric and adolescent GCTs (Figure 1.12). Fold change (FC) is a measure describing how much a quantity changes going from an initial value to a final value.

1.7.7 MiRNAs as potential therapy

1.7.7.1 MiRNA inhibition therapy

MiRNAs are of increasing interest in the clinical field, as they may be used as potential therapeutic strategies. Considerable strides have been made to develop and study anti-miRNA oligonucleotides that bind overexpressed miRNAs in cancer cells. These antagonists of miRNAs have been used both in *in vitro* and *in vivo* models of breast cancer (Obad *et al*., 2011), B-cell lymphoma (Zhang *et al*., 2012) and medulloblastoma (Murphy *et al*., 2013). There is also an extended Phase II clinical trial of an inhibitor against miR-122 (Miravirsen) in patients with chronic hepatitis C, who are not responding to current therapies, and it has been showing very promising early results (Lanford *et al*., 2010; Lindow and Kauppinen, 2012; van der Ree *et al*., 2016). A more recent clinical trial (LO7) showed that using RG-101 (a hepatocyte

targeted carbohydrate conjugated oligonucleotide with potent antagonist activity against miR-122) resulted in significant reduction of viral load in treated patients (Van Der Ree *et al*., 2015).

1.7.7.2 **MiRNA replacement therapy**

The feasibility of miRNA replacement strategies has been investigated using different delivery systems (Höbel and Aigner, 2013; Aigner and Fischer, 2016). Local delivery has been performed using different delivery methods – e.g. (i) a cholesterol-conjugated 2'-O-methyl-modified miR-375 mimics were used in a hepatoma xenograft resulting in suppression of liver cancer cell growth (He *et al*., 2012); (ii) a lentiviral vector was used to deliver let-7 in non-small cell lung cancer (Trang *et al*., 2010); and (iii) polyethylenimine (PEI)-mediated delivery of unmodified miRNAs was adopted to deliver miR-145 and miR-33a into a mouse model of colon carcinoma (Ibrahim *et al*., 2011). Recently, nanoparticle techniques have been used, such as a gold-nanoparticle vehicle to prevent metastasis in breast cancer. MiR-96/miR-182 was delivered locally in a breast cancer mouse model and the level of metastasis derived from the primary tumour was significantly decreased (Gilam *et al*., 2016). Systemic delivery has also been successfully performed, for example, using adeno-associated viruses containing miR-26a in a hepatocellular carcinoma murine model (Kota *et al*., 2009) or *in silica* nanoparticle-based targeted delivery of miR-34a to neuroblastoma tumours in a murine orthotopic xenograft model (Tivnan *et al*., 2012). The latter method used engineered nanoparticles to specifically target cells with high expression of disialoganglioside (GD2) antigen. This antigen is overexpressed on the surface of neuroblastoma cancer cells, making the therapy more targeted towards that specific subtype of cells (Tivnan *et al*., 2012).

In 2013, the first Phase I clinical trial involving miRNA-replacement therapy for cancer in humans commenced, utilising a nanoparticle-derived minicell delivery system to replenish levels of miR-16 in mesothelioma and non-small cell lung cancer (TargomiRs) (Reid *et al*., 2013). This system used mini-cells derived from an asymmetric cell division in bacteria and engineered with epidermal growth factor receptor (EGFR)-specific antibodies. Recently, the outcome of the clinical trial was published with positive results and an acceptable safety profile,

supporting additional studies of miRNA replacement in combination with chemotherapy agents (van Zandwijk *et al*., 2017a) (Figure 1.14).
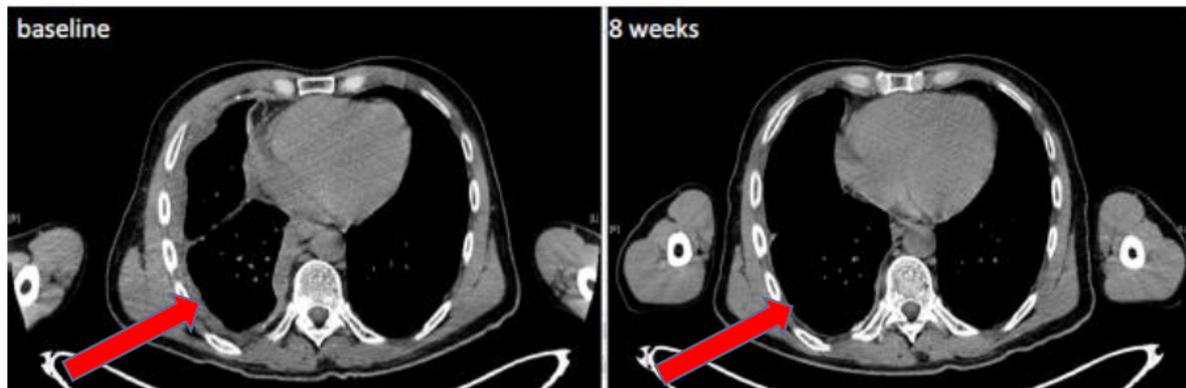
Figure 1.14: **TargomiR treatment of mesothelioma.** Computed tomography (CT) scans from patient lungs with mesothelioma before (on the left) and after (on the right) TargomiR treatment. The red arrow shows the thickness of the lung due to mesothelioma, which disappeared after 8 weeks of treatment. Adapted from (Reid *et al*., 2016).

1.7.7.3 **MiRNAs in body tissues**

Prior to such further work on therapeutic strategy, it is important to make predictions about the tolerance of normal peripheral tissues to exogenously delivered miRNAs. It has been demonstrated that the delivery of miRNAs, that are abundant in normal tissues but lost in diseased cells, can be successfully replenished with minimal toxicity (Kota *et al*., 2009). In order to interrogate this further, in the case of the miRNAs under current interrogation in mGCTs, expression data in most other human tissues was retrieved from the *miRNA body map project* (Mestdagh *et al*., 2011). Fortunately, this shows high expression of miR-99a-5p, miR-100-5p, miR-125b-5p, miR-152-3p and miR-204-5p in most body tissues, and provides a good predictor for the global therapeutic tolerance of replenishment of these miRNAs (Figure 1.15).
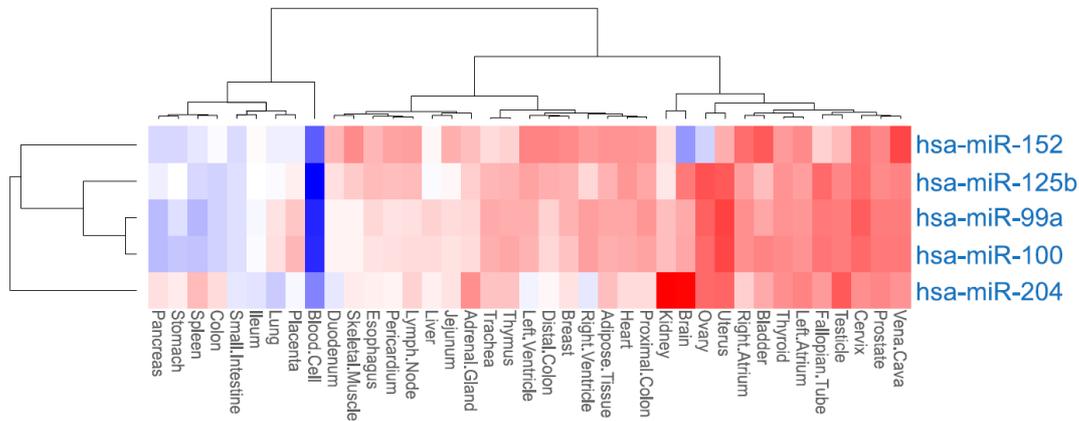
Figure 1.15: **The miRNA body map of the five downregulated miRNAs relevant to GCTs.** High level of expression is colored in red, and is confirmed in most normal tissue types. All miRNAs in the picture are referred as -5p, except for miR-152 which is referred as -3p.

1.8   Aims of the project

The overall aim of this project is to understand the mechanism and functional significance of miRNAs downregulated in mGCTs. As the understanding of the biogenesis and mechanisms of action of miRNAs has grown in the past two decades, their importance in regulating gene expression and essential role in controlling key cellular mechanisms has been highlighted, including contributing to cancer progression when deregulated (Peng and Croce, 2016). As global miRNA profiling segregates mGCTs from non-malignant controls, new avenues for developing miRNA-targeted therapy in the future has been opened for this tumour type. However, in order to develop this, first it's crucial to understand the causes of miRNA dysregulation and their effect on mGCT progression.

The investigation was divided into four major parts:

- **Part I- Providing evidence for the importance of miRNA downregulation in mGCTs compared with non-malignant controls.**

Levels of the most downregulated miRNAs were reanalysed from previous microarray expression data performed in the lab (Palmer *et al*., 2010). From these data, miRNA downregulation was compared between clinical samples derived from mGCTs and representative mGCT cell lines and non-malignant control samples. Subsequently, the relevance of these miRNAs was assessed by looking at the specific seed abundance decrease in the mRNA profile of mGCTs. Raw data from the mRNA microarray were available in the lab (Palmer *et al*., 2010).

- **Part II- Elucidating the cellular level where miRNA downregulation occurs**

  MiRNA deregulation may occur in a variety of ways (Rupaimoole *et al*., 2016). A multitude of steps need to be tightly regulated for normal miRNA production and expression levels. For instance, at the genomic level, an alteration in the copy number of specific loci could affect the quantity of miRNA production (Kloosterman and Hochstenbach, 2014). Similarly, alterations in the biogenesis pathway of miRNA could also affect its maturation (Ha and Kim, 2014).

  In my study, I have focused on the following-

  i)   Genomic DNA copy number analysis was performed to look for evidence of copy number loss at these loci in mGCT cell lines as compared with non-malignant samples.

  ii)   qRT-PCR was used to quantify primary miRNA transcripts (pri-miRNA), -3p counterpart and host genes, to identify relative contributions of the miRNAs and to elucidate the biogenesis stage in which miRNA downregulation occurs.

- **Part III- Understanding the contribution of methylation to miRNA downregulation.**

  Given the importance of DNA methylation in the regulation of miRNAs (Loginov *et al*., 2015), and the role of methylation previously described in mGCTs (Shen *et al*.,

2018), studies were conducted to clarify a possible involvement of methylation in the control of miRNA expression.

i)   DNA demethylation experiments were performed using 5-azacytidine (5-aza) followed by qRT-PCR to quantify the levels of the downregulated miRNAs.

ii)   Semi-quantitative analysis such as bisulfite conversion was performed followed by quantitative analysis such as pyrosequencing to accurately quantify the percentage of methylation on CpG islands.

- **Part IV- Evaluation of downstream consequences of the replenishment of these putative tumour-suppressor miRNAs *in vitro* using mGCT cell lines.**

The next part of the study was to understand the effects of miRNAs replenishment *in vitro* on mGCTs to characterize their importance in the progression of cancer. This step was essential to set the basis for future *in vivo* work.

i)   Phenotypic effects on cancer cell proliferation was assessed, along with the miRNA, mRNA target and protein alterations that underpin these changes.

ii)   Further, in order to identify downregulated protein-coding targets following replenishment, assessment of global mRNA levels was performed using microarray/sequencing technology and systematic interrogation using the bioinformatic tool Sylamer (van Dongen, Abreu-Goodger and Enright, 2008a).

iii)   Pathway analysis of the putative gene target was analysed using different bioinformatic tools and the most significant ones were carried forward for further characterisations.

Part I and Part II of my investigations are presented in Chapter 3 of this thesis, part III is presented in Chapter 4 and Part IV in Chapter 5.

Chapter 2   Materials and Methods

2.1   Cell culture

Reagents for Cell Culture	
Reagent	**Company**
Absolute ethanol	VWR International Ltd, Radnor, PA,USA
Dulbecco's Modified Eagle's Medium (DMEM) with GlutaMAZ-1	Gibco, Thermo Fisher Scientific, Waltham, MA, USA
Dulbecco's Modified Eagle's Medium (DMEM) with GlutaMAZ-1, no sodium pyruvate	
Opti-MEM Reduced serum media	
10X Trypsin-EDTA: porcine trypsin (5 g/ml) and EDTA (2 g/ml)	
RPMI-1640 medium	Sigma-Aldrich, St. Louis, MO, USA
Dimethyl sulphoxide (DMSO)	
Penicillin-Streptomycin: 10,000 units' penicillin + 10 mg streptomycin per ml in 0.9% NaCl (100X stock)	
200 mM L-Glutamine (100X stock)	
Dulbecco's Phosphate Buffered Saline modified without calcium chloride and magnesium chloride (PBS)	
USA HI fetal calf serum	Hyclone; Thermo Fisher Scientific, Waltham, MA, USA

FBS Good	PAN Biotech, Aidenbach, Germany
Trypan blue stain (2X) 0.4% solution for use with Countess cell counter	Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA

Table 2.1: **Reagents used in tissue culture.**

2.2   Cell lines

For *in vitro* studies, four human malignant germ cell tumours (mGCTs) cell lines were selected that reflected the range of malignant GCT histological subtypes commonly observed in clinical practice:, namely Embryonic carcinoma (EC) [2102Ep (Damjanov and Andrews 1983)], yolk sac tumour (YST) [GCT44 (Pera, Blasco Lafita et al. 1987) and 1411H (Vogelzang, Bronson et al. 1985)] and seminoma (Sem) [TCam2 (de Jong, Stoop et al. 2008)].

2102Ep and 1411H were provided by Professor Peter Andrews, University of Sheffield, UK. GCT44 and TCam2 were provided by Dr. Janet Shipley, Institute of Cancer Research, Marsden, Surrey, UK.

Cells were cultured in their respective media (Table 2.2). All cells were grown in T175 or T75 plastic-culture flasks (Corning) at 37°C in an incubator with 5% carbon dioxide (CO_2) in air. 2102Ep were passaged every 4-5 days when they reached over confluency to maintain their embryonal state. TCam2, 1411H and GCT44 were passaged every 3-4 days when 80-90% confluency was reached. For all cell lines, cell passage was performed by removal of supernatant followed by a wash with 10 ml of sterile 1X phosphate buffered saline (PBS,) per T175 flask. After the wash, 5 ml of pre-warmed 0.5% trypsin-ethylenediaminetetraacetic acid (trypsin-EDTA) was added to the monolayer (Table 2.1). Cells were then incubated at 37°C and visually inspected under microscope until they had fully detached from the culture surface. The cells were then centrifuged for 5 min at 300 g. The supernatant was aspirated

and cells resuspended with 6 ml of pre-warmed media, 2 ml of this suspension was added to T175 containing 20 ml of pre-warmed medium.

Cell Lines	Description	Media used	10% FCS	Penicillin (100 U/ml) and Streptomycin (0.1 mg/ml)	Reference
2102Ep (EC)	EC-derived from adult teratocarcinoma with mixed mGCT (mainly EC)	Dulbecco's Modified Eagle Medium (DMEM) (+ 2 mM Glutamine; no sodium pyruvate)	Hyclone Heat-Inactivated fetal calf serum (HI-USA FCS) (10X)	1%	(Andrews *et al*., 1982)
1411H (YST)	YST–derived from a testicular teratomacarcinoma	DMEM+Glut-pyr	Hyclone Heat-Inactivated fetal calf serum (HI-USA FCS) (10X)	1%	(Vogelzang *et al*., 1985)
GCT44 (YST)	YST–derived from metastatic site: para-aortic lymph node	Dulbecco's Modified Eagle Medium (DMEM)	Standard fetal calf serum (FCS) (10X)	1%	(Pera, Köberle and Masters, 1995)
TCam2 (Sem)	Sem–derived from primary lesion of a left testicular seminoma	RPMI-1640 + Glut	Standard fetal calf serum (FCS) (10X)	1%	(Mizuno *et al*., 1993)

Table 2.2: **Cell lines with their respective culture methods used.**

2.2.1 **Cell counting and seeding for experiments**

For cell counting, the cells were trypsinised and resuspended into pre-warmed medium. 10 µl of cells suspension was mixed 1:1 volume with 0.4% trypan blue dye (Thermo Fisher Scientific) and 10 µl of the mixed solution was added to each of the counting chambers A and B of a Countess automated cell counter (Invitrogen). Trypan blue stains dead cells blue and allows distinction between live and dead cells. For transfection experiments, the cells were

seeded in 6-well pates; density at which cells were plated differed between cell lines as indicated in (Table 2.3).

Cell line	Culture dish	Growth area (cm^2)	Cell seeding number x 10^5	Medium volume (ml)
2102Ep	24-well plate	1.9	1.5	0.5
	12-well plate	3.8	3	1
	6-well plate	9.6	7.5	2
1411H	24-well plate	1.9	0.2	0.5
	12-well plate	3.8	0.4	1
	6-well plate	9.6	1	2
GCT44	24-well plate	1.9	0.2	0.5
	12-well plate	3.8	0.4	1
	6-well plate	9.6	1	2
TCam2	24-well plate	1.9	0.2	0.5
	12-well plate	3.8	0.4	1
	6-well plate	9.6	1	2

Table 2.3: **Cell culture seeding density and volumes.**

2.2.2 Cell freezing and thawing

For freezing cells, after trypsinising and counting, the cell suspension was centrifuged at 300 g for 5 min and the supernatant subsequently removed. The cells were then resuspended in freezing medium [90% volume of cell line specific FCS and 10% volume of dimethyl sulphoxide (DMSO)] to achieve a concentration of $1x10^6$ cells/ml and aliquoted into cryogenic vials ('cryovials', Nunc A/S). These vials were placed into a Mr FrostyTM (Thermo Fisher

Scientific) container and stored overnight at 80°C to achieve a gradual cooling process. The following day the cells were moved to a liquid nitrogen tank.

For thawing cells, the cryovials were first transferred on dry ice from the liquid nitrogen container to the tissue culture facility. The cryovials were gently agitated in a water bath at 37°C until almost no ice was visible. The cell suspension was then poured into a 15 ml tube containing 10 ml of pre-warmed culture medium and spun at 300 g for 5 min in a Sorvall Legend T centrifuge (Thermo Scientific, Epsom, Surrey, UK). The supernatant was carefully aspirated and the cell pellet resuspended in 2 ml of the appropriate pre-warmed medium (Table 2.1). This suspension was used to seed a 25 cm^2 (T25) flask containing a further 4 ml of pre-warmed medium. Cells were incubated at 37°C in 5% CO_2.

2.2.3 Cell treatment using mimic transfection

2.2.3.1 Volume of transfection reagents and duration of transfection

The cell lines 2102Ep, 1411H, GCT44 and TCam2 were seeded in 2 ml of growth medium without penicillin-streptomycin (P/S) in 6-well plates (Corning) and allowed to grow to around 30~40% confluency before being treated in triplicate at day 0 (d0) with the transfection reagent Lipofectamine RNAiMAX (Thermo Fisher Scientific, Cat. #13778030). Mimics (mirVana®, Thermo Fisher Scientific) were transfected at three different concentrations: 8.3 nM, 16.7 nM and 33.3 nM. 8 µl of Lipofectamine RNAiMAX was added to 200 µl of specific OptiMEM® transfection media and incubated at room temperature for 20 min. At d0, triplicate wells were treated by transfection with 16.7 nM of miR-100-5p (mirVana® miRNA mimic, Cat. #4464066), 16.7 nM of miR-125b-5p (mirVana® miRNA mimic, Cat. #4464066), 16.7 nM of non-targeting control (NTC) (mirVana® miRNA mimic negative control #1, Cat. #4464058), an equal combination of 8.3 nM miR-100-5p mirVana® miRNA mimic and 8.3 nM miR-125b-5p mirVana® miRNA mimic, a combination of 16.7 nM miR-100-5p mirVana® miRNA mimic and 16.7 nM miR-125b-5p mirVana® miRNA mimic or 33.3 nM of NTC (Table 2.4).

Mimics used	Single mimic concentration	Final concentration of all mimics	Lipofectamine RNAiMAX	Opti-MEM
miR-100-5p	16.7 nM	16.7 nM	8 µl	200 µl
miR-125b-5p	16.7 nM	16.7 nM	8 µl	200 µl
(miR-100-5p+miR-125b-5p)	8.3 nM	16.7 nM	8 µl	200 µl
(miR-100-5p+miR-125b-5p)	16.7 nM	33.3 nM	8 µl	200 µl
NTC	16.7 nM	16.7 nM	8 µl	200 µl
NTC	33.3 nM	33.3 nM	8 µl	200 µl

Table 2.4: **Conditions used for mimic experiment**.

Two 6 well-plates of untreated cells were used for comparison of cell numbers at day 0 (d0), day 2 (d2), day 4 (d4) and day 7 (d7). These studies identified that 8 µl of Lipofectamine RNAiMAX was well tolerated, as long as the duration of transfection was adjusted in a cell line dependent fashion (8 hrs for 2102Ep and GCT44, 6 hrs for 1411H and 4 hrs for TCam2). Transfection was performed using normal media without P/S, as previously described (Murray *et al.*, 2013), before a media change, which was repeated daily thereafter. From the second day after the transfection, the media included P/S. At d2, d3, d4 and d7, cells were trypsinised and RNA extracted.

2.3 DNA work

2.3.1 DNA sample

In the first instance, DNA was extracted from four representative mGCT cell lines: 2102Ep, TCam2, GCT44, 1411H and from testis of a healthy patient samples as normal gonadal control (Genomic DNA – human adult normal tissue: testis, BioChain, Cat. #D1234260).

2.3.2 DNA extraction

For DNA extraction, cells were pelleted in a 15 ml falcon tube and 5 ml of lysis buffer (10 mM Tris-Cl pH 8, 10 mM EDTA pH 8, 150 mM NaCl, 0.4% SDS) with 1 mg/ml protease K was added. Samples were incubated at 37°C overnight. The next day the adequate digestion of cells was confirmed and if not sufficient an extra volume of lysis buffer was added and samples were incubated for an extra 4 hrs at 37°C. Equal volumes of phenol-chloroform were subsequently added to the solution of lysed cells and phase separation performed by centrifugation for 10 min at 1,800 g resulting in a lower organic phase and an upper aqueous phase. The upper aqueous phase was carefully removed and two further washes with phenol/chloroform performed. The phenol-chloroform combination reduces the partitioning of poly(A) and mRNA into the organic phase and reduces the formation of insoluble RNA-protein complexes at the interphase. The pH of phenol determines the portioning of DNA and RNA into the aqueous and the organic phase, respectively.

After the third wash with phenol/chloroform, 5 ml of chloroform was added to completely remove any phenolic and organic contaminants, the sample was subsequently mixed and centrifuged at 1,800 g for 10 min. The top layer was transferred in a clean 15 ml falcon tube and 2 volumes of 100% ethanol were added to one volume of the aqueous solution. Finally, 1/10th of 3 M sodium acetate (pH 5.2) was added to one volume of the aqueous solution and samples were stored at -20°C overnight.

The next day the samples were spun at 4°C, 12,000 g for 10 min to pellet the DNA, supernatant was discarded. The pellet was then washed with 70% cold ethanol and centrifuged at 12,000 g

for 10 min, supernatant was discarded. Finally, the pellet was air-dried for 15 min, re-suspended with 20 µl of nuclease free water, and left overnight at 4°C. DNA was subsequently quantified using a NanoDrop (Thermo Scientific) spectrophotometer to confirm satisfactory purity and obtain the resultant DNA concentration. The samples were diluted if needed and stored at -20°C.

2.3.3 **Primer design**

All primers to assess for genomic loss were designed using the website: 'Primer3' (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) and subsequently ordered from Sigma Aldrich. Genomic sequences were downloaded from the Ensembl Genome Browser (http://www.ensembl.org/index.html) according to Human release GRCh38. Forward and reverse primers were designed to have a length between 18 and 23 nt (optimal size 20 nt), a product size between 80 and 100 nt and a melting temperature between 59°C and 62°C. The designed primer pairs were screened using PRIMER BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). I identified primer pairs that were specific to the intended PCR target. This method uses the Basic Local Alignment Search Tool (BLAST) and a global alignment algorithm to screen primers against a user-selected database in order to avoid primer pairs (all combinations including forward-reverse primer pair, forward-forward as well as reverse-reverse pairs) that may cause non-specific amplification. The optimal primer pairs for each genomic locus (minimum two pairs) were then formally tested by qPCR and assessed for target abundance (i.e. Ct value), primer efficiency calculation and melting curve assessment for the presence of a single product.

2.3.4 **Primer efficiency calculator**

Total DNA from the malignant GCT cell line, 1411H was used to calculate the primer efficiency for each pair of primers. Primer efficiency is the multiple by which the target DNA is amplified per PCR cycle. Under ideal conditions each pair of primers should double the amount of DNA target per cycle, giving an optimal primer efficiency of 2. In reality, values ≥1.80 are considered acceptable. These efficiencies were calculated using a cell line known to

express the target gene sufficiently to be detected within 45 PCR cycles. Six different DNA dilutions were performed (starting from undiluted DNA at a concentration of 10 ng/µl and then having 1:2, 1:4, 1:10, 1:50 and 1:100 dilutions).

All reactions were performed in triplicate in 96-well plates, using the SYBR green qPCR protocol described below and using the Realplex Mastercycler ep-gradient S machine (Eppendorf) (PCR conditions described in Table 2.5).

Stage	Time (min)	Temperature (°C)
Hold	2:00	95°
45 cycles	0:15	95°
	0:20	58°
	0:15	72°
	0:08	76°
Melting curve	8:00	78°
	0:15	95°
	0:15	65°
	0:15	95°

Table 2.5: **PCR conditions for primary efficiency calculations.**

The raw data derived from the fluorescent calculation was exported in Microsoft Excel (.xls) format and the primer efficiency was calculated by using the 'Primer Efficiency Calculator' program designed by Dr. Ian Roberts (ex-member of Coleman's lab). This program plots Ct values for each reaction versus the $\log_{10}$ (relative amount of target amplicon). Linear regression was performed to determine a line of best fit for these data points, and the slope of the line was calculated. Primer efficiency was then determined using this equation:

$$primer\ efficiency = 10 \times \left(-\frac{1}{slope}\right) - 1$$

For each primer pair, target abundance (Ct value), primer efficiency and melting curves (to confirm the presence of only one product) were considered in order to choose the most reliable pair of primers.

2.3.5 **Polymerase chain reaction**

The PCR reactions were conducted using JumpStart™ Taq DNA polymerase (Sigma-Aldrich, #D9307), to find optimal PCR condition to obtain a suitable concentration of DNA to enable sequencing of the fragment of interest and were performed in a MJ Research Tetrad PTC-225 Peltier Thermal Cycler. The reagent for the PCR were combined in 0.2 ml non-flex PCR tube (Starlab) for a total concentration of 25 μl or 50 μl. The reaction components (Table 2.6) and the cycling conditions (Table 2.7) are listed below.

Reagent	Volume per reaction
Buffer 10X	2.5 μl
dNTP (2 μM)	2.5 μl
Forward primer (2 μM)	2.5 μl
Reverse primer (2 μM)	2.5 μl
JumpStart Taq DNA polymerase	0.5 μl
DNA (12.5 ng/μl)	2 μl
Nuclease-free water (NFW)	12.5 μl
Total Volume	25 μl

Table 2.6: **Reagents required for PCR.**

PCR Step	Time (min)	Temperature (°C)
Initial Denaturation	2	94°
40-45 cycles		
Denaturation	0:30	94°
Annealing	0:30	50-60°
Extension	1	72°
Final extention	7	72°

Table 2.7: **PCR reaction conditions.**

The resultant PCR samples were kept at 4°C until further use. The PCR products were then mixed with 1X Orange G tracking dye (SigmaAldrich, Cat. #O3756) and run on a 1% agarose gel alongside a 100 bp ladder (New England Biolabs). The resultant gel was visualised using a Quantity One UV machine (Bio-Rad).

2.3.6 **DNA gel extraction**

Specific DNA bands identified were cut out and gel purified in order to isolate the desired DNA fragment required for sequencing. Extraction of the DNA from the agarose gel was completed using the QIAquick® Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. Briefly, the excised band was placed in a 1.5 ml Eppendorf tube and the mass of the gel fragment determined. 3X gel volumes of Buffer QG were added to each tube, where 100 mg = 100 μl, and the sample incubated at 50°C for 10 min. Once dissolved, 1X gel volume of isopropanol was added and mixed. The sample mixture was then applied to a QIAquick column in a 2 ml collection tube and centrifuged at 17,900 g for 1 min. The flow through was discarded and the column washed with 500 μl Buffer QG followed by 750 μl Buffer PE. The column was then placed into a clean microcentrifuge tube and DNA eluted in 30 μl ddH$_2$O. The concentration of DNA was then determined using a Thermo Scientific Nanodrop 2000 Spectrophotometer. The samples were then placed in a -20°C freezer until further use. The DNA sequencing was completed by the Department of Biochemistry, University of Cambridge according to the specifications of the institute (DNA at 20 ng/100 bp in 10 μl).

2.3.7 **DNA sequencing analysis**

The resultant sequencing files were viewed using DNA sequencing software ChromasPro (Technelysium). Chromatograms were used to assess the quality of the sequence data and the alignment tool was used to compare samples as well as to detect any changes to the expected sequence.

When DNA sequencing analysis was used in the biosulfite experiment, DNA sequences were compared with look at CpG changes between mGCTs and non-malignant controls.

Chromatograms were analysed using ImmageJ and when a double peak was detected, usually C and T in a CpG island, the relative contribution of the C or T was calculated by equating the two peaks to 100% and then assign a proportion for each nucleotide depending on the peak's height.

2.3.8 Quantitative-PCR (qPCR) analysis of gDNA samples

To determine genomic copy number of chromosome 11 and chromosome 21, primers adjacent to the area of miR-100/-125b-1 and miR-99a/-125b-2, were created (Table 2.9). DNA was extracted and measured in triplicate in four representative mGCT cell lines: 2102Ep (EC), 1411H (YST), GCT44 (YST) and TCam2 (Sem) and in a testis patients' samples as a control. Each experiment was performed in triplicate using SYBR$^{®}$ Green JumpStart™ Taq ReadyMix™ (Sigma-Aldrich, Cat. #S4438) (Table 2.8). Test reactions and non-template controls were run in triplicate on a realplex real-time PCR system (Mastercycler$^{®}$ Eppendorf). QPCR cycle conditions were the same as those used for the primer efficiency assay described in Table 2.5.

Reagents used	Volume per reaction
SYBR Green 2.5X	10 μl
Primer mix (4 μM)	2 μl
gDNA (25 ng)	8 μl
Total volume	20 μl

Table 2.8: **Reagents used for qPCR of gDNA.**

Realplex qPCR software was used to automatically determine the fluorescence threshold and return a cycle threshold number (Ct). Melting curve analyses were obtained after each PCR run at 65°C to 90°C and showed a single PCR product, confirming specificity of amplification. Expression ratios of the genomic sequences were calculated using the comparative threshold cycle (Ct) method described by (Pfaffl, 2001), with normalisation to four reference genes

(*B2M, GAPDH, 18A and 18B*). 18A (at position chr18: 9103147-9103241 of the Ensembl version GRCh38) and 18B (at position chr18: 6401843-6401934 of the Ensembl version GRCh38) represent two regions known not to be deregulated in GCTs on chromosome 18. Cts for each particular DNA sample then referenced to the control samples (Table 2.9).

CHROMOSOME 11		
Name	**Forward primer (5' to 3')**	**Reverse Primer (5' to 3')**
MIR100HG	AGACATCTGAGGGGCAACC	GCGAGGAAGCCAAGTTTATG
50Kb upstream miR-100	GTGTATGGGGAAGGCAATGT	GCACGAAGGAAGGATGGTAA
10Kb upstream miR-100	CTAGCCATGTTGCTTGTTTCTTT	AGACACCTTTCATTCTACCACCA
5Kb upstream miR-100	GATAAAACGTGGCACGACAAT	TTTGCCAAAAGATCCAATCAC
miR-100	CCCGTAGATCCGAACTTGTG	TATAAGCAAAGCCCCAGGTC
Let-7a-2	ATGCTCCCAGGTTGAGGTAGTA	GTTTAGTGCAAGACCCAAGGAA
BLID	TGCAAGACACATTCCTTCTCA	AGCAGCAGCAGAGAGGCTAT
10Kb upstream miR-125b-1	CGGCTAGATTGAATGAAAGACC	CATAGTTGGGAACTGAGGCAAT
miR-125b-1	AGAAAACATTGTTGCGCTCCT	AGAGCCTAACCCGTGGATTTA
CHROMOSOME 21		
Name	**Forward primer (5' to 3')**	**Reverse Primer (5' to 3')**
MIR99AHG	AGAGAAGCCGGCACTCAGAT	CAGTTGCGTAGAGTAAAAGTGTGC
50Kb upstream miR-99a	TGAAAACATTGTGTCAGCTTCA	TTAATCCGCAGAGGCAATTC
10Kb upstream miR-99a	AAATTGTGAAACGTTCCCTCTTC	GCATCTGGCCATTATTGAAAGT
5Kb upstream miR-99a	GATAAAACGTGGCACGACAAT	TTTGCCAAAAGATCCAATCAC
miR-99a	CCCGTAGATCCGAACTTGTG	TATAAGCAAAGCCCCAGGTC
Let-7c	ATGCTCCCAGGTTGAGGTAGTA	GTTTAGTGCAAGACCCAAGGAA
10Kb upstream miR-125b-2	CGGCTAGATTGAATGAAAGACC	CATAGTTGGGAACTGAGGCAAT
miR-125b-2	AGAAAACATTGTTGCGCTCCT	AGAGCCTAACCCGTGGATTTA
HOUSE KEEPER GENES		
Name	**Forward primer (5' to 3')**	**Reverse Primer (5' to 3')**
GAPDH	TTGTCAAGCTCATTTCCTGGT	GAATGAGGAACCTCCGGTACA
B2M	CGGCTCTGCTTCCCTTAGACT	TCACAGCCAAGCATTCTACAAAC
18 A	GAGGCACCTCTCGTTTCAAG	TTCTCCCTAAGCCTTCACGA
18 B	GCAGTAGGGCAAGGGTGTAG	TGGAGGTTTGGATTGAGGTC

Table 2.9: **Primers used for genomic copy number assessment in mGCTs cell lines and normal testis.**

2.4   RNA work

2.4.1 RNA extraction (miRNA and total RNA)

MiRNA and total RNA extractions for the various GCT cell lines were performed using TriReagent (Sigma-Aldrich, St Louis, Missouri, USA), following the TriZol RNA extraction protocol (Rio *et al*., 2010). Experiments carried out in 6-well plates were directly treated with 1 ml of TriReagent per well and the resulting suspension transferred to a 1.5 ml Eppendorf tube. 200 μl of saturated chloroform was added, left briefly at room temperature and then spun at 16,000 g for 40 min. The upper aqueous phase containing RNA was then carefully separated from the interphase and pink organic layer into a clean Eppendorf tube. RNA was precipitated by adding 2-propanol, left to stand and then spun in a microcentrifuge at 16,000 g for 5 min. The supernatant was removed and the resultant pellet washed with ice-cold 75% ethanol, and then respun for 5 min at 16,000 g. After removing all liquid with a pipette, the pellet was air-dried for 5 min, dissolved in nuclease-free water and placed on a heat block set at 55°C for 10 min. RNA was then quantified using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA).

2.4.2 cDNA synthesis

Synthesis of complementary DNA (cDNA) from RNA was done using the QuantiTect® Reverse Transcription Kit (Qiagen, Cat. #205310) according to the manufacturer's instructions. In order to ensure the removal of any contaminating genomic DNA, 1 μg of RNA template and 2 μl of Genomic DNA Wipeout buffer were added to an RNase-free 0.2 ml microcentrifuge tube together with nuclease-free water up to a final volume of 14 μl. The reaction was incubated for 2 min at 42°C and then samples placed immediately on ice. Samples underwent the reverse transcription step as per the manufacturer's instructions (Table 2.10).

Reagents	Volume per reaction
Quantiscript RT buffer	4 µl
RT primer mix (polyT plus random hexamer primers)	1 µl
Quantiscript reverse transcriptase	1 µl
Total volume	6 µl

Table 2.10: **Reagents used for mRNA cDNA synthesis.**

The mixture was added to 0.2 ml microcentrifuge tubes containing RNA template. The mixture was then incubated in a GeneAmp 9700 PCR Thermal Cycler at 42°C for 30 min to generate complementary DNA (cDNA), followed by 3 minutes at 95°C to inactivate the enzyme. Following the reverse transcription reaction, the synthesised cDNA was diluted 1:80 using ddH$_2$O, and was immediately used for qPCR amplification or placed at -20°C for long-term storage.

2.4.3 RT-qPCR analysis of cDNA samples

RT-qPCR was used to quantify the level of relative cDNA levels between samples. RT-qPCR was performed using SYBR® Green JumpStart™ Taq ReadyMiX™ Green Master Mix (Sigma), as per manufacturer's instruction (Table 2.11).

Reagents	Volume per reaction
Sigma SYBR Green Buffer	10 µl
Specific primer pair (4uM)	2 µl
(5 ng) cDNA (1:80)	8 µl
Total volume	20 µl

Table 2.11: **Reagents used for mRNA RT-qPCR.**

The RT-qPCR reaction was carried out in triplicate for each primer pair in 96-well white PCR plates (Starlab) using an Eppendorf Mastercycler ep gradient S realplex2. The cycling conditions followed the manufacturer's instructions (Table 2.12).

Stage	Temperature (°C)	Time (min)
Denaturation	95°	2
Cycle (40 Cycles)	95°	0:15
	58°	0:20
	72°	0:15
Final extension	76°	8

Table 2.12: **Reaction conditions required for mRNA RT-qPCR**.

The primer efficiency of each qRT-PCR primer was determined as previously described. Primers sequences are described below (Table 2.13).

Name	Forward primer (5' to 3')	Reverse primer (5' to 3')	Reference/Supplier
LIN 28A	GAAGCGCAGATCAAAAGGAG	CTGATGCTCTGGCAGAAGTG	(Murray *et al.*, 2013)
H1_ARID3A	CATCATGTACACAGGAGTTC	AAGGCAACGAGTTATTTGAG	Sigma-Aldrich KiCqStart™
H3_ARID3B	TCAACAAGAAGATCTGGAGG	CTTTCTTCTCACACTCATAGG	Sigma-Aldrich KiCqStart™
H1_TRIM71	TGGAAGGTAGAAAAGATCCG	TCAAGCTTGTTGAGGTTTTG	Sigma-Aldrich KiCqStart™
H1_CTDSPL	ATGGAAAGAAATGTGTGGTC	ACCTGATGTATAGTTCCATCG	Sigma-Aldrich KiCqStart™
H1_E2F7	CTACAGCTTGATGTTGTTGG	GTAGTTTTCTCAGTTGACAAGG	Sigma-Aldrich KiCqStart™
mTOR	ACAGCCCAGGCCGCATTGTC	TCCAGGGACTCCGTCAGGCG	Persoal design

H1_FGFR 3	GAAGATGCTGAAAGACGATG	GCAGGTTGATGATGTTTTTG	Sigma-Aldrich KiCqStart™
GAPDH	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGA G	(Vandesompele *et al*., 2002)
ACTB	CTGGAACGGTGAAGGTGACA	AAGGGACTTCCTGTAACAAT GCA	(Vandesompele *et al*., 2002)
RPLI13A	CCTGGAGGAGAAGAGGAAA GAGA	TTGAGGACCTCTGTGTATTTG TCAA	(Vandesompele *et al*., 2002)
YWHAZ	ACTTTTGGTACATTGTGGATT CAA	CCGCCAGGACAAACCAGTAT	(Vandesompele *et al*., 2002)

Table 2.13: **Primers used for RT-qPCR analysis of cDNA samples.**

2.4.4 Quantification of transcript level changes

Based on Vandesompele and co-worker's findings (Vandesompele *et al*., 2002), a range of housekeeping genes were tested for each different experimental condition in order to obtain a baseline level of expression for comparative analysis of gene expression. Gene expression ratios were calculated using the comparative threshold cycle (Ct) method described by Pfaffl (Pfaffl, 2001). The house keeping genes used for normalisation were: *GAPDH, YHWAZ, RPL13A* and *ACTB* (Table 2.13).

2.4.5 Pri-miRNA assay

Levels of pri-miR-100; pri-miR-99a, pri-miR-125b-1 and pri-miR-125b-2 were quantified in quadruplicate. The total RNA extracted from cells firstly underwent treatment with TURBO DNA-free kit (Part number AM1907, Ambion, Thermo Fisher Scientific Inc, USA) as per the manufacturer's recommendation. Initially, 100 ng of RNA was used and diluted to a final volume of 10 µl with nuclease-free water. However, due to very low abundance of pri-miRNA product at this concentration of RNA (Ct threshold values generally >40 cycles or undetectable), more RNA input was needed in order to accurately quantify pri-mRNA levels

(1000 ng). Briefly, 1 μl of 10X TURBO DNAse (2U) was added to the initial RNA and 2 μl of DNase inactivation reagent was subsequently added added at the mix. The sample was then incubated for 30 min at 37°C. Subsequently, the samples were incubated for 2 min at room temperature and then centrifuged at 12,000 g for 1.5 min and the supernatant containing RNA was transferred to a fresh 1.5 ml Eppendorf tube.

Samples then underwent reverse transcription using Taqman High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Warrington, UK, Cat. #4368814), which is specifically designed for low abundance targets, as per the manufacturer's instructions (Table 2.14). The samples were then placed in a GeneAmp 9700 PCR Thermal Cycler with the cycling condition listed in (Table 2.15).

Reagents	Volume per reaction
RT buffer	2 μl
Random primers	2 μl
Multiscribe reverse transcriptase	1 μl
Deoxynucleotide triphosphate (dNTP)	0.8 μl
Nuclease-free water (NFW)	4.2 μl
TURBO DNase-treated RNA	10 μl

Table 2.14: **Reagents used for pri-miRNA cDNA synthesis.**

Temperature (°C)	Time (min)
25°	5
42°	30
85°	5

Table 2.15: **Reaction conditions for pri-miRNA cDNA synthesis**.

cDNA was then used to quantify the levels of pri-miRNA levels by qRT-PCR using Taqman 2X Universal primer master mix (No AmpErase UNG) and the relevant pri-miRNAs (pri-miR-100, pri-miR-99a, pri-miR-125b-1 and pri-miR-125b-2) or housekeeping assays: RPLO, GUS-B and 18S (Applied Biosystem) (Table 2.18).

The necessary mix (Table 2.16) was prepared in half-skirted clear 96-well PCR plates (Axygen Biosciences, Union City, California, US), and run on the Realplex Mastercycle ep-gradient S machine (Eppendorf) according to the cycling conditions listed below (Table 2.17).

Reagents	Volume per reaction
Taqman 2X Universal mastermix	10 µl
Relevant pri-miRNA	1 µl
Nuclease-free water (NFW)	5 µl
cDNA product	4 µl

Table 2.16: **Reagents required for TaqMan pri-miRNA qRT-PCR.**

Stage	Temperature (°C)	Time (min)
Hold	95°	10
Cycle (40 Cycles)	95°	0:15
	60°	1

Table 2.17: **Reaction conditions required for TaqMan pri-miRNA qRT-PCR**.

Finally, Ct values were exported in Microsoft Excel (.xls) format from the Realplex Mastercycler and the expression ratios was calculated using the comparative Ct method and normalised to the three housekeepers.

Taqman assay	miRBase accession	Assay ID
hsa-pri-miR-99a	MI0000101	Hs03302729_pri
hsa-pri-miR-100	MI0000102	Hs03302731_pri
hsa-pri-miR-125b-1	MI0000446	Hs03303095_pri
hsa-pri-miR-125b-2	MI0000470	Hs03303224_pri
RPLO	N/A	Hs99999902_m1

GUS-B	N/A	Hs99999908_m1
18S	N/A	Hs99999901_s1

Table 2.18: **Taqman assays used for quantification of pri-miRNA levels.**

2.4.6 **Taqman qRT-PCR**

MiRNA expression levels were quantified using Taqman quantitative real-time PCR (qRT-PCR) assays in two separate steps (Applied Biosystems, Warrington, UK), as per the manufacturer's instructions. The first step was necessary in order to reverse transcribe miRNA into single stranded cDNA. Total RNA was quantified and diluted to 5 ng/µl where necessary. Reverse transcription (RT) was performed using Taqman microRNA Reverse Transcription Kit (Applied Biosystems), adopting specific stem-loop primers for each of the miRNAs tested and housekeeping gene (Cat. #4427975 for all assays, Applied Biosystems) (Table 2.23).

The necessary mix (Table 2.19) was prepared in half-skirted clear 96-well PCR plates (Axygen Biosciences, Union City, California, US), and run on a GeneAmp 9700 PCR Thermal Cycler (Applied Biosystems) according to the cycling conditions listed below (Table 2.20).

Reagent	Volume per reaction
dNTP (100 mM)	0.1 µl
Multi-scribe RT enzyme (50 U/µl)	0.67 µl
RT buffer (10X)	1 µl
RNase inhibitor (20 U/µl)	0.13 µl
Nuclease-free water (NFW)	2.77 µl
Specific primer	2 µl
Relevant RNA	3.33 µl
Total volume	10 µl

Table 2.19: **Reagents required for mature miRNA cDNA synthesis**.

PCR Step	Time (min)	Temperature (°C)
Pre-incubation	30	16°
Annealing/extension	30	42°
Enzyme inactivation	5	85°
Hold	∞	4°

Table 2.20: **Reaction conditions for miRNA cDNA synthesis.**

The second step was undertaken using the TaqMan qPCR kit (Applied Biosystems, Warrington, UK) to quantify the cDNA RT products in triplicate on a Starlab 96-well plate. Initially the concentration to detect miRNA-5p and miRNA-3p used was 15 ng, but further optimisation was necessary to detect miRNA-3p strand as the product at this concentration of RNA was undetectable (Ct values > 37 cycles) and the concentration was increased to 100 ng. The necessary mix (Table 2.21) was prepared in half-skirted clear 96-well PCR plates (Axygen Biosciences, Union City, California, US), and run on a Realplex Mastercycle ep-gradient S machine (Eppendorf) according to the cycling conditions listed below (Table 2.22).

Reagent	Volume per reaction
Taqman Universal PCR 2X Master Mix	7.5 μl
Nuclease-free water (NFW)	5.42 μl
Specific 20X miRNA assay probe	0.75 μl
Specific RT product (cDNA)	1.33 μl
Total volume	15 μl

Table 2.21: **Reagents required for TaqMan miRNA qRT-PCR step.**

qRT-PCR Step	Temperature (°C)	Time (min)
TaqMan activation (Hold)	95°	10
Denaturation X40	95°	0:15
Annealing/extension X40	60°	1

Table 2.22: **Reaction conditions for TaqMan miRNA qRT-PCR**.

Each of the two steps were run with RT primers and qPCR probes specific to each of miR-99a-5p/3p, -100-5p/3p, -125b-5p, -125b-1-3p and -125b-2-3p, as well as *RNU24*, which was previously selected as an abundant and stable 'housekeeping' gene for specific use in GCT tissues and cell lines (Torres *et al.*, 2012), and is commonly used as a reference gene for miRNA studies.

Ct values (calculated automatically by the Realplex Mastercycler) were transferred to Microsoft Excel for analysis. Quantification of expression levels of selected miRNAs was performed using the ΔΔCt method (Livak and Schmittgen, 2001). Expression levels were normalised to *RNU24*.

Taqman assay	miRBase accession	Assay ID
hsa-miR-99a-5p	MIMAT0000097	000435
hsa-miR-99a* (-3p)	MIMAT0004511	002141
hsa-miR-100-5p	MIMAT0000098	478224
hsa-miR-100-3p	MIMAT0004512	478619
hsa-miR-125b-5p	MIMAT0000423	477885
hsa-miR-125b-1-3p	MIMAT0004592	478665
hsa-miR-125b-2-3p	MIMAT0004603	478666
hsa-miR-204-5p	MIMAT0000265	000508
hsa-miR-152-3p	MIMAT0000438	000475

hsa-miR-26b-3p	MIMAT0004500	483077
hsa-miR-137-3p	MIMAT0000429	001129
hsa-miR-500-3p	MIMAT0002871	001046
hsa-let-7b	MIMAT0000063	002619
hsa-let-7d	MIMAT0000065	000380
hsa-let-7e	MIMAT0000066	000381
RNU24	N/A	001001
RNU38B	N/A	001004
RNU44	N/A	001094

Table 2.23: **Taqman assays used for quantification of miRNA levels.**

2.5   Protein analysis

2.5.1 Total protein extraction

Cells were washed with cold PBS and lysed in cold radioimmunoprecipitation assay (RIPA) lysis buffer (Thermo Fisher Scientific) (approximately 80 µl per well of a 6-well plate) supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche Diagnostics Ltd., Burgess Hill, UK) (1:25 dilution). Cells were then removed from the dish using a cell scraper and were transferred into a pre-cooled Eppendorf tube. After 15 min of continuous agitation on a vibrating shaking platform at 4°C, cellular debris was removed by centrifugation (15 min, 4°C, 14,000 g). The supernatant was transferred to a fresh 1.5 ml Eppendorf tube and immediately quantified or placed at -80°C for long term storage.

2.5.2 **Protein quantification**

Total protein concentration was determined using Pierce bicinchoninic acid (BCA) protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instruction with some modification. Briefly, protein samples were diluted with PBS (2 μl of protein sample added to 12 μl of PBS, 1:7 dilution) and 3 μl of diluted sample added to the well of 96-well plate in triplicate. The BSA stock standard at 1 mg/ml concentration was used to generate a standard curve and was added to each well at 5 different volumes in triplicate: 0 μl (Blank), 1 μl, 3 μl, 5 μl, 7 μl and 10 μl. The provided Reagent A and Reagent B were mixed together in a ratio of 1:50 and 200 μl of the mixture was added to each well and incubated for 30 minutes at RT in the dark. Absorbance was measured at a wavelength of 570 nm using a Dynex Technologies plate reader and Revelation software. Protein concentrations were determined using a standard curve (Figure 2.1).
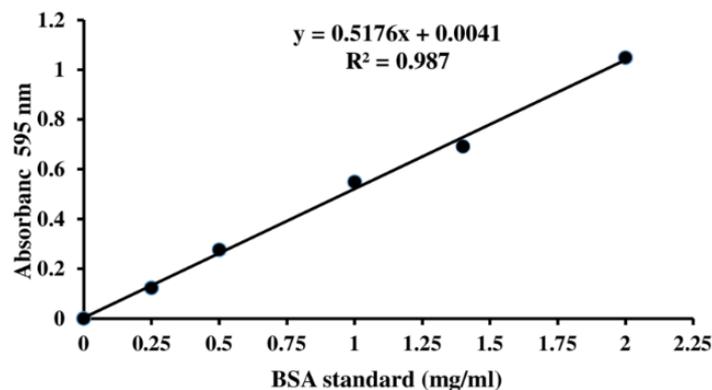
Figure 2.1: **Example of a standard curve used for BCA assay to determine protein concentration based on Absorbance values.**

2.5.3 **Protein sample preparation and SDS-PAGE separation**

Quantified protein extract (25 μg) was added to 5 μl of NuPAGE™ LDS Sample Buffer and 2 μl of NuPAGE™ Sample Reducing Agent (both from Novex, Thermo Fisher Scientific Cat. #NP0008, Cat. #NP0004, respectively) and made up to a final volume of 20 μl with 1X PBS.

Samples were denatured for 10 min at 95°C and run immediately on a gel or stored at -20°C. Pre-cast 4-12% NuPAGE™ Bis-Tris Mini Gels (Thermo Fisher Scientific) were placed in an XCell SureLock™ Mini-Cell (Thermo Fisher Scientific) and the inner tank filled with NuPAGE™ MOPS SDS Running Buffer 20X (Thermo Fisher Scientific Cat. #NP0001) containing NuPAGE™ Antioxidant (Thermo Fisher Scientific, Cat. #NP0005) and the outer tank filled to the level of the well base with NuPAGE MOPS SDS Running Buffer without Antioxidant. The samples were loaded, along with 10 µl of See Blue molecular weight ladder (Novex, Thermo Fisher Scientific, Cat. #LC5625). Proteins were separated by electrophoresis at 150 V until the desired protein separation was achieved (approximately 1.5 hrs).

2.5.4 Western blotting

A PVDF membrane (Thermo Fisher Scientific) and two pieces of 3MM Whatman filter paper were cut to the size of the SDS-PAGE gel. The PVDF membrane was activated by incubation in 100% methanol for 15 sec followed by a 2 min rinse in distilled water (dH2O). The PVDF membrane was then transferred to ice-cold Transfer Buffer (Thermo Fisher Scientific) (supplemented with 20% methanol and NuPAGE Antioxidant) for at least 5 min to equilibrate. The whatman filter paper and four blotting sponges were also pre-soaked in Transfer buffer.

Following electrophoresis, protein samples separated by SDS-PAGE were transferred to the PVDF transfer membrane using the XCell SureLock Mini-Cell and XCell II™ blot module. The gel cassette was removed from the mini-cell and the two plates of the cassette separated using a gel knife. Wells of the gel were subsequently removed with a gel knife and a piece of pre-soaked filter paper was placed on top of the gel. The gel and filter paper were carefully peeled away from the cassette and the equilibrated PVDF membrane placed on top of the gel, ensuring no air bubbles were present. An additional piece of pre-soaked filter paper was placed on top of the membrane.

Two pre-soaked blotting sponges were placed on top of the cathode core of the XCell II™ blot module. The filter paper – gel – membrane – filter paper 'sandwich' was placed on top of the sponges and an additional two pre-soaked sponges placed on top of the 'sandwich'. The lid of the blot module was placed on top and the whole assembly inserted into the XCell

SureLockTM Mini-Cell. The blot module was filled with Transfer Buffer, and ice-cold dH20 was added to the outer reservoir. The protein transfer was performed for 2 hrs at 30 V at 4°C.

After completion of the transfer, the blot module was dismantled and the PVDF membrane stained with Ponceau-S Staining Solution (Sigma-Aldrich) to assess the quality of the protein transfer. Membranes were washed first with dH2O, then three times for 5 min with TBS with 0.5% Tween-20 (TBS-T) and incubated for 1 hour at room temperature with blocking buffer (5% skimmed milk in 1% TBS-T). Membranes were incubated at 4°C overnight with primary antibodies diluted in antisera buffer (5% skimmed milk in 0.1% TBS-T) (Table 24). The following day, the membranes were washed three times for 5 min with TBS-T and incubated for 1 hr at room temperature in species specific HRP-conjugated secondary antibodies diluted in blocking buffer. Following three further washes with TBS-T, the membrane was developed using enhanced standard chemiluminescence (ECL) or ECL prime western blotting detection reagents (Amersham, GE Healthcare, Little Chalfont, UK) according to the manufacturer's instructions and exposed to autoradiography films at various time-points. If required, the blots were treated with Stripping Buffer (Thermo Fisher Scientific) for 15 min incubation at 37°C in the dark; and blots incubated with blocking buffer and re-probed with different primary antibodies. The images were scanned and densitometry analyses were performed using FluorChem-9900 imaging system software (Alpha-Innotech, San Leandro, U.S), with normalisation to a loading control (Actin β or β Tubulin) and referenced to the chosen control samples set to 1.

A PVDF membrane (Thermo Fisher Scientific) and two pieces of 3MM Whatman filter paper were cut to size of the SDS-PAGE gel. The PVDF membrane was activated by incubation in 100% methanol for 15 sec followed by a 2 min rinse in distilled water (dH2O). The PVDF membrane was then placed in ice-cold Transfer Buffer (Thermo Fisher Scientific) (supplemented with 20% methanol and NuPAGE Antioxidant) for at least 5 min to equilibrate along with the 3MM Whatman paper and 4 sponges. Protein samples separated by SDS-PAGE were transferred to the PVDF transfer membrane using the XCell SureLock Mini-Cell and XCell IITM blot module. Firstly, XCell SureLock Mini-Cell was dismantled, gel separated from one of its plates and top lanes and bottom part of the gel cut away. A sheet of 3MM filter paper was then adhered to the gel and gel carefully peeled away from second plate. The

equilibrated PVDF membrane was then carefully placed on top of the gel ensuring that no air bubbles were present. Another sheet of 3MM filter paper was placed on top of the PVDF membrane and the `sandwich' placed on top of two pre-soaked 2 sponges that had been placed on top of the cathode core of the XCell IITM blot module. Another 2 pre-soaked sponges were then placed on top of the `sandwich', the lid of the blot module placed on top and the whole assembly inserted into the XCell SureLockTM Mini-Cell. The blot module was topped up with Transfer Buffer, the outer reservoir was filled with ice-cold dH2O and whole device placed in a 4°C cold room. The protein transfer was performed for 2 hrs at 30 V. After completion of the transfer, the blot module was dismantled and the membranes were stained with Ponceau-S Staining Solution (Sigma-Aldrich) to assess the quality of the protein transfer to PVDF membranes. Membranes were washed first with dH2O, then three times for 5 min with TBS with 0.5% Tween-20 (TBS-T) and incubated for 1 hour at room temperature with blocking buffer (5% skimmed milk in TBS-T). Membranes were incubated at 4°C overnight with primary antibodies diluted in blocking buffer (Table 2.24). The following day, the membranes were washed three times for 5 min with TBS-T and incubated for 1 hr at room temperature in species specific HRP-conjugated secondary antibodies diluted in blocking buffer. Following three further washes with TBS-T, the membrane was developed using enhanced standard chemiluminescence (ECL) or ECL prime western blotting detection reagents (Amersham, GE Healthcare, Little Chalfont, UK) according to the manufacturer's instructions and exposed to autoradiography films at various time-points. If required, the blots were stripped with Stripping Buffer (15 min incubation at 37°C in the dark; Thermo Fisher Scientific) and blots incubated with blocking buffer and re-probed with different primary anti-bodies. The images were scanned and densitometry analyses were performed using FluorChem-9900 imaging system software (Alpha-Innotech, San Leandro, U.S), with normalisation to a loading control (Actin β or β Tubulin) and referenced to the chosen control samples set to 1.

Target	Antibody catalog number	Host species	Supplier	Dilution
LIN-28	ab46020	rabbit	abcam	1:5000
LIN-41 (E-1)	sc-393352	mouse	Santa Cruz Biotechnology	1:300
ARID3A (A-4)	sc-398367	mouse	Santa Cruz Biotechnology	1:300
CTDSPL	NBP1-53169	rabbit	Novusbio	1:300
Caspase-3	#9662	rabbit	Cell signalling Technology	1:500
Cleaved caspase-3 (Asp175)	#9661	rabbit	Cell signalling Technology	1:500
β tubulin	ab6046	rabbit	abcam	1:10000
Rabbit	P044801	goat	Dako	1:2000
Mouse	P044701	goat	Dako	1:2000

Table 2.24: **Antibodies used for Western Blotting.**

2.6   Methylation experiment

2.6.1 Demethylation using 5-azacytidine

For this work, four GCT cell lines: 1411H (YST), 2102Ep (EC), GCT44 (YST) and TCam2 (Sem) were used. Cells were seeded at a low concentration ($7x10^4$ cells/well) to allow them to grow for up to 7 days in triplicate in six-well plates. 48 hrs after seeding, the cells were treated with 10 μM of the demethylation reagent 5-azacytidine (5-aza) (Sigma-Aldrich, Cat. #A2385) which was previously dissolved at the working concentration in different media depending on the cell line treated (Table 2.2). Complete media change was performed daily for four days.

Negative controls (appropriate media without 5-aza) were performed in parallel using exactly the same conditions as described above. 96 hrs after drug exposure, cells were trypsinised and RNA extracted in order to evaluate any changes in miRNA levels and host gene caused by demethylation. Levels of miR-99a-5p/-100-5p, miR-125b-5p, MIR100HG, BLID and MIR99AHG were evaluated. As a control, levels of miRNAs known not to be deregulated in mGCTs were tested. The results were then normalised; for this work three different housekeeping genes were used: *RNU24*, *RNU44* and *RNU38B*.

2.6.2 **Bisulfite conversion**

For this work, we used DNA derived from four mGCT cell lines: 1411H (YST), 2102Ep (EC), GCT44 (YST) and TCam2 (Sem). Due to the lack of cell lines representing normal testis or ovaries, DNA from one human adult normal testis tissue (BioChain, Cat. #T1234260), three normal cervix cell line (95:14, 95:15 and NCX6), one normal fibroblast cell line (HFF-1) and one normal retina cell line (RPE) were used as a control.

1000 ng of DNA was bisulfite converted as per the manufacturer's instructions using either EpiTec Bisulfite Kit (Qiagen, Hilden, Germany, Cat. #59104) or EZ DNA Methylation kit (Zymo research, D5001). In order to obtain bisulfite converted DNA, a 140 µl PCR reaction was prepared using the reagents in Table 25 and the samples were then placed in a MJ Research Tetrad PTC-225 Peltier Thermal Cycler, which was run according to the cycling conditions listed in (Table 2.26) below.

Reagent	Volume per reaction
DNA solution	Variable (max 20 µl)
RNase-free water	Depending on DNA quantity
Bisulfite Mix	85 µl
DNA protect Buffer	35 µl
Total volume	140 µl

Table 2.25: **Reagents used to perform biotin conversion using EpiTec Bisulfite Kit.**

PCR Step	Time (min)	Temperature (°C)
Denaturation	5	95°
Incubation	25	60°
Denaturation	5	95°
Incubation	85	60°
Denaturation	5	95°
Incubation	175	60°
Hold overnight	∞	20°

Table 2.26: **Reaction conditions to perform biotin conversion using EpiTec Bisulfite Kit.**

Unmethylated cytosine residues were deaminated to uracil while methylated cytosine (5-mC) residues remained unaffected, enabling PCR amplification to recognize uracils as thymines and 5-mC as cytosines.

The subsequent morning the Cleanup of bisulfite converted DNA was performed as per the manufacturer's instructions. Briefly, 560 µl of Buffer BL (all the buffer listed are proprietary of Qiagen) was added to the PCR reaction. The sample mixture was then applied to an EpiTect spin column and placed in a centrifuge and spun at 17,900 g for 1 min. The flow through was discarded and the column washed with 500 µl of Buffer BW followed by 500 µl of Buffer BD and incubated for 15 min at room temperature. The column was then spun at 17,900 g for 1 min and washed twice with 500 µl of Buffer BW. The column was then placed into a clean microcentrifuge tube and bisulfite converted DNA eluted in 30 μl ddH2O. The samples were then placed in a -20°C freezer until further use.

2.6.3 **Primer design specific for CpG-rich island**

All primers to assess for CpG enrichment were designed using the website: 'MethPrimer' (http://www.urogene.org/methprimer/) and subsequently ordered from Sigma Aldrich. Genomic sequences were downloaded from the Ensembl Genome Browser (http://www.ensembl.org/index.html) according to Human release GRCh38.p12 and bisulfate converted manually. Forward and reverse primers were designed in order to have a length of -

24 nt a product size of 200-500 nt and a melting temperature of 62-68°C. Additionally, the primers were designed in order to give two different melting curve products depending on the methylated or unmethylated status of the amplified sequence.

A range of primers spanning up to 10 kilo bases upstream the area of interest at chromosome 11 and chromosome 21 were used (Table 2.27).

CHROMOSOME 11		
Name	Forward primer (5' to 3')	Reverse Primer (5' to 3')
MIR100HG UP TO 10KB #1	TTTTTTTATTATATGTTAATTTTTTTAGTT	AAACACATTCTAATATTCTCTACCATTATA
MIR100HG UP TO 10KB #2	GGTTAGGAGTGGTGGTTGGA	TCCTACCTCAACCTTACTAAACAAC
MIR100HG UP TO 10KB #3	GTTGAGGTAGGAGAATTGTTTGAAT	CCACTAAACCCCACTAACTTTAATAAA
MIR100HG UP TO 5KB	AGAGGGTTGGTTGGTGTATTT	TCACTCTCTACCCCTCCTCA
miR-100 UP TO 10KB	AGGGGTATGGAGAAATGAAGGA	ACTCAACTCCCATTTCACTCAA
miR-100 UP TO 5KB	TTATTTGAAATGAAAAATTTTTTTT	TTATAATAAAACAAAATAAACCCAAC
miR-125b-1 UP TO 5KB #1	GGGGTAGTGTGGTTTAGGAAA	CCCCACCTTCAAAACAACTTT
miR-125b-1 UP TO 5KB #2	TTTATTTTTAGTTTGATGAGGAAAG	CACCAAACTATCATTTAATAAACAC
miR-125b-1 UP TO 5KB #3	GATTTAGTTAAGTAGTTGATGTAGA	AACCTCCTATAACACCATCC
miR-125b-1 UP TO 5KB #4	GGATGGTGTTATAGGAGGTT	CTTTCCTTCACAAAATTAAAATTC
BLID up to 5Kb	GGAGTGTAGTGGTGTGATTTTA	AACAACATACAAAAAATACATCCTA

CHROMOSOME 21		
Name	Forward primer (5' to 3')	Reverse Primer (5' to 3')
MIR99AHG UP TO 10KB	TGGGTATTAGTAGTGGTGGTTGTAG	CTACCTCCTCAAATAAATCCCTAAC
miR-99a up to 5Kb	TTAAAATTGTTATTGTGAAAATAATA	TCTAAAAAAAACATCTAAATAAATC
miR-99a up to 10Kb #1	GGTTGAAATTTGGATATTTTAGGTATTAG	AAACCACAATAAAAACAACAACAAAC
miR-99a up to 10Kb #2	GTTAATGAATGTGTTATTTGTTTT	ATTACCCCTTTTCCATATATAAAC
miR-125b-2 up to 5Kb	TTGTAGTTTAGTAGGTTTGGGAGAG	TAACTAACTAACTTTTACCCTCCAC

Table 2.27: **List of primers used for the assessment of CpG rich region on chromosome 11 and 21.**

2.6.4 Pyrosequencing

Fragments for pyrosequencing were generated by polymerase chain reaction, using the settings previously described (Table 2.6, Table 2.7). Primers used were the same as shown in Table 2.27, but one of the primers (Forward or Reverse) was biotinylated (from biomers.net). Fragments from the PCR products were then run on a 1% agarose gel to assess the concentration. For purification of biotinylated fragments, 10 to 15 µl of PCR product was added to a mix consisting of 2 µl Streptavidin Sepharose HP™ Beads (Amersham Biosciences, Freiburg, Germany) and 40 µl binding buffer (Qiagen). The mix was shaken at a maximum speed (1,400 rpm) for at least 10 min on a BioShake iQ (Qinstruments). Single-stranded fragments were purified using the Vacuum Prep Tool™ (Qiagen). Sepharose beads with the single-stranded templates attached were added to a PSQ 96 Plate Low™ (Qiagen, Hilden, Germany) containing a mix of 24.25 µl annealing buffer (Qiagen) and 0.75 µl of 10 µM pyrosequencing primer (final concentration 0.3 µM/well) (Table 2.28). Pyrosequencing was performed in a PyroMark MD™ Q24 System (Qiagen) with the PyroGold Q24™ Reagent Kit (Qiagen) containing nucleotides and prepared mixtures with enzymes and substrates, respectively. For pyrogram exposure including CpG-site methylation calculation, the

PyroMark Q24 Software (Qiagen, version 2.0.8) was applied. Only pyrograms including sharp peaks with sufficient height for each injected nucleotide of interest and without background controls were considered.

Island	Biotinylated primer	Sequence primer assay #1	Sequence primer assay #2	Sequence primer assay #3	Sequence primer assay #4
Isla#4 (MIR100 HG.11.4)	TGAGGAGGGGTA GAGAGCGA	TAAATTTTTAGT TAGATATAGAA T	ATAGAAAAGTTT TTTAAGTTTTTA T	TAATGGTATT TGTAGTGGG A	TGTTTTTTAGT AGTTTAAAGG
Isla#4 (MIR100 HG.11.4)	AGAGGGTTGGTT GGTGTATTT	ATTTAGAAATC GAGTTAGTTG	AGTTAGTTTTAG AGAGGGGTT	GTTGAGGAG GGGTAGAGA G	
Isla#5 (miR-100.11.1)	TTGAGTGAAATG GGAGTTGAGT	GTTAAGTTTGA AAGATGATGA	GATTAATGTAGA AGGTTGG	GATTAGAAT ATAAGATGG A	
Isla#8 (miR-125b-1.11.1)	AAAGTTGTTTTG AAGGTGGGG	GTGTGGTTTAG GAAAATTT	GATTAAATGTGT TTTTTAAAG	AATGGTAAA AGAGAAAAG G	AAGGGAGAGA AAAAGGAGAG
Isla#9 (miR-125b-1.11.2)	GTGTTTATTAAAT GATAGTTTGGTG	TTTTTAGTTTGA TGAGGAAA	GGTTAAAGGTTT TAAAGAATT		
Isla#16 (miR-125b-2.21.1)	GTGGAGGGTAAA AGTTAGTTAGTT A	GAGTTAATGTT AATGATGAG	GTTTTGAGTTAG TATTAGAT		

Table 2.28: **Sequence primers design for pyrosequencing.** Depending on the quantity of CpG sites embedded in the PCR fragment generated using primers described in Table 2.27, a different number of assays was designed in order to cover as many CG as possible. Pyrosequencing design is described in detail in the section 4.2.4 of this thesis.

2.7   Global mRNA expression profiling

2.7.1 Samples used

Global mRNA profiling was performed on a total of 24 samples treated with miRNA mimic as per section 1.1.4.1 at day 7 (Table 2.29):

- 3 x untreated
- 3 x treated non-targeting control using a total concentration of 16.7 nM
- 3 x treated using combination of miR-100-5p + miR-125b-5p for a final concentration of 16.7 nM

Global mRNA profiling was performed on a total of 12 samples treated with miRNA mimic as per section 1.1.4.1 at day 2 (Table 2.29):

- 3 x treated non-targeting control using a total concentration of 16.7 nM
- 3 x treated using combination of miR-100-5p + miR-125b-5p for a final concentration of 16.7 nM

Cell Line	Number of Replicates	Day	condition	Single mimic concentration	Final concentration of all mimics
1411H	3	2	Non targeting control	16.7 nM	16.7 nM
	3	2	Combination of miR-100-5p + miR-125b-5p	8.3 nM	16.7 nM
Tcam2	3	2	Non targeting control	16.7 nM	16.7 nM
	3	2	Combination of miR-100-5p + miR-125b-5p	8.3 nM	16.7 nM
1411H	3	7	untreated	16.7 nM	16.7 nM
	3	7	Non targeting control	16.7 nM	16.7 nM
	3	7	Combination of miR-100-5p + miR-125b-5p	8.3 nM	16.7 nM
Tcam2	3	7	untreated	16.7 nM	16.7 nM

	3	7	Non targeting control	16.7 nM	16.7 nM
	3	7	Combination of miR-100-5p + miR-125b-5p	8.3 nM	16.7 nM
2101Ep	3	7	Non targeting control	16.7 nM	16.7 nM
	3	7	Combination of miR-100-5p+miR-125b-5p	8.3 nM	16.7 nM

Table 2.29: **Samples used for Illumina array.** The table shows number of replicates and the different conditions per cell line used in the array performed at day 2 and day 7.

2.7.2 mRNA microarray protocol

RNA was assessed for concentration and quality using a SpectroStar (BMG Labtech, Aylesbury, UK) and a Bioanalyser (Agilent Technologies, Cheadle, UK). Microarray experiments were performed at Cambridge Genomic Services, University of Cambridge, using the HumanHT-12 v4 Expression BeadChip (Illumina, Chesterford, UK) according to the manufacturer's instructions. Briefly, 200 ng of total RNA underwent linear amplification using the Illumina TotalPrep RNA Amplification Kit (Life Technologies, Paisley, UK) following the manufacturer's instructions. The concentration, purity and integrity of the resulting cRNA were measured by SpectroStar and Bioanalyser. Finally, cRNA was hybridised to the HumanHT-12 v4 BeadChip overnight followed by washing, staining and scanning using the Bead Array Reader (Illumina).

2.7.3 Bioinformatics analysis

Bioinformatics analysis was performed by the Cambridge Genomic Services at University of Cambridge. Briefly, the data are loaded in Illumina proprietary software, GenomeStudio. No background correction or normalisation is applied at this stage. A "final report" is generated, generating a text file containing the sample probe profile and the control probe information. The report is then processed in R using the lumi package (Du, Kibbe and Lin, 2008) and the limma package (Ritchie *et al.*, 2015). The data are loaded and divided into subsets according to the groups being compared; only the samples involved in a given comparison are used. Subsets were then filtered to remove any non-expressed probes using the detection p-value

from Illumina. Across all samples probes for which the intensity values were not significantly different (p>0.01) from the negative controls were removed from the analysis. Following filtering the data were transformed using the Variance Stabilization Transformation (VST) from lumi and then normalised to remove technical variation between arrays using quantile normalisation. Comparisons were performed using the limma package with results corrected for multiple testing using False Discovery Rate (FDR) testing. Finally, the quality of the data were assed and the correlation of the samples in the groups compared.

2.7.4 mRNA microarray data processing using Sylamer

Sylamer analysis was conducted with Keaan Amin and Dr. Anton Enright, member of the Cambridge Genomic Services at University of Cambridge. The *Sylamer* analysis algorithm (http://www.ebi.ac.uk/enright/sylamer/) (van Dongen, Abreu-Goodger and Enright, 2008b) was used to assess enrichment and/or depletion of nucleotide words of specific length complementary to elements of the seed region (nucleotide positions 1-8) of microRNAs (i.e. SCRs) in the 3'UTRs of genes within ranked lists, with significance calculated using hypergeometric statistics. The primary aim of *Sylamer* was to identify whether changes in microRNA expression are of biological significance, with the secondary aim of producing lists of target genes for further validation.

2.7.5 Algorithm details

For each ranked genelist derived from the mRNA expression data, *Sylamer* analysis was undertaken for the six SCR elements of increasing size: three hexamers (corresponding to microRNA seed positions 1-6 nt, 2-7 nt and 3-8 nt); two heptamers (positions 1-7 nt and 2-8 nt); and one octamer (position 1-8 nt). Due to over-representation of conserved adenosines flanking SCRs in mRNAs (Lewis, Burge and Bartel, 2005b), the complementarity criterion was discarded for SCR position 8 (seed position 1), where the nucleotide was always set to be adenosine, irrespective of the actual nucleotide at that position. For each comparison analysed, the mRNA genelist was first ranked from down-regulated (to the left) to up-regulated (to the right). For each SCR under consideration, an enrichment/depletion *p*-value was computed at different cut-offs in the ranked gene list. At each cut-off, an SCR was either enriched in the

3'UTRs of the genes to the left and accordingly depleted in the 3'UTRs of the genes to the right, or conversely depleted on the left and enriched on the right. An event of enrichment on one side and corresponding depletion on the other side of the cut-off was associated with a single *p*-value. Varying the cut-off resulted in a set of *p*-values for each SCR (*y*-axis) visualised on a landscape plot (van Dongen, Abreu-Goodger and Enright, 2008a), in which the $\log_{10}$-transformed *p*-values were sign-adjusted and plotted against the ranked genelist (*x*-axis). Sign-adjustment depended on the specific enrichment/depletion status of the pertinent SCR. A point plotted along the positive *y*-axis signified that the SCR was enriched in the genes to the left and depleted in the genes to the right, whereas a point plotted along the negative *y*-axis conversely signified depletion to the left and enrichment to the right. The displacement along the *y*-axis identified the significance of the joint enrichment/depletion *p*-value for the SCR at that cut-off, according to the sign-adjusted $\log_{10}$ transformation.

2.7.6 Single-summed significance score

The single summed significance score is an integration of Sylamer significant scores and an overall evaluation of the enrichment or depletion of nucleotide sequences with enhanced sensitivity. The scores were calculated by combining the Sylamer results for one hexamer (2-7 nt), two heptamers (1-7 nt, 2-8 nt) and the octamer of the miRNA SCR, because the middle 2- to 7-nucleotide region is critical for the mRNA binding function of miRNAs (Lewis, Burge and Bartel, 2005a). Based on the calculation method of single summed significance scores, Enright's lab developed an algorithm to extract Sylamer results for individual nucleotide sequences and visualised the single summed significance scores with line charts for the first time.

2.7.7 Functional enrichment analysis

Reactome, KEGG and Gene Ontology (GO) are the three major tools used in this project for functional enrichment analysis of lists of Ensembl gene IDs produced from the Sylamer and Illumina analysis. Reactome is a publicly-available database containing manually curated, peer-reviewed information of biological processes and pathways (Fabregat *et al.*, 2018). KEGG (Kyoto Encyclopedia of Genes and Genomes) is an integrated database resource consisting of

eighteen databases. They are broadly categorized into: systems information, genomic information, chemical information and health information, which are distinguished by color coding of web pages. The KEGG database has been in development by Kanehisa Laboratories since 1995, and is now a prominent reference knowledge base for integration and interpretation of large-scale molecular data sets (Kanehisa and Goto, 2000). Gene Ontology (GO) performs enrichment analysis on gene sets (Ashburner *et al.*, 2000). We used Reactome, KEGG and Gene Onthology website to identify the pathways and functions associated with the genes of interest which may imply any changes of gene expression contributing to the initiation and progression of tumorigenesis of mGCTs.

2.8   FACS

2.8.1   Cell cycle analysis

For cell cycle analysis, cells were analysed at d1, d2, d3 and d4 after transfection using Click-iT™ EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Thermo Fisher Scientific, #C10419). Reagents were prepared according to manufacturer's instructions. A stock solution of 10 mM EdU was prepared which was then conserved at -20°C for up to 1 year. EdU was used at a working concentration of 10 μM by dilution of the stock solution. Cells were incubated for up to 3 hrs with 10 μM of EdU. Cells from the same population not treated with EdU acted as a negative controltreated. After the optimal incubation time, cells were washed with 3 ml of 1% BSA in PBS, pelleted and supernatant removed. The remaining EdU stained pellet was then washed again with 3 ml of 1% BSA, pelleted and supernatant removed. The pellet was then dislodged using 100 μl of Click-iT™ fixative (4% paraformaldehyde in PBS). Cells were incubated for 15 min protected from light. Cells subsequently underwent a 1% BSA wash, were centrifuged and the supernatant was removed. The cell pellet was resuspended in 100 μl of 1X Click-iT™ saponin-based permeabilisation and wash reagent. 500 μl of Click-iT™ reaction cocktail containing Fluorescent dye azide (Alexa Fluor® 647 azide, #C10424), reaction buffer and copper (II) sulfate ($CuSO_4$) in a 100 mM aqueous solution, were added to the tube and samples incubated for 30 min in the dark. Cells were then washed once with 3 ml of 1X Click-

iT™ saponin-based permeabilisation and wash reagent, centrifuged and the pellet resuspended in 500 μl of Click-iT™ saponin-based. 1 μl of FxCycle™ Violet (Thermo Fisher Scientific, #F10347) was then added to the mixture to stain the DNA content of the cells and samples analysed using flow cytometer BD LSRFortessa machine (BD Biosciences) at the Wellcome-MRC Cambridge Stem Cell Institute. For the detection of EdU with Alexa Fluor® 647 azide 633/635 nm excitation with a red emission filter (660/20 nm) was used. Flow cytometer data were analysed using FlowJo® (version 10.5.0).

2.8.2 Apoptosis analysis

Cell death and apoptosis was measured using annexin V-FITC (BioLegend, Cat. #640905). Cells were washed twice with cold BioLegend cell staining buffer (Cat. #420201) and then resuspend in Annexin V Binding Buffer (Cat. #422201) at a concentration of 1×10^6 cells/ml. Cells were then transferred in a 5 ml FACS test tube and 5 μl of FITC Annexin V and 1 μl of PI (Cat. #421301) were added. The cells were then incubated for 15 min at 37°C in the dark. 400 μl of Annexin V Binding Buffer (Cat. #422201) was added to each tube and samples analysed by flow cytometry.

2.8.3 Statistical analysis

Statistical analyses were performed using GraphPad Prism 6 software (GraphPad Software, La Jolla, U.S.). For comparisons between groups, an unpaired, two-tailed Students t-test was used. P-values <0.05 were considered statistically significant. Data are presented as mean ± SEM.

Chapter 3   Understanding the origin of the upstream regulation of miR99a-5p/-100-5p and miR-125b-5p

3.1   Introduction

MiRNAs dysregulation has been shown to affect most of the hallmarks of cancer, including continuous proliferative signalling, evasion of growth inhibition, resisting cell death, activating invasion and metastasis and inducing angiogenesis (Cimmino *et al.*, 2005; Klein *et al.*, 2010).

Since their discovery, it has become apparent that dysregulation of miRNAs is complex and can occur at different levels. One of the most common mechanism of dysregulation takes place at genomic level. Most human miRNAs are encoded in introns of both coding and non-coding transcripts although they can also be found in exons. MiRNAs, including some which are downregulated in GCTs such as *miR-100~ let-7~ miR-125b,* are often clustered. Although the distance between clustered miRNAs is generally less than 1 Kb, larger clusters do exist, and a distance as far as 50 Kb between clustered miRNAs has been described (Baskerville and Bartel, 2005; Yuan *et al.*, 2009). Clustered miRNAs have been shown to be generally transcribed as polycistrons and have similar expression patterns. Some miRNAs, especially when they are located in introns of genes, are expressed with their host gene, implying that they share the same promoter (Baskerville and Bartel, 2005).

More than half of the human miRNAs are located in genomic regions of high instability – such as fragile sites and cancer susceptibility loci – and structural genetic changes have been

described in a variety of solid tumours including breast, lung, ovarian, colon, gastric and hepatocellular carcinoma as well as in leukemias and lymphomas (Calin *et al*., 2004b). Genetic changes identified, include partial chromosome deletions, loss of heterozygosity (LOH), amplification of common minimal region of amplification, and mutations at common breakpoint region. Natural sequence variation, such as SNPs, have also been found to affect miRNA biogenesis, for example changes in the composition of the sequence recognised by Drosha-mediated processing complex can lower the production of a mature miRNA (Kumar *et al*., 2008; Volinia *et al*., 2010), but tumour specific mutations appear to be infrequent.

Transcription factor such as p53, MYC, ZEB1 and ZEB2 are important regulators of miRNAs expression by their direct binding to the promoter of miRNAs genes. They induce both repression or upregulation of transcription, and their function can also be differential in a tissue-specific manner (Kim, Han and Siomi, 2009).

MiRNAs dysregulation can also occur at post-transcriptional level due to aberrant transcription factor activity. For example, MYC mediated activation of LIN28A and B results in reduced levels of mature let-7 due to the binding of LIN28A and B to the terminal loop of pri- and pre-let-7 which in turn blocks the cleavage by Drosha and Dicer respectively needed for miRNA maturation (Murray *et al*., 2013).

Recent studies have also found a role for epigenetic mechanisms in miRNAs silencing, due to promoter hypermethylation, which is a well-known mechanisms that cancer cells adopt to suppress the production of genes (Lopez-Serra and Esteller, 2012).

Interestingly, dysregulation of miRNAs is seen across the clinical spectrum of mGCTs. This was originally established in our group by profiling 615 miRNAs in paediatric mGCTs, control tissue and cell lines (48 samples in total) by microarray. It was found that most of the differentially-expressed miRNAs were down-regulated in GCTs compared with controls and benign tumours, suggesting an overall tumour suppressing role for miRNAs in these tumours which is common to other malignancies (Palmer *et al*., 2010).

The five-top-ranking putative tumour suppressor miRNAs identified in this study were then chosen for further investigation. The five most downregulated miRNAs, ranked by adjusted p-value, with at least a mean 4-fold decrease in paediatric GCTs compared with non-malignant samples are: miR-152-3p, miR-204-5p, miR-99a-5p, miR-100-5p and miR-125b-5p (in descending order of significance). Moreover, a similar pattern of downregulation of miR-99a-5p/-100-5p and miR-125b-5p, was noticed which might suggest co-regulation of these miRNAs. The mechanisms directing the downregulation of miRNAs in these tumours are not fully understood and are the subject of work presented in Chapter 4.

The aim of this work is to further characterise downregulated miRNAs in mGCTs. Firstly, reanalysis of the five-top-ranking putative tumour suppressor miRNAs will be performed across a panel of mGCT tissue samples and cell lines when compared with control and benign samples. Secondly, the importance of each of them will be assessed by studying specific seed abundance decrease in mRNA profile available in the lab (Palmer *et al*., 2010), their genomic characteristic and their expression profile derived from the previously published miRNAs heatmap (Palmer *et al*., 2010). Finally, investigation to understand at which level the miRNAs dowregulation occur will be evaluated.

3.2 Results

3.2.1 MiRNA expression in mGCTs

3.2.1.1 MiRNA levels in a panel of tissues from GCTs and representative GCT cell lines

In previous work using microarray, five miRNAs, namely miR-99a-5p, miR-100-5p, miR-125b-5p, miR-204-5p and miR-152-3p, were identified as the most downregulated miRNAs in mGCTs after analysis of expression data for 615 miRNAs in 42 paediatric samples comprising 12 yolk sac tumours (YST), 13 seminomas (Sem), 3 embryonal carcinomas (EC), 6 teratomas (Ter), 8 normal gonadal tissues as well as 6 cell lines derived from adult and paediatric teratomas (Tera2, and PA1 respectively), adult embryonal carcinomas (2102Ep), yolk sack cell tumours (GCT44, 1411H), and adult seminoma (TCam2) (Palmer *et al*., 2010). Analysis of the individual values of the negative log2 fold-change (log$_2$FC) for each miRNA demonstrated a significant downregulation of miR-99a-5p, miR-100-5p, and miR-125b-5p across all mGCT tissue samples and cell lines compared with control and benign samples (teratomas) (Figure 3.1). In contrast, miR-204-5p and miR-152-3p, appeared to be variably downregulated in the mGCT tissue samples, and no significant changes were observed in the representative cell lines, making future replenishment strategies *in vitro* and *in vivo* models unlikely to be successful.

a)

b)

c)

d)

e)
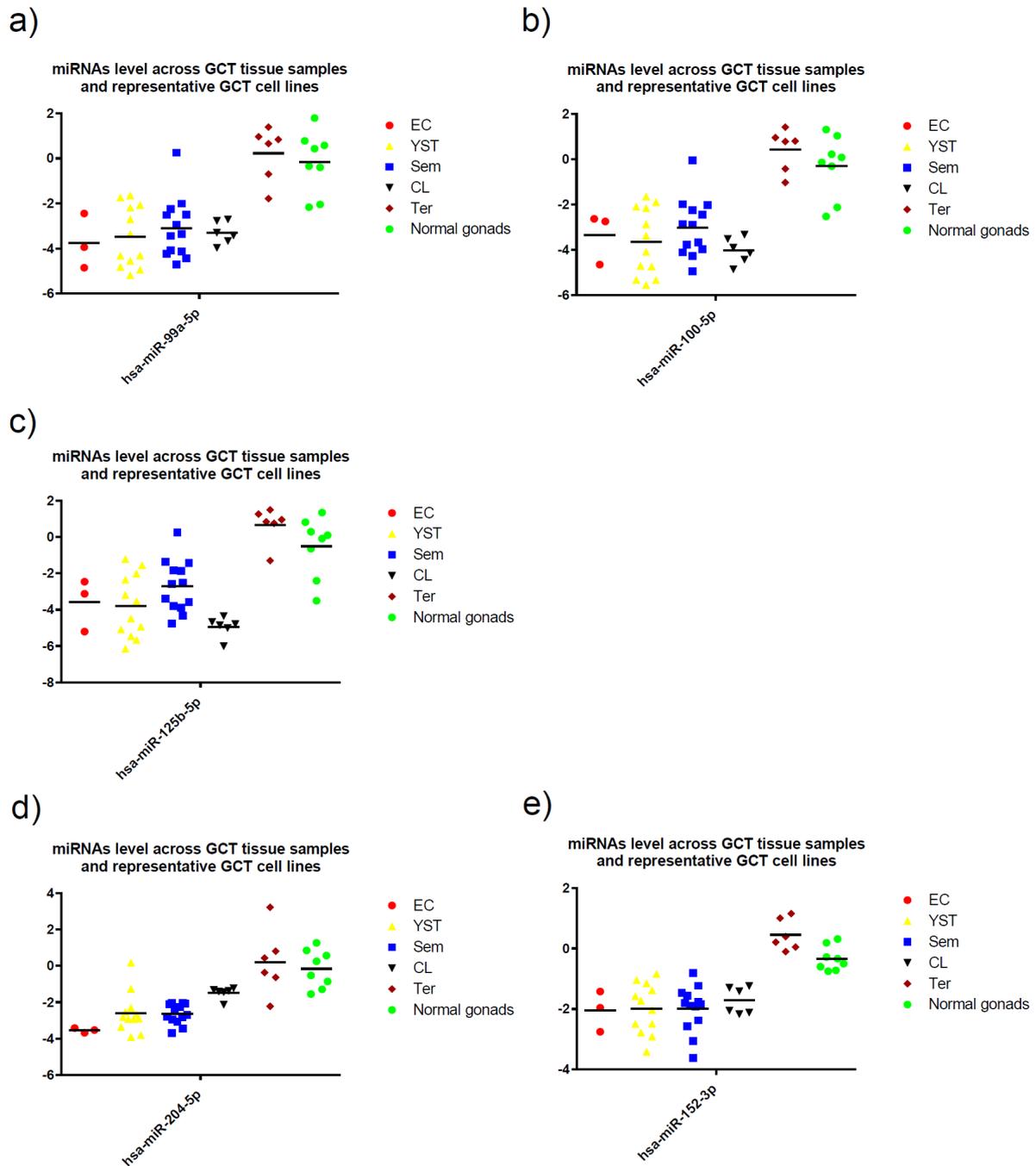
Figure 3.1: **Relative expression of five selected miRNAs in GCT tissues.** a) hsa-miR-99a-5p, b) hsa-miR-100-5p, c) hsa-miR-125b-5p, d) hsa-miR-204-5p and e) hsa-miR-152-3p. Samples are colour coded as follows: embryonal carcinoma (EC) in red, yolk sac tumour (YST) in yellow, seminoma (Sem) in blue, teratoma (Ter) in brown. Normal gonads which were used as a control are shown in green and cell lines (CL) in black. The black line indicates the average log2 fold-change (log2FC) for the specific miRNA. Data from Palmer *et al.*, 2010.

3.2.1.1 **Abundance of miRNAs 2-7 nt seed regions across GCT samples**

In the last ten years, it has become apparent that there is a great deal of redundancy in miRNA regulation. In addition to the existence of multiple genomic loci producing the same miRNA, many miRNAs are part of miRNA "families," which target the same seed sequence. Hence, there is a great deal of redundancy at the genomic level. Because the seed sequence is small, a single miRNA can target a large number of mRNAs. The opposite is also true, that a single mRNA can be targeted by numerous miRNAs, including those from different families. The miRNAs identified as the most downregulated in mGCT share seed sequences with other miRNA but these families comprise no more than 3 members. In detail, the 2-7 nt seed of miR-99a-5p/-100-5p is identical to that of miR-99b; the same is true for miR-125b-5p and miR-125a; miR-152-3p and miR-148a and miR-148b; and miR-204-5p and miR-211 and miR-623. The small size of these miRNA families suggests that any perturbation of any of the miRNAs belonging to the same family is likely to have a significant effect on the relevant overall seed abundance and result in a biological effect. Interestingly, miR-99a-5p and miR-100-5p not only share the 2-7 nt seed sequence, but their entire mature sequence is almost identical and only differs by 1nt towards the 3' end of the sequence (Table 3.1).

Putative tumour suppressor miRNAs	Sequence
miR-99a-5p	AACCCGUAGAUCCGA<u>U</u>CUUGUG
miR-100-5p	AACCCGUAGAUCCGA<u>A</u>CUUGUG

Table 3.1: **Sequences of mature mir-99a-5p and miR-100-5p**. The sequences are almost identical, underlined is the only nucleotide that is different.

In view of the similarity between the two miRNAs, the specificity of the TaqMan qRT-PCR detection assays for each individual miRNA was tested in the lab. Briefly, miRNAs were reverse transcribed using the primer for a specific assay (e.g. miR-99a-5p or miR-100-5p), and then amplified using the correct probe/primer mix for the assay of interest (e.g. miR-99a-5p). Reverse transcription with an irrelevant primer (i.e. miR-125b-5p) was included to control for aspecific amplification (Figure 3.2).
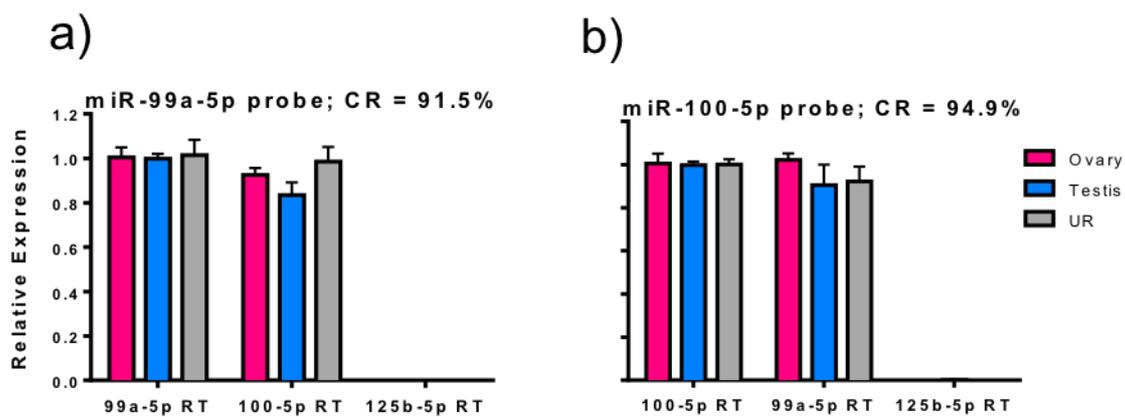
Figure 3.2: **MiR-99a-5p and miR-100-5p cross-reactivity assay**. Samples were reverse transcribed with primers specific for miR-99a-5p, miR-100-5p and miR-125b-5p (x-axis) and then amplified with either a probe/primer mix specific for miR-99a-5p (a) or miR-100-5p (b). Expression is shown relative to the matched combination of primer in RT and primer/probe PCR assay. Templates used were a reference RNA from human ovarian tissue, a reference RNA from human testis (both from Ambion) and the Universal Reference (UR) which is derived from a mix of 14 commonly used cell lines. Headers show mean percentage cross-reactivity (CR). Adapted from Christos Panayi's NST, Part II Pathology Project Report, 2014.

These tests demonstrated that the individual assays are not able to discriminate between the two miRNAs. Therefore, either assay quantifies both miR-99a-5p and miR100-5p, making the individual contribution for each miRNA unquantifiable. These data however suggest that these miRNAs are likely to have a complete functional overlap.

To confirm that not only the individual miRNAs but also the overall 'family' seed abundance was decreased, a reanalysis of the raw microarray data (Palmer *et al*., 2010) was performed with bioinformatic support (Enright group, Department of Pathology, Cambridge). Individual

data from miRNAs that share the same seed sequence, were pooled together and a normalised intensity ratio of the seed was calculated (Figure 3.3).
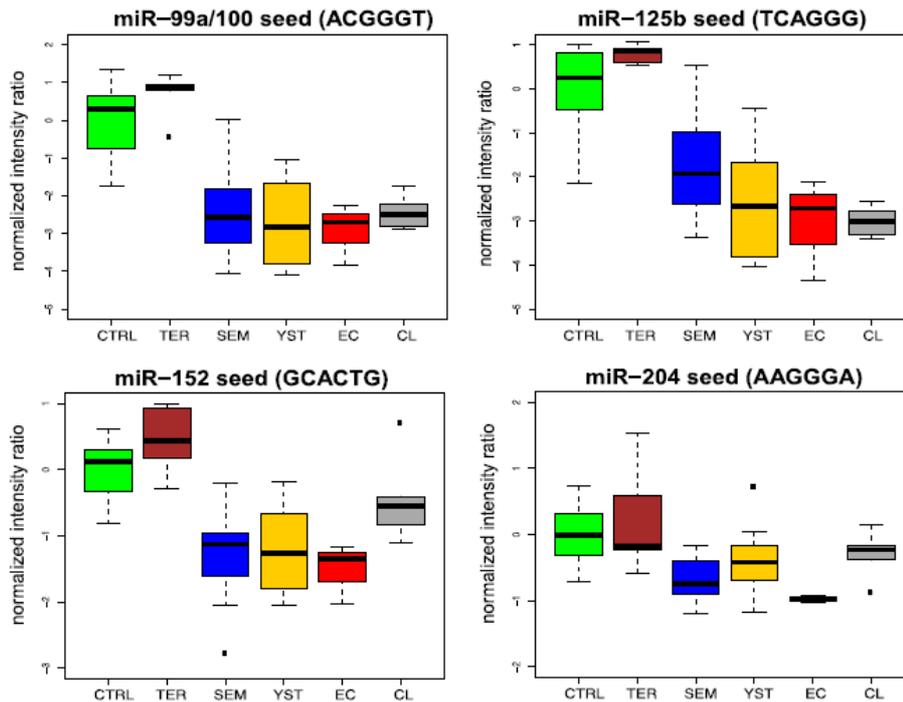
Figure 3.3: **Seed abundance in GCT samples.** Samples are (left to right): CNTRL (normal gonads, control tissue), TER (teratomas, benign GCTs), SEM (germinomas, mGCT), YST (yolk sac tumours, mGCTs), EC (embryonal carcinomas, mGCT) and CL (GCT derived cell lines). All miRNAs listed in this figure are quantified at the -5p end except miR-152 which is quantified at the -3p end. Data analysed with support by Dr Harpreet Saini.

In agreement with the data on the individual miRNAs, this analysis demonstrates that the 2-7 nt seed sequences of miR-99a-5p/miR-100-5p and miR-125b-5p are substantially decreased in both clinical mGCT samples and representative GCT cell lines, compared with non-malignant control samples (normal gonadal controls and benign teratomas). Moreover, the 2-7 nt seed sequence of miR-152-3p is not as highly downregulated across the cell lines, there is no significant downregulation in mGCT samples and cell lines of the 2-7 nt seed of miR-204-5p compared with non-malignant control samples. Overall these data suggest a more biologically

relevant role for the downregulation of miR-99a-5p/miR-100-5p and miR-125b-5p, than for the other two miRNAs originally identified from the array data.

3.2.2 Investigation of potential genomic alteration as the reason for miRNAs downregulation

3.2.2.1 Genomic context

As a result of the array data analysis, further work was carried out investigating the effects and mechanisms of the downregulation in mGCTs of only three of the putative tumour suppressor miRNAs: miR-99a-5p, miR-100-5p and miR-125b-5p.

Interestingly, the mature form of miR-125b-5p is the product of two genomic loci, each of which is physically clustered with either miR-99a-5p or miR-100-5p and a *let-7* family member. The cluster on chromosome 11q24.1 contains miR-100 and miR-125b-1, while the cluster on chromosome 21q21.1 contains miR-99a and miR-125b-2 (Figure 3.4).
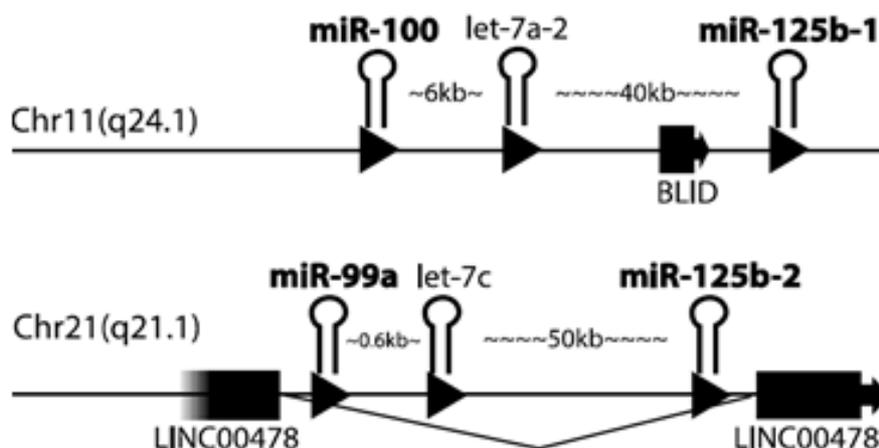
Figure 3.4: **Genomic view of the two miRNA clusters.** Chromosome 11 containing miR-125b-1/-100 and chromosome 21 containing miR-125b-2/-99a. Taken from (Roush and Slack 2008).

Interestingly, the original microarray data (Figure 1.12) showed a very similar pattern of expression of these three miRNAs in each tissue examined, which may suggest co-regulation (Figure 3.5).

Figure 3.5: **MiR-99a-5p/-100-5p and miR-125b-5p microarray data**.

Furthermore, linear regression analysis of previous qRT-PCR validation data, revealed a very strong positive correlation between miR-99a-5p/-100-5p and miR-125b-5p levels, again suggesting co-regulation of the 11q and 21q loci (Figure 3.6).
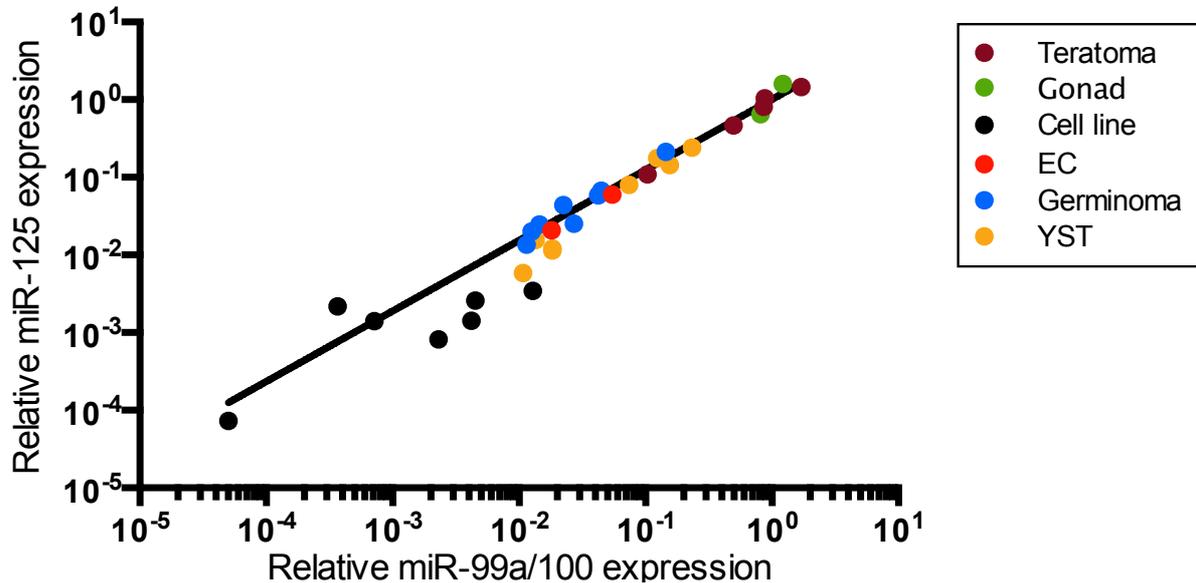
Figure 3.6: **Linear regression analysis of the relative expression of miR-99a-5p/-100-5p (x-axis) and miR-125b-5p (y-axis).** Adapted from Christos Panayi's NST, Part II Pathology Project Report, 2014.

3.2.3 **Identification of potential genomic imbalances**

3.2.3.1 **Identification of reference genes for genomic copy number quantification**

To study the mechanisms mediating the downregulation of these miRNAs, we investigated whether any genomic deletion or insertion might have occurred at genomic level. This was carried out using a qPCR method to establish relative copy number of the miRNA genomic loci relative to other reference genes, for which copy number is known.

To identify which genes were suitable for data normalisation, four candidate housekeeper genes (HK) were selected from regions known to be copy number neutral in mGCTs, namely, *GAPDH*, *B2M*, *18A* and *18B* (Terashima, Yu, Wing-Yuk T. Chow, *et al*., 2014), and levels analysed using q-PCR. This was performed on DNA from four representative GCT cell lines and DNA extracted from a normal testis tissue sample. The reliability of the HK genes was evaluated using two criteria: the first one considered the standard deviation between treated and untreated samples, while the second assessed the consistency of expression of the HK genes within individual samples by looking at relative expression of the housekeeping genes vs each other. For an HK to be acceptable therefore, the standard deviation (σ) of its expression in different samples has to be ≤ 1, while its expression ratio to the other HKs has to be similar in all samples. The four reference genes, showed small variation in the raw Ct values obtained between cell lines (Figure 3.7) and overall an intra cell line deviation standard lower than 1 ($\sigma < 1$), and they all had the same expression ratio across the samples. Therefore, these genes were used as reference for further experiments.
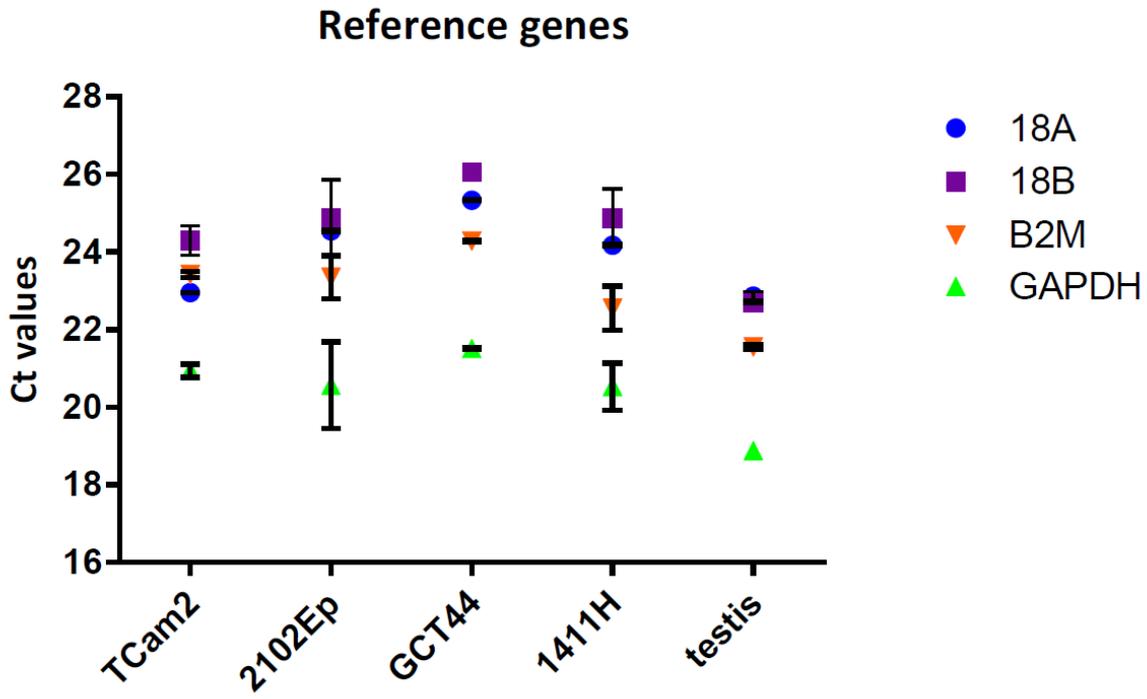
Figure 3.7: **Raw Ct values obtained after amplification of genomic regions of four candidate reference genes in four representative GCT cell lines and normal testis tissue samples.** Data presented as mean ± SEM of n=3 biological replicates.

3.2.3.1 Assessment of genomic copy number in mGCT cell lines

Potential genomic copy number imbalances were assessed by q-PCR for both chromosome 11, where the miR-100, *let-7a-2* and miR-125b-1 cluster is located, and chromosome 21 where the miR-99a, *let-7c* and miR-125b-2 cluster resides. To this end, seven different primers were used to screen the entire region of interest on either chromosome 11 and 21.

Primers were designed to quantitate areas 50Kb, 10Kb and 5Kb upstream of the miRNAs of interest as well as specifically hsa-pri-miR-100 (ch11) and hsa-pri-miR-99a (ch21), let-7a-2 (ch11) and let-7c (ch21) and hsa-pri-miR-125b-1 (ch11) and hsa-pri-miR-125b-2 (ch21) (Figure 3.8).
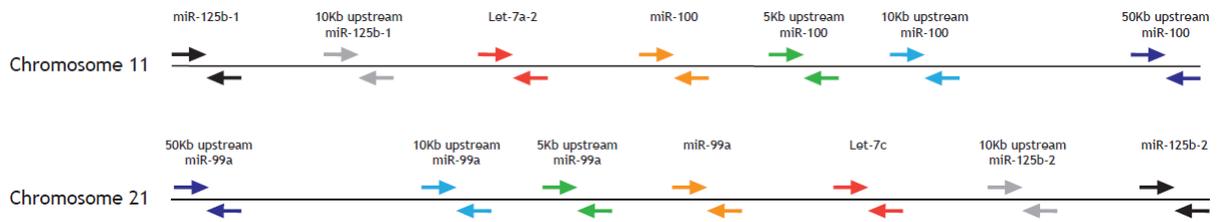
Figure 3.8: **Location of primer pairs used for copy number assessment on either chromosome 11 or 21.** Chromosome distribution is not to scale.

Quantification was performed across five different DNA samples obtained from four representative GCT cell lines and one normal testis tissue. Normal testis tissue DNA was used as genomic diploid reference and copy number for the four cell lines established relative to this, after appropriate normalisation to their housekeeper genes. For all cell lines, there was no consistent variation in the genomic copy number across the region of interest (ROI) on both chromosome 11 and chromosome 21 (Figure 3.9). These values showed significant correlation ($p < 0.05$) across all the cell lines.
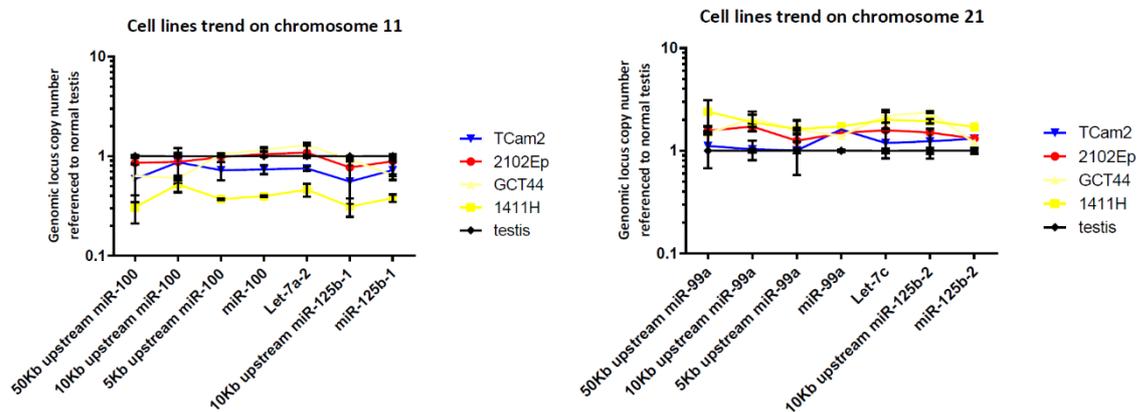
Figure 3.9: **Genomic copy number quantification.** Summary genomic copy number quantification for the regions of interest on chromosome 11 (left panel) and chromosome 21 (right panel) for four mGCT cell lines and control testis tissue DNA. Data presented as mean ± SEM of n=3 biological replicates.

3.2.4 **Assessment of pri-miRNA levels**

Having excluded genetic aberration as a potential cause of miRNAs downregulation, we investigated at which level of miRNAs biogenesis the downregulation occurred. The first step in the biogenesis of miRNAs is the production of primary miRNAs transcripts (pri-miRNAs). Importantly, the pri-miRNAs of each individual miRNAs studied is different as it includes not only the sequence of the mature miRNA, but also regulatory sequences that are unique to each chromosome. This allows the study of each loci separately (Table 3.2).

pri-miRNA name	Sequence
hsa-pri-miR-100	CC<u>UG</u>UUG<u>CC</u>A<u>C</u>AAACCCGUAGAUCCGA<u>A</u>CUUGUGGU<u>AUUAGUCCG</u>C<u>A</u>CA<u>AG</u> <u>CUUGUAUCU</u>A<u>UA</u>G<u>GUA</u>U<u>GUGUC</u>U<u>GUUA</u>G<u>G</u>
hsa-pri-miR-99a	CCCAUUGGCAUAAACCCGUAGAUCCGAUCUUGUGGUGAAGUGGACCGCACA AGCUCGCUUCUAUGGGUCUGUGUCAGUGUG
hsa-pri-miR-125b-1	<u>UG</u>C<u>GCUC</u>C<u>U</u>C<u>U</u>CA<u>GUCCC</u>UGAGA<u>CCC</u>UAACUUGUGA<u>U</u>G<u>UUUA</u>C<u>CGUUUAAA</u> <u>UCCACGGGU</u>UA<u>GG</u>C<u>UCUU</u>GGG<u>AG</u>C<u>UGC</u>G<u>AGUC</u>G<u>UGC</u>U
hsa-pri-miR-125b-2	ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUAUUUUAGUAA CAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA

Table 3.2: **Pri-miRNAs sequences.** Nucleotide differences in pri-miR-100 and pri-miR-125b-1 using pri-miR-99a and pri-miR-125b-2 respectively as the canonical sequence are underlined.

3.2.4.1 **Optimisation of pri-miRNA experimental conditions**

Various optimisation steps were performed in order to achieve accurate quantification of pri-miRNA levels. Most importantly optimisation of input was key due to the low abundance of these transcripts. Input as high as 1000 ng was eventually established as the optimal amount for detection.

To identify which HK genes were suitable for data normalisation, three candidates from previously published work were selected (Murray *et al.*, 2013) and levels analysed using qRT-PCR (*18s*, *GUSB* and *RPLO*). To ensure these genes were suitable across cell lines, the four representative GCT cell lines (2102Ep, 1411H, GCT44 and TCam2) as well as RNA from two gonadal control samples (ovary and testis mRNA, Ambion) were tested. *18S* was the only HK with a higher standard deviation between samples (σ=1.23) compared with *GUSB* (σ=0.5) and *RPLO* (σ=1). However, *18S* maintained the same expression ratio between HKs across all the samples and because of the low number of possible HK genes available for the experiment, all of the three HKs were used (Figure 3.10).
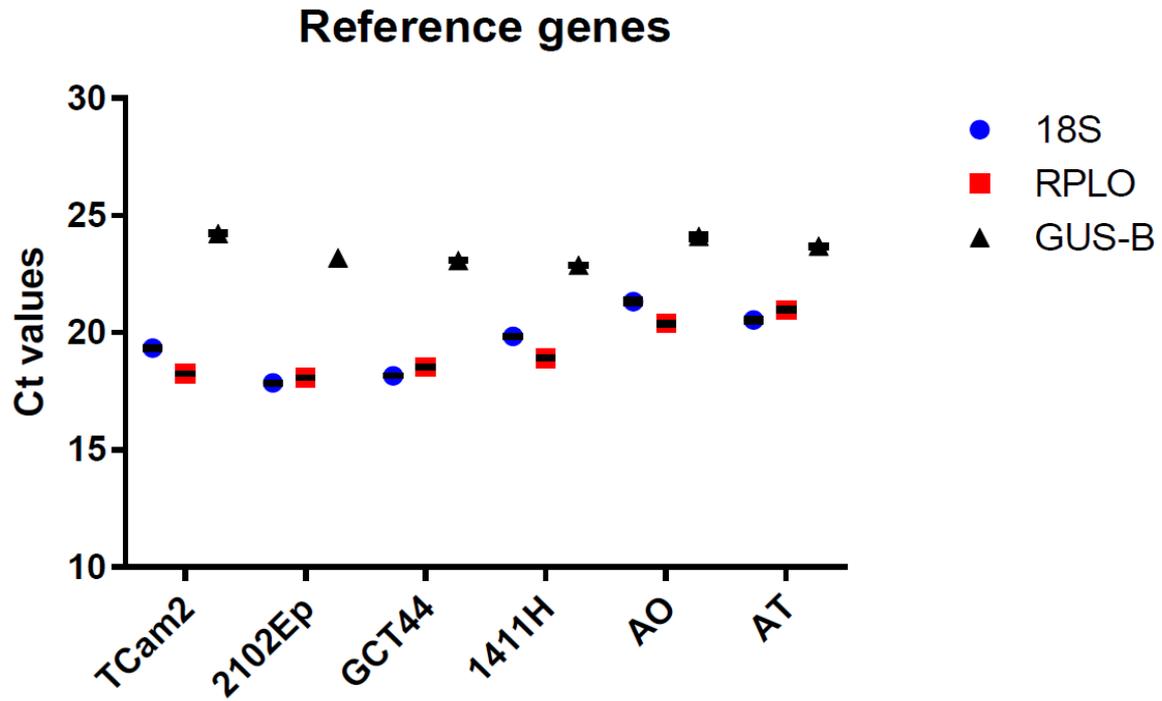
Figure 3.10: **Raw Ct values of three candidate housekeeping genes for pri-miRNA analysis in four representative GCT cell lines and two controls**. AO is Ambion Ovary RNA and AT is Ambion Testis RNA. Data presented as mean ± SEM of n=3 biological replicates.

3.2.4.2 **Pri-miRNA expression in mGCTs cell lines**

The level of expression of the pri-miRNAs from each individual chromosome, was normalised to the pooled control mRNAs (i.e. testis and ovary gonadal controls) using the three housekeeping genes identified and plotted for each individual cell line. While the expression of the miRNA was downregulated in a similar manner irrespective of the genomic location in the cell line TCam2 (a seminoma cell line), the other cell lines showed a differential downregulation from the two chromosomal loci. In 2102Ep (an EC derived cell line) there was greater downregulation of the chromosome 21 derived pri-miR-99a and pri-miR-125b-2 transcripts, while in 1411H and GCT44 (both cell lines derived from YSTs) there was a greater downregulation of the chromosome 11 derived pri-miR-100 and pri-miR-125b-1 (Figure 3.11).

This suggests that transcription of pri-miRNA from both chromosomes is consistently affected in mGCT cell lines, although in some cases there appears to be a preferential choice of which transcript is targeted for downregulation. Whether this preference is associated with the subtype of tumour is unclear, more cell lines and tissue samples would need to be studied for firmer conclusion, but it is interesting to note that both YST derived cell lines studied here showed the same pattern and that this pattern was different from that seen in other GCT subtypes.
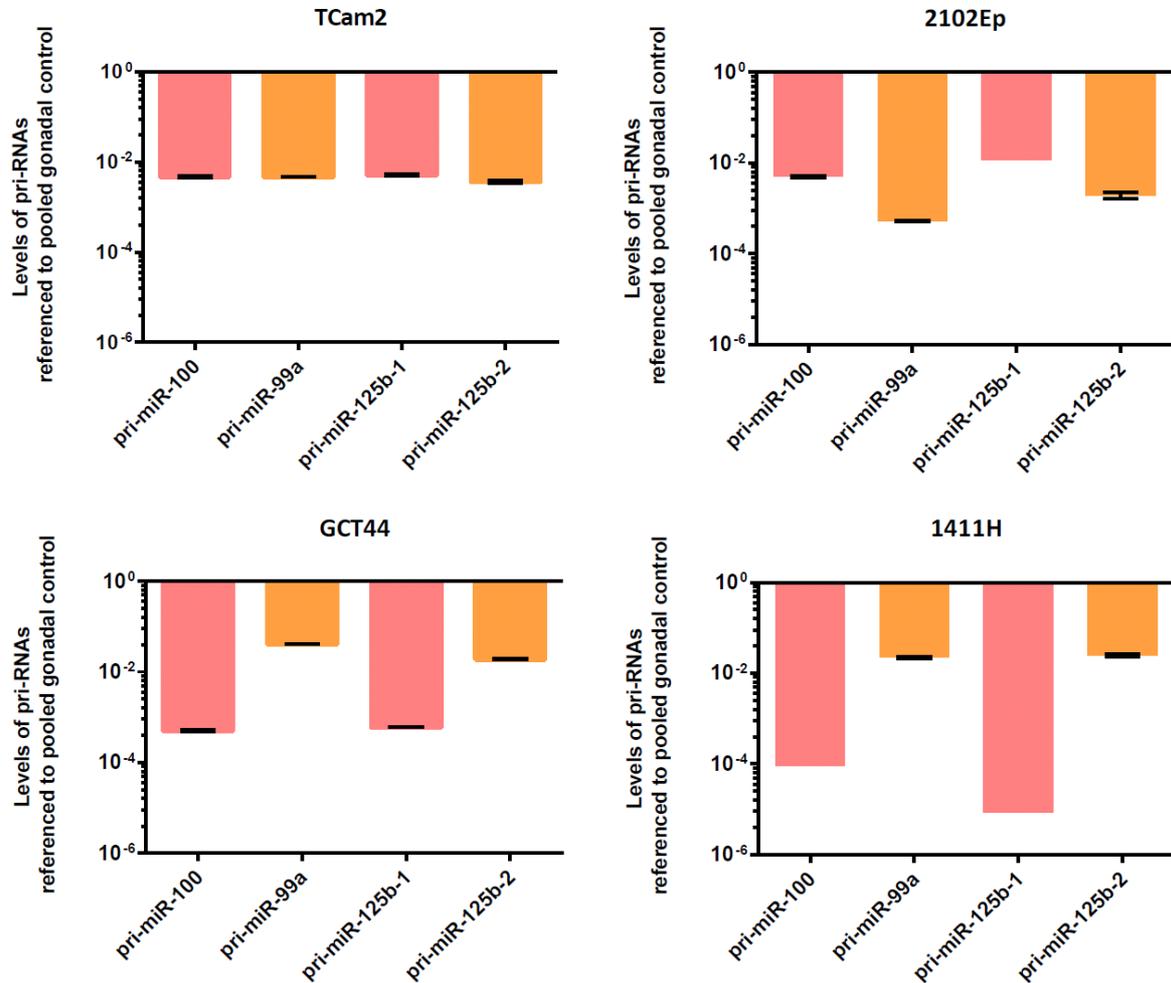
Figure 3.11: **Expression of pri-miRNAs in different GCT cell lines**. Pink bars represent contributions from chromosome 11, where miR-100 and miR-125b-1 are located, while orange bars represent chromosome 21 contributions, where miR-99a and miR-125b-2 are located. Data presented as mean ± SEM of n=3 biological replicates.

3.2.4.3 Correlation of the expression of pri-miRNA on chromosome 11 and 21

To assess if expression of the pri-miRNA showed intra-chromosomal correlation (i.e to establish if pri-miR-125b-1 transcription correlated to pri-miR-100 transcription and pri-miR-125b-2 transcription correlated to pri-miR-99a transcription) expression levels were compared in four different mGCTs cell line and two normal gonadal controls. Significant correlation was found (*p*-values: 0.0015 and 0.0006, respectively) suggesting that both transcripts on the same chromosome are similarly expressed (Figure 3.12).
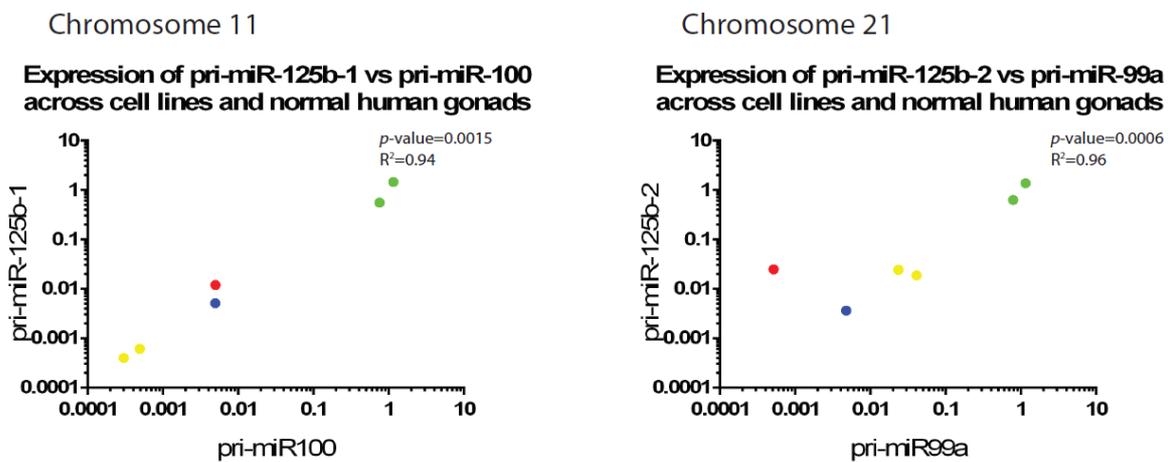
Figure 3.12: **Correlation of expression of pri-miRNAs located on the same chromosome.** Four different mGCTs cell line were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour) and in green two normal gonadal control (ovaries and testis).

Correlation between expression from the two chromosomes was also assessed (Figure 13). High level of correlation was found (p-values: <0.0001), suggesting a co-regulation in the mechanism of downregulation.
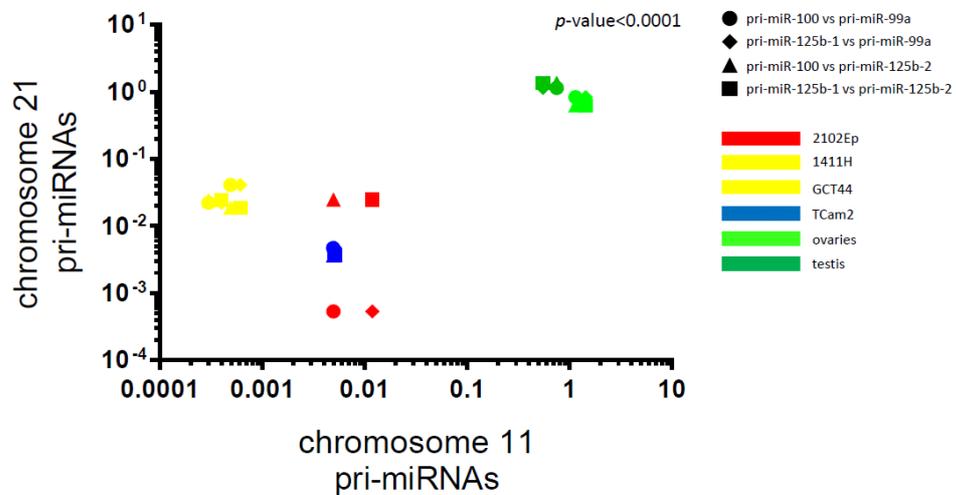
Figure 3.13: **Correlation of expression of pri-miRNAs located on the different chromosome**. Four different mGCTs cell line were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour) and in green two normal gonadal control (ovaries and testis). The different symbols represent different correlations. The circle symbolises the correlation between pri-miR-100 and pri-miR-99a, the rhombus symbolises the correlation between pri-miR-125b-1 and pri-miR-99a, the triangle symbolises the correlation between pri-miR-100 and pri-miR-125b-2 and the square symbolises the correlation between pri-miR-125b-1 and pri-miR-125b-2.

3.2.5 Assessment of miRNA -3p abundance

Pri-miRNAs are further processed to a miRNA:miRNA* duplex, an immature miRNA which is found in the cytoplasm. This immature miRNA is formed by two strands, the guide strand, which will become the mature miRNA and as such is loaded into the miRNA-induced silencing complex, and the passenger strand or miRNA*. Whether the -5p strand or the -3p strand becomes the mature miRNA is generally depended on which of the strands is more stable. In the case of miR-99, miR-100, and miR-125b the mature miRNAs are the -5' strand. To further investigate the relative contributions to overall miRNA abundance from the two loci on chromosomes 11 and 21, the -3p strands of the four miRNAs were quantified (Table 3.3). This is due to the fact that the -3p strands of each miRNA are unique because their sequence depends on whether the miRNA is derived from Chromosome 11 or 21. Indirectly this makes possible to further study the relevant contribution from the two different loci.

miRNA name	Sequence
hsa-miR-99a-3p	CAAGCUC̲G̲C̲U̲UCUAU̲G̲GGU̲C̲UG
hsa-miR-100-3p	CAAGCUUGUAUCUAUAGGUAUG
hsa-miR-125b-1-3p	A̲C̲G̲GGUUAGGCUCU̲U̲GGG̲A̲GCU
hsa-miR-125b-2-3p	UCACAAGUCAGGCUCUUGGGAC

Table 3.3: **MiRNAs-3p sequences for the miRNA of interest**. Nucleotide differences in miR-99a-3p and miR-125b-1-3p using miR-100-3p and miR-125b-2-3p respectively as the canonical sequence are underlined.

RNU24, was used as HK gene to normalise expression level of -3p strand of the miRNAs of interest. -3p strands are present in the cells in low abundance, making their detection difficult (Anglesio *et al.*, 2013). Optimisation were undertaken as later described in the paragraph 3.2.5.1. Because of this reason, when RT-qPCR was performed, the input of RNA was higher for mGCTs and universal reference (UR, UR derived from a mix of 14 commonly used cell lines and one of the used cell lines in a mGCTs, for this reason in this experiment the amount of UR RNA used was the same as the one used for mGCTs), than non-malignant control and as a result, RNU24 showed a high standard deviation (σ=1.3) when compared among all of the sample. When standard deviation was assessed only among mGCTs cell lines and UR or among non-malignant control (normal ovaries and testis) the standard deviation become acceptable (σ=0.4) (Figure 3.14).
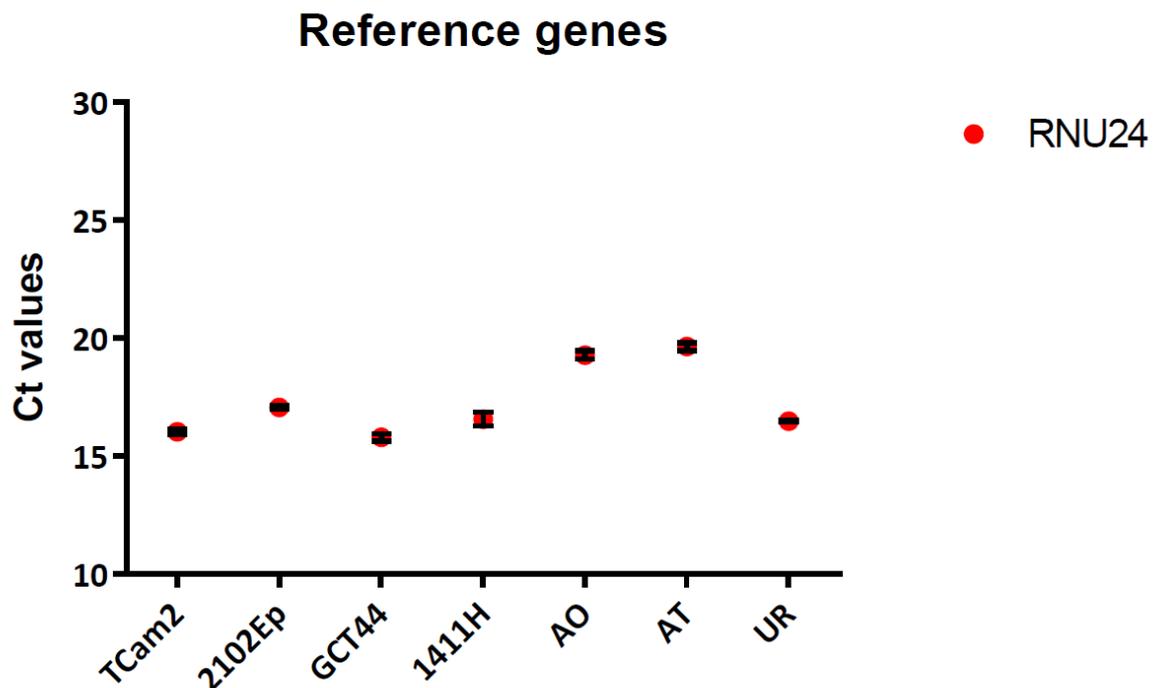
Figure 3.14: **Raw Ct values of the housekeeping gene RNU24 in four representative GCT cell lines**. AO is Ambion Ovary RNA, AT is Ambion Testis RNA and UR is Universal Reference. Data presented as mean ± SEM of n=3 biological replicates.

3.2.5.1 **Expression of miRNA-3p in Germ Cell Tumour cell lines**

As for the pri-miRNA quantitation, some optimisation had to be carried out to be able to quantitate the -3p miRNAs. Consistent with the fact that -3p strand miRNAs are generally found at much lower abundance in cells than their -5p counterparts (Anglesio *et al*., 2013), the input RNA had to be increased from the standard 15 ng to 100 ng, to detect reliably in the GCT cell lines the -3p strands using Taqman miRNA assays (Figure 3.15).
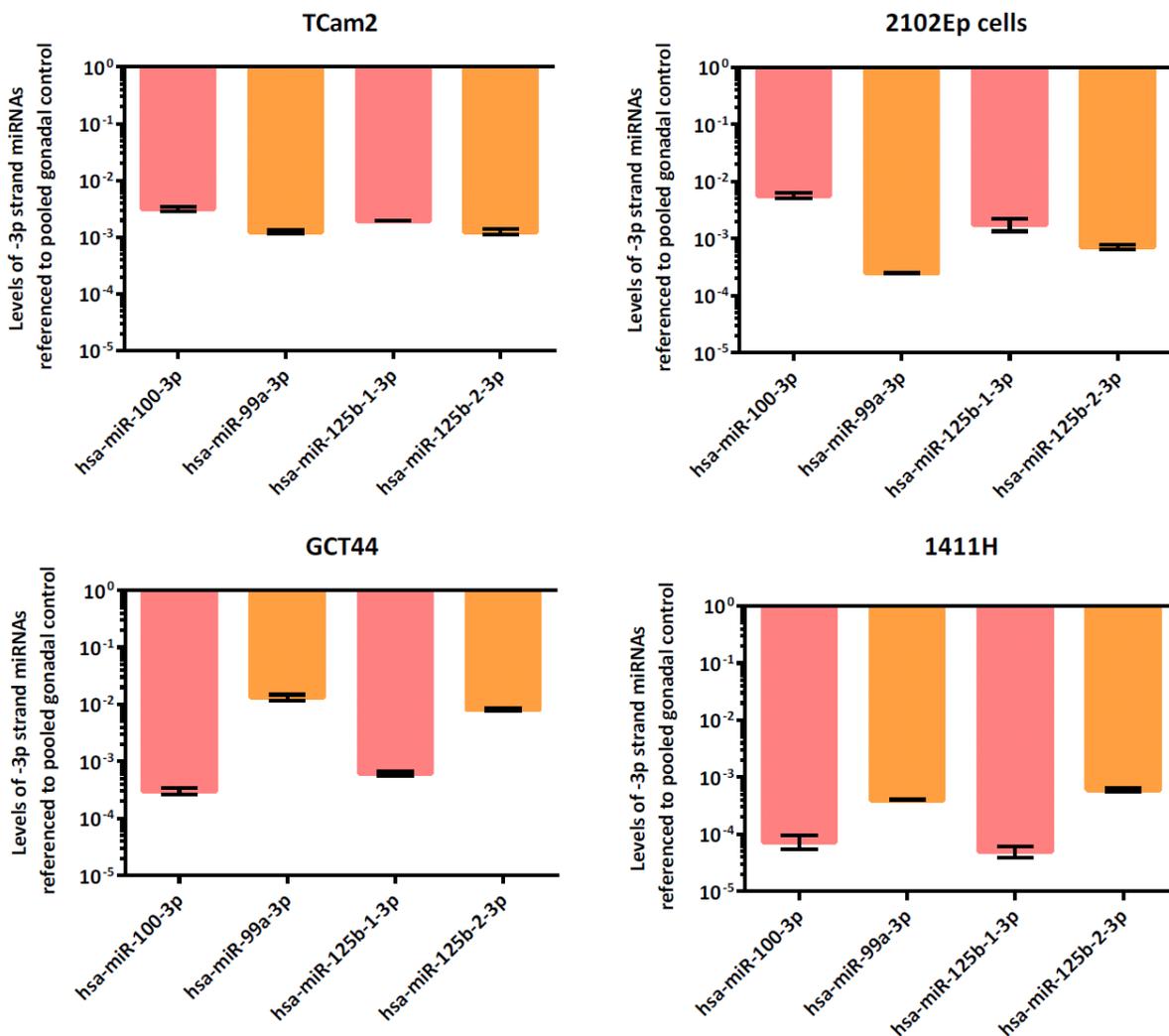
Figure 3.15: **Expression of -3p strand miRNAs in different GCT cell lines.** Pink bars represent contributions from chromosome 11, where miR-100 and miR-125b-1 are located, while orange bars represent chromosome 21 contributions, where miR-99a and miR-125b-2 are located. Data presented as mean ± SEM of n=3 biological replicates.

The level of expression of the -3p miRNAs from each individual chromosome, was normalised to the pooled control mRNAs (i.e. testis and ovary gonadal controls) using RNU24 as housekeeping gene and plotted for each individual cell line. Universal reference control (UR) was excluded in this analysis. UR is a pool of 14 cancer cell lines and one of these is mGCT cell line, hence it cannot be used as a negative control.

Consistent with the differential expression of pri-miRNA described in 3.2.4.2, the seminoma cell line TCam2 showed a similar level of downregulation of the miRNAs located on chromosomes 11 and 21, whereas the other cell lines showed differential regulation. In 2102Ep (the EC derived cell line) there was greater downregulation of the chromosome 21 derived miR-99a-3p and miR-125b-2-3p transcripts, while 1411H and GCT44 cells (both cell lines derived from YSTs) showed a greater downregulation of the chromosome 11 derived miR-100-3p and 125b-1-3p (Figure 3.15).

These results are consistent with the observation made during the analysis of pri-miRNA expression that transcription of -3p miRNAs from both chromosomes is consistently affected in mGCT cell lines, and that in some cases there appears to be a preferential choice of which transcript is targeted for downregulation. These observations support the idea that this preference is associated with the subtype of tumour, as observed with pri-miRNA expression described earlier, but more cell lines and tissue samples would need to be studied.

3.2.5.2 Correlation of the expression of miRNA-3p on chromosome 11 and 21

Post-transcriptional mechanisms inhibiting the maturation of miRNA have been shown (Gulyaeva and Kushlinskiy, 2016), therefore correlation between the expression of -3p strands derived miRNAs from each chromosome was determined. Significant correlation in the expression of -3p transcripts from within the same chromosome, as assessed in four different mGCTs cell line, two normal gonadal controls, and the Universal RNA Reference (Ambion), was found (p-values: 0.0032 and 0.0028, respectively) (Figure 3.16).
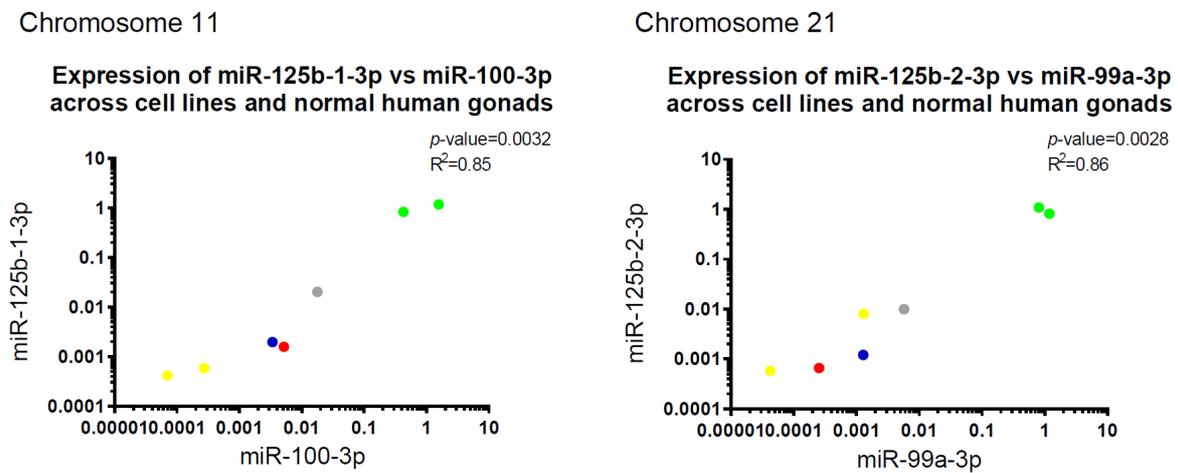
Figure 3.16: **Correlation of expression of -3p miRNAs located on the same chromosome.** Four different mGCTs cell line were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour). In green two normal gonadal control (ovaries and testis) and in grey the Universal RNA Reference.

As for the pri-miRNA experiments, correlation between chromosomes was also assessed. Again, significant correlation was found (p-values: <0.0001), in agreement with the potential co-regulation suggested for the pri-miRNAs (Figure 3.17).
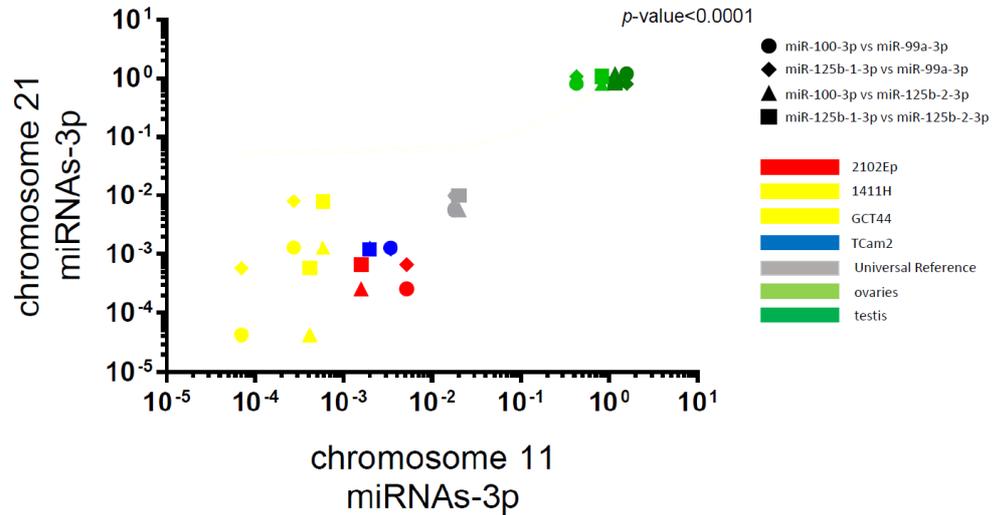
Figure 3.17: **Correlation of expression of -3p miRNAs located on the different chromosome**. Four different mGCTs cell line were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour). In green two normal gonadal control (ovaries and testis) and in grey the Universal RNA Reference. The different symbols represent different correlations. The circle symbolises the correlation between miR-100-3p and miR-99a-3p, the rhombus symbolises the correlation between miR-125b-1-3p and miR-99a-3p, the triangle symbolises the correlation between miR-100-3p and miR-125b-2-3p and the square symbolises the correlation between miR-125b-1-3p and miR-125b-2-3p.

3.2.6 **Assessment of miRNAs host genes abundance**

MiR-125b-1 and miR-100 are encoded in the intron of a large polycistronic transcript derived from Chromosome 11, the long non-coding RNA (lincRNA) MIR100HG. Similarly, miR-125b-2 and miR-99a are encoded in the intron of another large polycistronic transcript derived from Chromosome 21, MIR99AHG (LINC00478). In order to better understand the regulation of transcription across the whole region where the two families of miRNA are located, the expression of the lincRNAs MIR100HG and MIR99AHG were also determined. Moreover, the expression of a coding gene, BLID (BH3-like motif-containing cell death inducer), a positive regulator of caspase activity which is located within the region on Chromosome 11 was also studied.

3.2.6.1 **MiRNAs host genes detection**

Four housekeeping genes, RPL13, HMBS, ACTβ and YWHAZ were tested to determine the most consistent ones to use for normalisation of the expression data. HMBS transcription had a high standard deviation between samples (σ=1.46), and the expression ratio between HMBS and the others HK genes varied across all the samples. This HK gene was discarded for subsequent data normalisation. The other HK genes had a lower variability (σ=0.85, 0.88 and 0.85, respectively), and were chosen as they all had the same expression ratio between HKs across the samples (Figure 3.18).
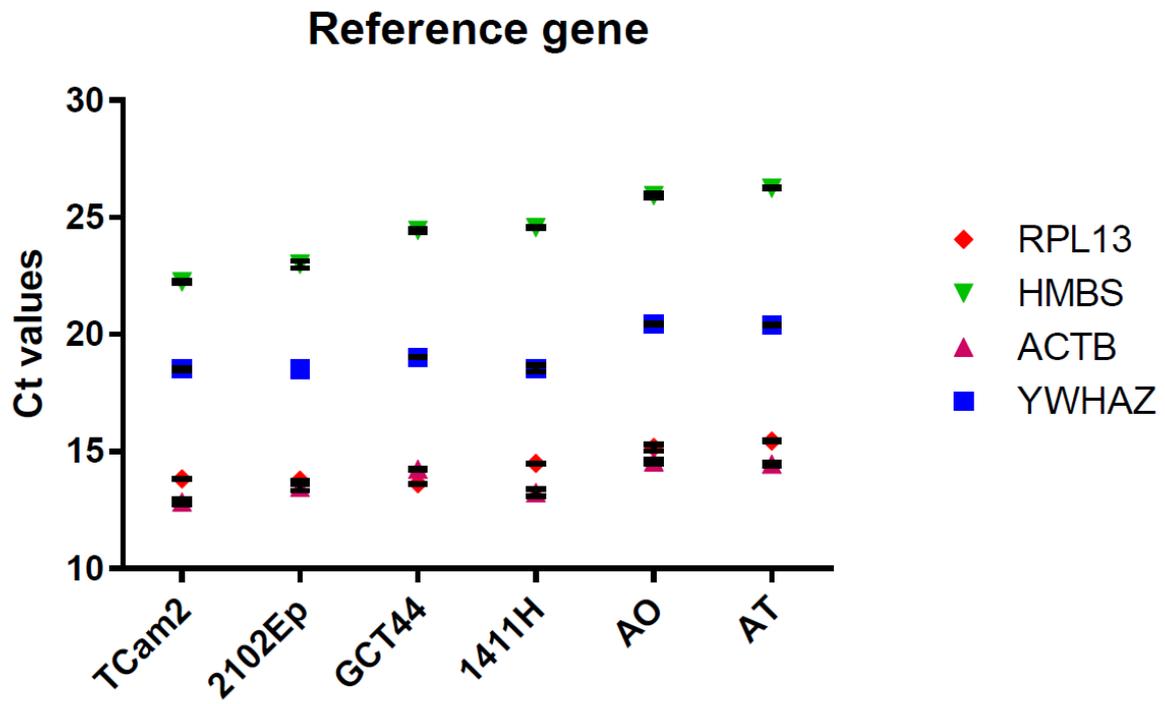
Figure 3.18: **Raw Ct values of four housekeeping genes in four representative GCT cell lines**. Data presented as mean ± SEM of n=3 biological replicates.

The level of expression of lincRNA MIR100HG, and BLID, and lincRNA MIR99AHG from each individual chromosome, was normalised to the pooled control mRNAs (i.e. testis and ovary gonadal controls) using the three housekeeping genes to normalise expression of both coding and miRNAs host transcripts and plotted for each individual cell line (Figure 3.19).
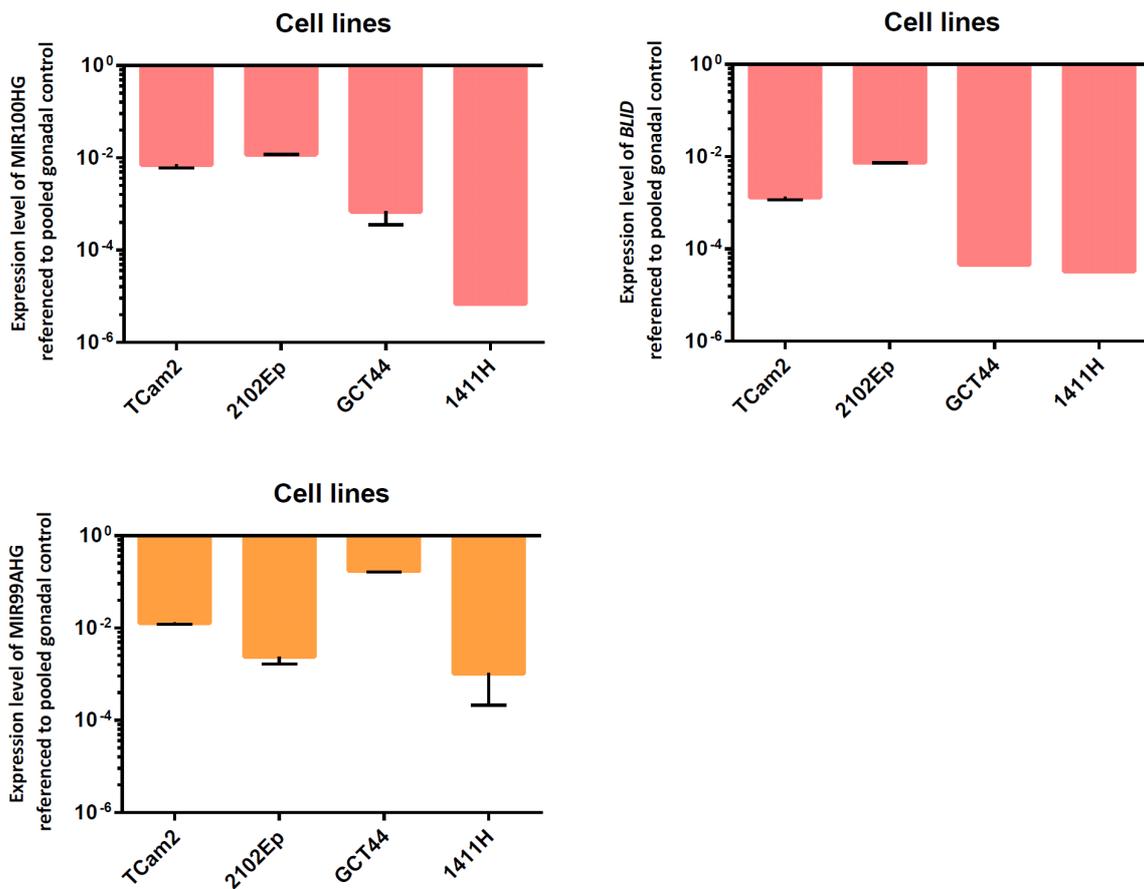
Figure 3.19: **Expression of lincRNAs and BLID gene in different GCT cell lines.** Pink bars represent contributions from chromosome 11, where miR-100 and miR-125b-1 are located, while orange bars represent chromosome 21 contributions, where miR-99a and miR-125b-2 are located. Data presented as mean ± SEM of n=3 biological replicates.

Consistent with previous results both chromosomal sites showed downregulation of lincRNA transcripts. Moreover, the data is consistent with the differential expression of pri-miRNA and -3p miRNA described in 3.2.4.2, and 3.2.5.1 respectively, the seminoma cell line TCam2 showed a similar level of downregulation of the lincRNAs located on chromosomes 11 and 21,

whereas the other cell lines showed differential regulation. In 2102Ep (the EC derived cell line) there was greater downregulation of the chromosome 21 lincRNA MIR99AHG transcripts, while 1411H and GCT44 cells (both cell lines derived from YSTs) showed a greater downregulation of the chromosome 11 derived lincRNA MIR100HG. Furthermore, the transcripts for the coding gene BLID were downregulated similarly to the lincRNA MIR100HG, which is transcribed independently from an overlapping region, with the exception of the GCT44 cell line where this transcript appear to be slightly more downregulated than the corresponding lincRNA. These data suggest that similar mechanisms of transcriptional regulation may occur across the whole area on both chromosomes.

3.2.7 Correlation of expression of pri-miRNA, miRNA-3p/-5p and MIR100HG host gene in chromosome 11

The data shown is consistent with similar regulation occurring on lincRNAs, pri-miRNAs and individual miRNAs, both -3p, and -5p as well as on the coding gene BLID, which is located in the same area but transcribed independently. To further investigate this transcriptional regulation, correlation between expression of MIR100HG and pri-miRNA, miRNA-3p, -5p and BLID gene, were analysed. The -5p data were previously presented in the Figure 3.6.

MIR100HG expression significantly correlated with expression of miR-100 and miR-125b-1 throughout the different stages of miRNA biogenesis studied. Expression of both primary transcripts pri-miR-100 and pri-miR-125b-1 showed significant correlation to MIR100HG expression (p-value: 0.0005 and <0.0001, respectively) (Figure 3.20 panel a). A similar significant correlation was seen for MIR100HG and miRNA -3p strands expression (p-value: 0.0009 and 0.0005, respectively) and MIR100HG and miRNA -5p strands (p-value: 0.0002 and <0.0001, respectively), (Figure 3.20 panels b and c respectively). Interestingly, the correlation between MIR100HG and BLID gene expression also appeared to be significant (p-value: <0.0001) (Figure 3.20 panel d).

a)

Expression of pri-miR-100 vs MIR100HG across cell lines and normal human gonads

p-value=0.0005
R²=0.96
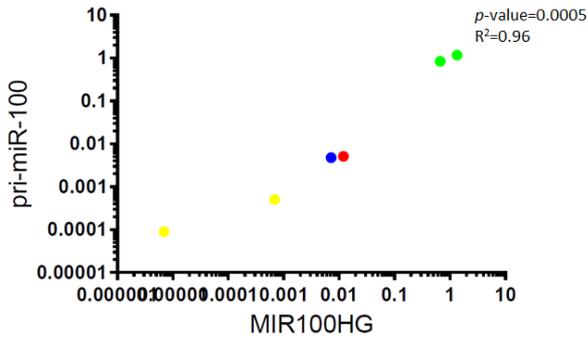
Expression of pri-miR-125b-1 vs MIR100HG across cell lines and normal human gonads

p-value<0.0001
R²=0.98
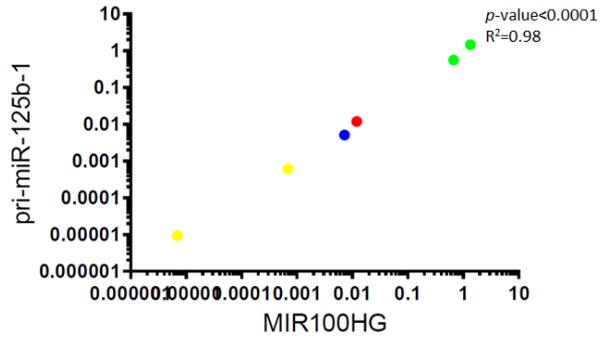
b)

Expression of miR-100-3p vs MIR100HG across cell lines and normal human gonads

p-value=0.0009
R²=0.95
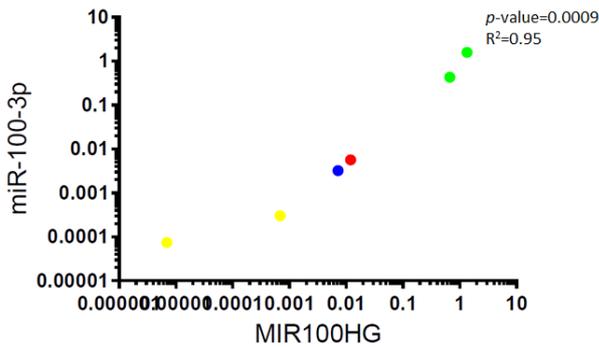
Expression of miR-125b-1-3p vs MIR100HG across cell lines and normal human gonads

p-value=0.0005
R²=0.96
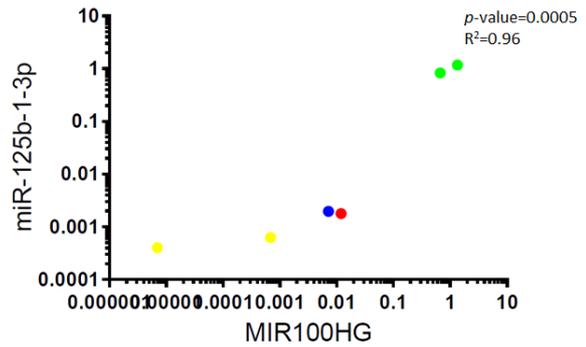
c)

Expression of miR-100-5p vs MIR100HG across cell lines and normal human gonads

p-value=0.0002
R²=0.98
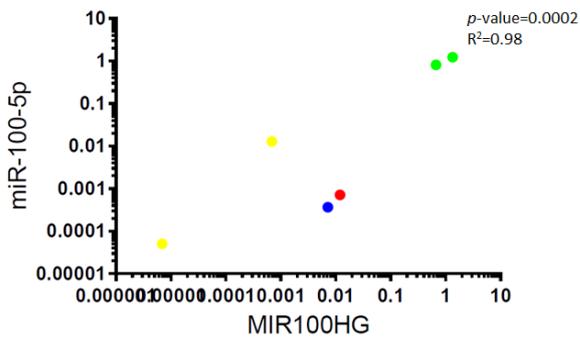
Expression of miR-125b-5p vs MIR100HG across cell lines and normal human gonads

p-value=<0.0001
R²=0.99

d)
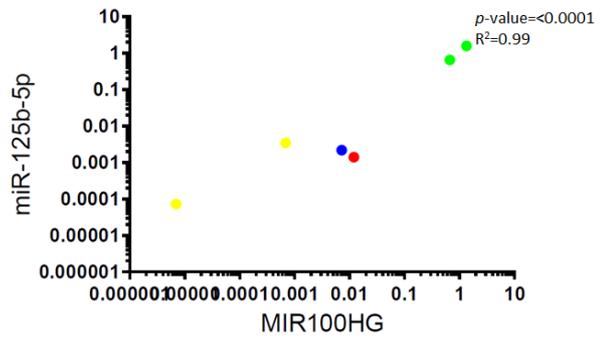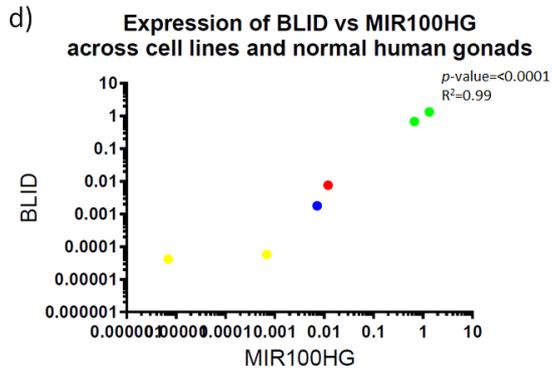
Figure 3.20: **Correlation of expression of various transcripts within chromosome 11.** Four different mGCTs cell lines were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour) and in green two normal gonadal control (ovaries and testis).

3.2.8 Correlation of expression of pri-miRNA, miRNA-3p/-5p with MIR099AHG host gene on chromosome 21

As for the chromosome 11 locus, the data shown for chromosome 21 loci is consistent with similar regulation occurring on lincRNAs, pri-miRNAs and individual miRNAs. To further investigate the transcriptional regulation on chromosome 21, correlation between expression of MIR99AHG and primary transcripts, miRNA-3p, -5p encoded at this locus were analysed. The -5p data were previously presented in the Figure 3.6.

MIR99AHG expression significantly correlated with expression of miR-99a and miR-125b-2 throughout the different stages of miRNA biogenesis studied. Expression of both primary transcripts pri-miR-99a and pri-miR-125b-2 showed significant correlation to MIR99AHG expression (*p*-value: 0.0003 and 0.0002, respectively) (Figure 3.21 panel a). A similar significant correlation was seen for MIR99AHG and miRNA -3p strands expression (p-value: 0.037 and 0.0005, respectively), however MIR99AHG and miRNA -5p strands did not correlate significantly (p-value: 0.07 and 0.137, respectively), probably due to a slightly higher level of MIR99AHG transcription in GCT44 cell line (Figure 3.21 panels b and c respectively).
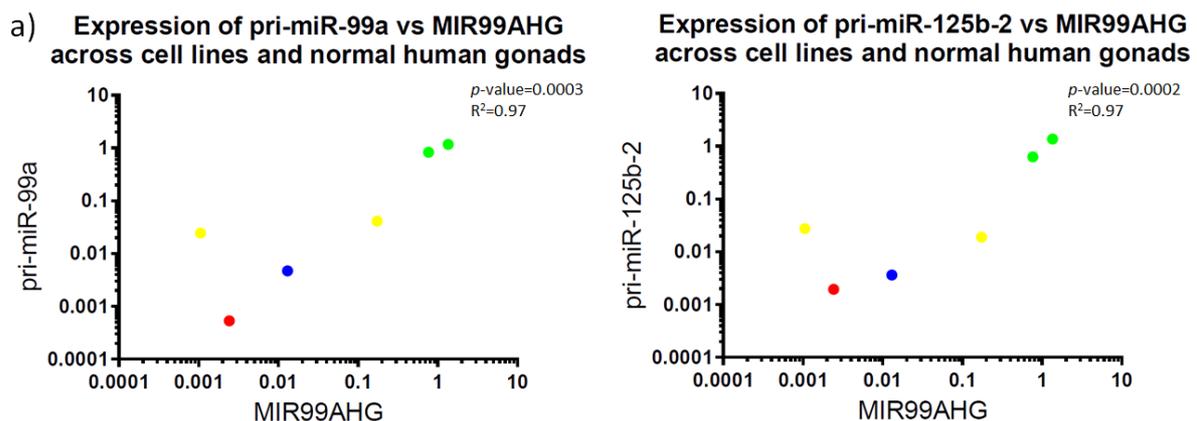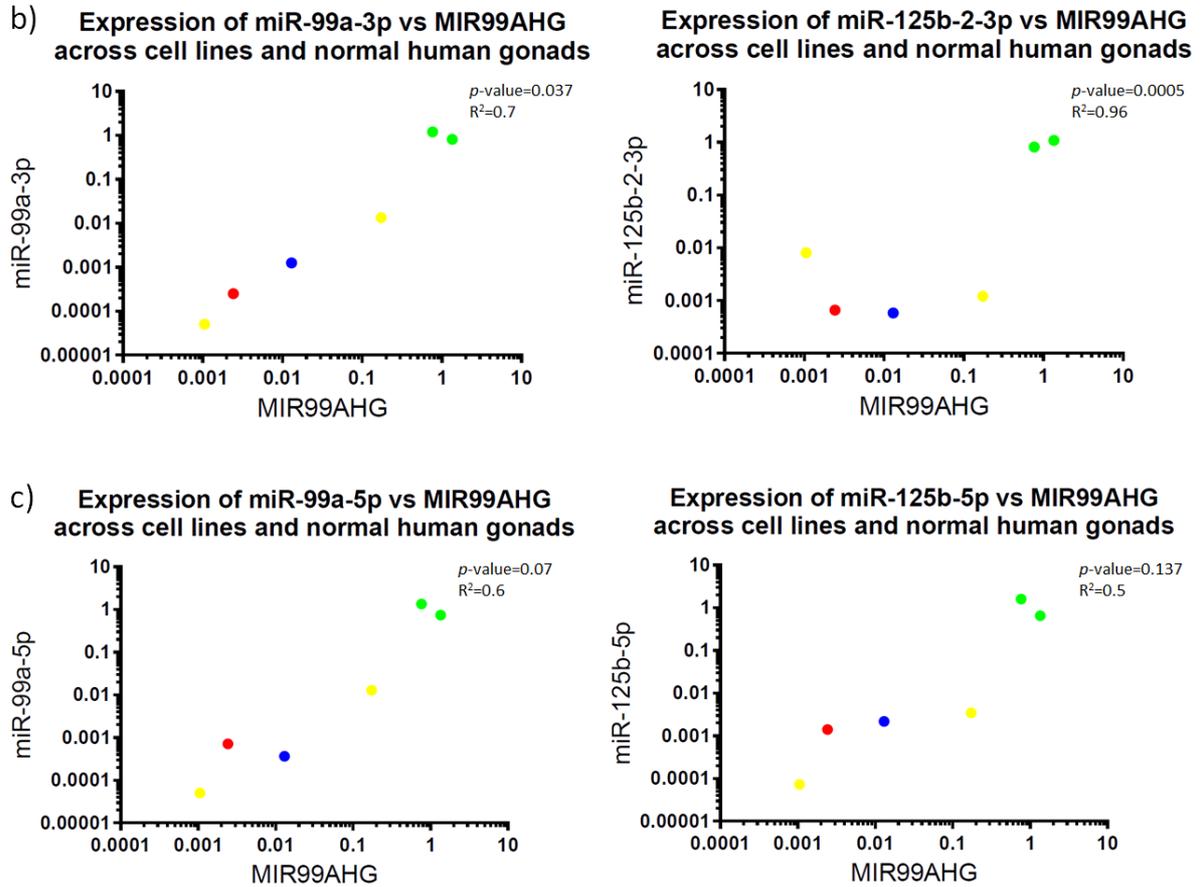
Figure 3.21: **Correlation of expression of various transcripts within chromosome 21**. Four different mGCTs cell line were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour) and in green two normal gonadal control (ovaries and testis).

3.2.9 Correlation between lincRNAs transcribed from two chromosomal loci

Having better clarified the regulation of transcription on each individual chromosome, the interplay between the two loci in different chromosomes was then explored.

The correlation between transcription of genes from the two chromosomes resulted to be significant, with a *p*-value of 0.005 (Figure 3.22).
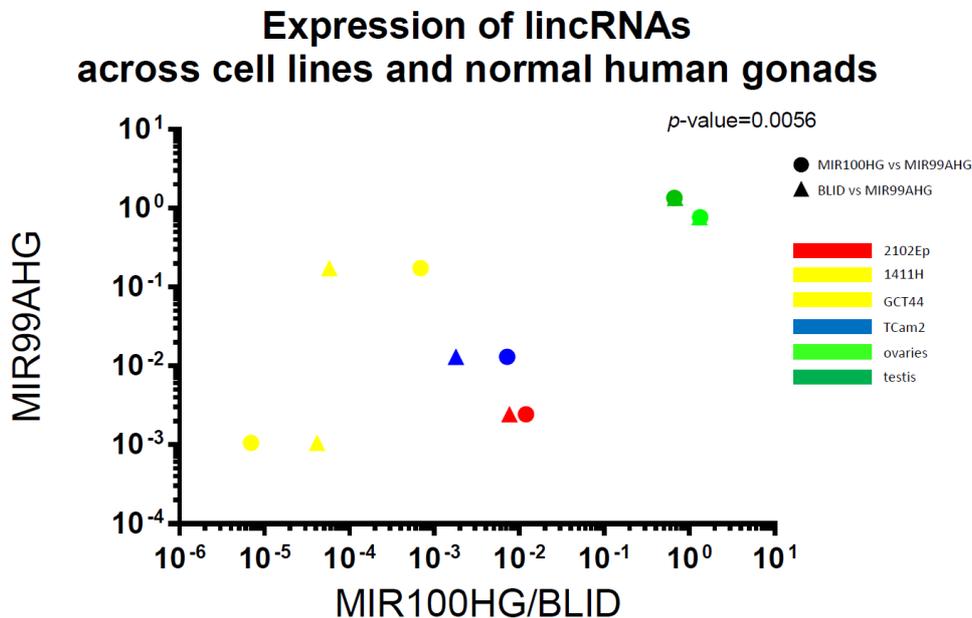
Figure 3.22: **Correlation of transcription across the regions encoding for the downregulated miRNAs on chromosome 11 or chromosome 21.** Four different mGCTs cell lines were used: in red 2102Ep (EC), in blue TCam2 (Seminoma), in yellow 1411H and GCT44 (both Yolk Sac Tumour) and in green two normal gonadal control (ovaries and testis). The different symbols represent different correlations. The circle symbolises the correlation between MIR100HG and MIR99AHG while the triangle symbolises the correlation between BLID and MIR99AHG.

All together these data point to an intra-chromosomal and inter-chromosomal co-regulation of the expression of these miRNA, the origins of which we investigated in further work presented in chapter 5.

3.3 Discussion

Microarray analysis of the expression of 246 miRNAs from RNA extracted from 42 paediatric mGCTs samples and compared with normal samples, demonstrated that the most downregulated miRNAs (adjusted p-values$<1x10^5$) were: miR-125b-5p, miR-100-5p, miR-99a-5p, miR-204-5p and miR-152-3p. Further confirmation of the importance of these miRNAs in GCT malignancy was shown by the analysis of the overall seed abundance in the 3'UTR of deregulated mRNA from mGCT samples and cell lines compared with non-malignant control samples (normal gonadal controls and benign teratomas) (Palmer *et al.*, 2010).

These data demonstrated that the 2-7 nt seed sequences of miR-99a-5p/-100-5p and miR-125b-5p are substantially decreased in both clinical mGCT samples and representative mGCT cell lines, compared with non-malignant control samples. These data also suggest that the observed downregulation of miR-99a-5p/-100-5p and miR-125b-5p is likely to have a significant functional effect in mGCTs.

Analysis of the microarray data showed downregulation of transcription of miR-204-5p and miR-152-3p in both mGCTs and benign GCTs tissue samples. This result is consistent with the reduction in seed abundance shown for miR-204-5p, which seems to be the same across all the samples and cell lines. The abundance of the seed of miR-152-3p is more significantly downregulated across the mGCT clinical samples compared with the normal control and the benign form of GCTs, but this downregulation is not mirrored in the cell lines we studied. Although it is probable that the downregulation of these miRNAs has a role in mGCTs, the cell lines we have available to study these malignancies in vitro would not be the best model to use. Consequently, further work was focussed on miRNAs miR-99a-5p, miR-100-5p and miR-125b-5p.

Additionally, a similar pattern of downregulation for miR-99a-5p/-100-5p and miR-125b-5p, was observed in the original microarray data, which may suggest that these miRNAs are potentially coregulated.

The mature form of miR-125b-5p can be derived from either miR-125b-1, which is transcribed from a locus on chromosome 11 or miR-125b-2, which is transcribed from a locus on chromosome 21. Interestingly, miR-100-5p cluster together with miR-125b-1 on chromosome 11 and miR-99a-5p cluster together with miR-125b-2 on chromosome 21 and these two miRNAs (miR-99a-5p and miR-100-5p) have a high level of homology, with an identical seed region shared by the -5p mature strands and differing only by one nucleotide towards the 3' end of the miRNAs. Cross-reactivity experiment also revealed that miR-99a-5p and miR-100-5p are indistinguishable using the most selective available qPCR detection system (Taqman miRNA assays), and it is therefore impossible to separate these transcripts from each other.

Previous studies investigating the role of the individual miRNA of the miR-100/-125b-1 and miR-99a/-125b-2 clusters, demonstrated conflicting functions. For example, miR-125b-5p overexpression has been linked to poor prognosis in HER2-positive patients with either gastric or breast cancer (Luo *et al*., 2017; Sui *et al*., 2017), but conversely in breast cancer, miR-125b-5p was reported down-regulated in biopsy specimens and as tumour suppressors (Iorio *et al*., 2005; Mattie *et al*., 2006) by mediating the ERBB2 and ERBB3 pathway (Scott *et al*., 2007) or by targeting the ETS1 gene (Y. Zhang *et al*., 2011b). Similarly, miR-99a-5p/-100-5p has been associated with a tumour suppressing or a tumour promoter role in cancer. For example, miR-100-5p overexpression was correlated with progression of renal cell carcinoma, acute myeloid leukaemia and epithelium-derived gastric cancer cell (Gu *et al*., 2015; Qin, Huang and Wang, 2015; Yang *et al*., 2015), but contrarily in gastric cancer, ovarian cancer and breast cancer miR-100 was acting as tumour suppressor, by targeting CXCR7 or by inhibiting Wnt/β-catenin signalling pathway (Jiang *et al*., 2016; Cao *et al*., 2018). However, there is a high phylogenetic conservation of the miR-99a/miR-100/miR-125b cluster, and it is thought that within the cluster there may be co-regulation in either the tumour suppressive or oncogenic role.

The results obtained with the microarray analysis of expression in mGCTs, demonstrated an almost identical pattern of downregulation of miR-99a-5p/-100-5p and miR-125b-5p. Furthermore, the genomic redundancy of the clusters suggests that the expression of the individual miRNA may be contributed to by one chromosome alone (11 or 21) or by the

simultaneous expression from both chromosomes (11 and 21), which are then potentially regulated by the same mechanisms.

MiRNAs dysregulation occurs in several ways, but the detailed mechanisms causing this inaccurate expression are still largely uncharacterised. Chromosomal alteration and genomic abnormalities have been shown to cause miRNAs dysregulation in ovarian cancer (37.1% of the overall cases), breast cancer (72.8%) and melanoma (85.9%) (Zhang *et al*., 2006b). In light of this evidences, potential copy number changes at the miRNAs miR-100 and miR-125b-1 cluster locus on chromosome 11 (chr11 q24.1), and at the miR-99a and miR-125b-2 locus on chromosome 21, (chr21 q21.1) were investigated by qPCR analysis on genomic DNA.

In mammals, promoter regions, enhancers, repressors or single nucleotide polymorphisms (SNPs) many kilobases (Kb) away from the transcription start site can affect the subsequent transcription of genes (Dobi and Winston, 2007), therefore primers for the analysis of potential genomic imbalances were designed up to 50Kb upstream of both loci.

Frequent chromosomal aberration have been previously reported in mGCTs, in particular copy number gains of chromosomes 1, 7, 8, 17, 21, X and copy number loss of chromosomes 4, 5, 11, 13, 18 and Y (Kraggerud *et al*., 2002; McIntyre *et al*., 2007). Consistent with these findings, our data showed copy number loss of chromosome 11 in some of the mGCTs cell lines (particularly in 1411H and TCam2) compared with normal tissue from testis, and copy number gains for chromosome 21 in mGCTs cell lines (in particular 1411H, GCT44 and 2102Ep) when compared with normal tissue from testis. However, no significant further change in chromosomal copy number was detected in the region of chromosome 11 up to 50Kb upstream of miR-100 and in the region of chromosome 21 up to 50Kb upstream of miR-99a in mGCTs cell lines, suggesting this is not the mechanism by which miRNA dysregulation is occurring.

Therefore, expression of the genes throughout the biogenesis of the miR-99a-5p/-100-5p and miR-125b-5p miRNA was further investigated with particular focus in identifying the relative contributions to transcription of the chromosome 11 and 21 loci.

The first step in miRNA biogenesis is the production of a primary transcript (pri-miRNA) from introns or exons of non-coding or coding protein genes. This process takes place in the nucleus

and is highly controlled and complex. Pri-miRNAs are difficult to quantify as they can be easily degraded, unless they have been modified by capping (polyadenylated cap) and are protected by the proteins complexes of the microprocessor (where the core components DROSHA and DGCR8 cleave the pri-miRNA into the precursor miRNA (pre-miRNA) (Cai, Hagedorn and Cullen, 2004). It is also crucial that DNA is removed prior to quantification, as the pri-miRNAs specific primers can also target gDNA sequences.

Specific primers were used to detect the appropriate pri-miRNA and RNA input optimised to allow detection. Pri-miR-99a/-miR-100/-miR-125b-1/-miR-125b-2 expression was quantified in mGCT cell lines and compared with gonadal control. Overall expression of pri-miRNAs was downregulated for all mGCTs cell lines. As downregulation of pri-miRNA can be due to several mechanisms, including downregulation of protein involved in their stabilization or maturation, transcription factor or epigenetic mechanisms which regulates their expression (Ha and Kim, 2014), we used data from mRNA microarrays that were performed at the same time of the miRNA microarrays discussed earlier in the chapter using the same tissue samples, cell lines and controls, to identify potential changes that may suggest potential mechanisms.

Comparison of the expression of genes involved in the maturation and processing of pri-miRNA such as PRL1, important for pri-miRNA stabilization, DROSHA or DGCR8, the main regulators of pri-miRNA maturation in mGCT samples and control tissues suggested that the lower expression of pri-miRNAs was unlikely to be due to deregulated expression of members of the microprocessor complex as no differences in the expression of the transcripts encoding the proteins involved in these processes were detected (data not shown). It cannot be excluded however, that post-transcriptional regulation may have an effect on the composition of these complexes.

Interestingly, although all mGCT cell lines demonstrated downregulation of pri-miRNAs derived from both chromosomes 11 or 21, there appears to be differential contribution from each chromosome in different cell lines. For instance, in 1411H and GCT44, both yolk sac tumours derived cell lines, transcription of pri-miR-100 and pri-miR-125b-1 was more reduced than transcription of pri-miR-99a and pri-miR-125b-2, perhaps reflecting the detection of a

lower copy number across chromosome 11 in these cells. On the other hand, the TCam2 cell line showed similar level of downregulation on both chromosomes, whereas in 2102Ep expression from chromosome 21 appeared less abundant than from chromosome 11. These results suggest that there is no individual chromosome that contributes to the downregulation of miR-99a-5p/miR-100-5p and miR-125b-5p, but that both are downregulated, though at different levels in different cell lines.

The significant correlation between pri-miRNAs derived from the same chromosome (pri-miR-100/-miR-125b-1 on chromosome 11 and pri-miR-99a/-miR-125b-2 on chromosome 21) is in agreement with the transcription of the single polycistronic sequence from which both miRNAs are derived. However, the significant correlation between the expression of the pri- transcripts from the two chromosomal loci (pri-miR-99a/-miR-125b-2 with pri-miR-100/-miR-125b-1 also implies a common regulatory mechanism of regulation for both chromosomes.

To further verify the contribution of both chromosomes 11 and 21 to the miRNA production and downregulation, and the possible common mechanisms of regulation of miRNAs expression on both chromosomes, the expression of the -3p passenger strand for each of the miRNAs of interest was studied. Since the passenger stands derived from each chromosomal site are different in sequence, this allowed the dissection of the contribution of each chromosome which could not be carried out if the -5p guide strand of the mature miRNAs under investigation alone was investigated, as these are identical between chromosomes,

Not dissimilarly from the pri-miRNA, the passenger strand is usually quickly degraded, and it is difficult to quantitate. Optimisation was required but, using 10 times the normal amount of RNA used to detect guide strands, miR-99a-3p/-100-3p/-125b-1-3p/-125b-2-3p were quantitated in both mGCT cell lines and normal controls. All -3p miRNAs tested were downregulated, when compared with gonadal control. Moreover -3p miRNA exhibited in the different cell lines the same downregulation pattern seen for their respective pri-miRNAs.

For instance, in 1411H and GCT44, both yolk sac tumours derived cell lines, transcription of miR-100-3p and miR-125b-1-3p was more reduced than transcription of miR-99a-3p and miR-125b-2-3p. On the other hand, the TCam2 cell line showed similar level of downregulation on

both chromosomes, whereas in 2102Ep transcripts appeared to origin less form chromosome 21 than from chromosome 11.

The differences observed between the patterns of transcription in the different cell lines could be linked to the tumour subtype the cell lines were derived from. Accordingly, both yolk sac tumour derived cell lines, GCT44 and 1411H, have the same pattern of regulation, while 2102Ep and TCam2 derived from embryonal carcinoma and seminoma respectively, appear to have their own different characteristic regulation. However, further subtype specific cell lines or tumour samples would have to be tested to produce conclusive data about subtype specific chromosomal regulation. In time and if confirmed such differences may be exploited to differentiate between mGCT subtypes, for example, for diagnostic purposes.

Significant correlation between expression of -3p miRNAs derived from the same chromosome (miR-100-3p/-125b-1-3p on chromosome 11 and miR-99a-3p/-125b-2-3p on chromosome 21) was seen, as well as between expression from the different chromosomal loci (miR-99a-3p/-125b-2-3p with miR-100-3p/-125b-1-3p). These significant correlations again suggest that each cluster is transcribed as a single transcript and that both chromosomal loci may be regulated by a common regulatory mechanism.

These findings together, strongly support the hypothesis that both chromosomes contribute to the deregulation of the miRNAs of interest and that this downregulation has an origin at the transcriptional level. This mechanism is therefore different to the post-transcriptional block, previously described by our group and mediated by LIN28, which regulates the expression of let-7a, another miRNA transcribed from the same two chromosomal loci (Murray *et al*., 2013). A differential regulation for *let-7a* and miR-100/-125b-1 has also been described recently in response to TGF-β signalling. In this setting, miR-100/-125b-1 are upregulated through SMAD2/3, whereas *let-7a* remains unchanged, due to the fact that TGF-β also induces LIN28B, consequently inhibiting *let-7a* maturation (Ottaviani *et al*., 2018).

MiR-99a/-125b-2 and miR-100/-125b-1 are embedded in the introns of a long non-protein coding genes (lincRNA) called MIR99AHG (LIN00478), and MIR100HG, respectively. A few recent papers have showed that the expression level of MIR99AH (LINC00478) and

MIR100HG were similar to the one of the mature form of miR-99a-5p/-100-5p and miR-125b-5p, suggesting that the miRNA of chromosome 11 and 21 were produced from a single primary transcript together with their host lincRNAs (Emmrich, Rasche, *et al.*, 2014; Emmrich, Streltsov, *et al.*, 2014; Lu *et al.*, 2017; Ottaviani *et al.*, 2018).

In agreement with these data, expression of the two lincRNAs was downregulated in all the four mGCT cell lines when compared with gonadal control. Significant levels of correlation were seen between the lincRNA (MIR99AHG or MIR100HG) with pri-miRNA, miRNA-3p and miRNA-5p originating from the same chromosomal locus, strongly suggesting that the miRNAs are transcribed all together with their host gene as a single transcript. Furthermore, the significant correlation between the expression of lincRNA MIR99AHG on chromosome 21 and the expression of lincRNA MIR100HG on chromosome 11 points once more towards a common mechanism of transcriptional downregulation acting on both chromosomal loci.

Finally, the analysis of the expression of BLID, a protein coding gene transcribed separately from the miRNAs at the same loci on chromosome 11, showed downregulation of the expression of this gene. This suggest that the whole region is downregulated in a similar manner and with a process that is not specific to miRNA biogenesis.

The findings presented in this chapter therefore point to three conclusions: first, the downregulation of the expression is most likely to occur at the transcriptional level, second, the miRNAs within a cluster are transcribed as a single transcript and third, there is likely to be a common mechanism of regulation between the two miRNA clusters chromosomal loci. The precise mechanism by which the downregulation of transcription occurs, is still unclear. It cannot be excluded that deregulation of proteins involved in pri-miRNAs processing may also have a role in the control of miRNAs transcription. However, the fact that this mechanism appears to control transcription of both coding and non-coding transcripts, suggest that the mechanism is unlikely to involve processes specific for miRNA biogenesis, and that further studies should focus on a process potentially resulting on broader changes, such as epigenetic modulation of transcription. The conclusions presented in this chapter formed the basis for the work presented in the following chapter.

Chapter 4   Investigation of the role of genomic methylation on the regulation of the expression of miRNA in mGCTs cell lines

4.1   Introduction

The data presented in the previous chapter suggested that the downregulation of miR-99a-5p/-100-5p and miR-125b-5p occurs at a transcriptional level. Interestingly, half of the miRNAs found in humans have been identified to be closely associated with CpG islands (Weber *et al.*, 2007), specific areas in which, in mammalian species, DNA methylation occurs (McCabe, Brandes and Vertino, 2009; Jones, 2012). Epigenetic mechanisms, such as promoter methylation or histone modifications, are one of the most common mechanisms of miRNA regulation (Saito and Jones, 2006; Suzuki *et al.*, 2012). In human cancer, CpG island promoter hypermethylation has been identified as one of the causes of miRNA loss of expression (Saito *et al.*, 2006b; Lujambio *et al.*, 2007; Toyota *et al.*, 2008).

Both miR-125b-5p and miR-100-5p have been previously described to be dysregulated by epigenetic mechanisms: silencing of miR-125b-5p was shown to be related to promoter hypermethylation in invasive breast cancer (Y. Zhang *et al.*, 2011a) and in colorectal cancer (Chen and Xu, 2015). Similarly, miR-100 has been shown to be downregulated due to hypermethylation of his host gene MIR100HG (Chen *et al.*, 2014).

Furthermore, methylation mechanisms of regulation play an important role in mGCT transformation, differentiation and treatment response (Koul *et al.*, 2002). Numerous studies,

have shown global methylation of GCTs, but also a substantial difference between GCT subtypes, with seminomas being overall less methylated than non-seminomas (Jeyapalan *et al*., 2011; Amatruda *et al*., 2013b; Shen *et al*., 2018). Many key regulators genes of methylation (such as DNMT3B) have been found to be increasingly expressed in non-seminomas cell lines and paediatric samples compared with seminoma, providing a possible explanation for this difference (Palmer *et al*., 2008).

Due to the known importance of methylation as a regulatory mechanism in mGCTs, and its role in miRNA deregulation, this chapter aims to investigate whether aberrant methylation in the promoter region or in areas upstream to the miRNAs of interest is a credible mechanism by which the expression of miRNA miR-99a-5p/-100-5p and miR-125b-5p is downregulated in mGCTs.

4.2   Results

4.2.1 Effects of treatment with 5-azacytidine on miRNAs

To investigate whether methylation played a role in the downregulation of miR-99a-5p/-100-5p and miR-125b-5p, I assessed the effects of DNA demethylation on the expression of these miRNAs. 2102Ep (derived from an embryonal carcinoma) and 1411H (derived from a yolk sac tumour) cell lines, were treated with 5-azacytidine (5-aza). Cell lines were treated for 4 days with 10 μM of 5-aza, without replenishment of the media, and miRNA expression measured on day 4 (d4).

Three housekeeping (HK) genes - RNU24, RNU38B and RNU44 - were assessed to ensure that their expression did not change as a result of the treatment. These snRNA (small nuclear RNA) were chosen as they have been shown by others to be stably and homogeneously expressed in malignant and normal tissue samples and cell lines (Palmer *et al.*, 2010; Genovesi *et al.*, 2012; Schwarzenbach *et al.*, 2015). Reliability of the HK genes was evaluated using two criteria: the first one considered the standard deviation between treated and untreated samples, while the second assessed the consistency of expression of the HK genes within individual samples by looking at relative expression of the housekeeping genes vs each other. For an HK to be acceptable therefore, the standard deviation (σ) of its expression in different samples had to be ≤1, while its expression ratio to the other HKs had to be similar in all samples.

RNU38B, did not meet either criteria, since had a standard deviation between treated and untreated 2102Ep cells of σ=1.18 and its expression ratio *vs* the other HKs was smaller in the treated than in the control samples. Therefore, this gene was discarded for subsequent data normalisation. The other two HK genes had a low standard deviation between untreated and treated samples in both cell lines (RNU24 σ= 0.21 and 0.11 for 2102Ep and 1411H respectively and RNU44 σ= 0.46 and 0.14 for 2102Ep and 1411H respectively) and consistent

expression ratios, and were therefore used to normalise the expression levels of the miRNAs (Figure 4.1).
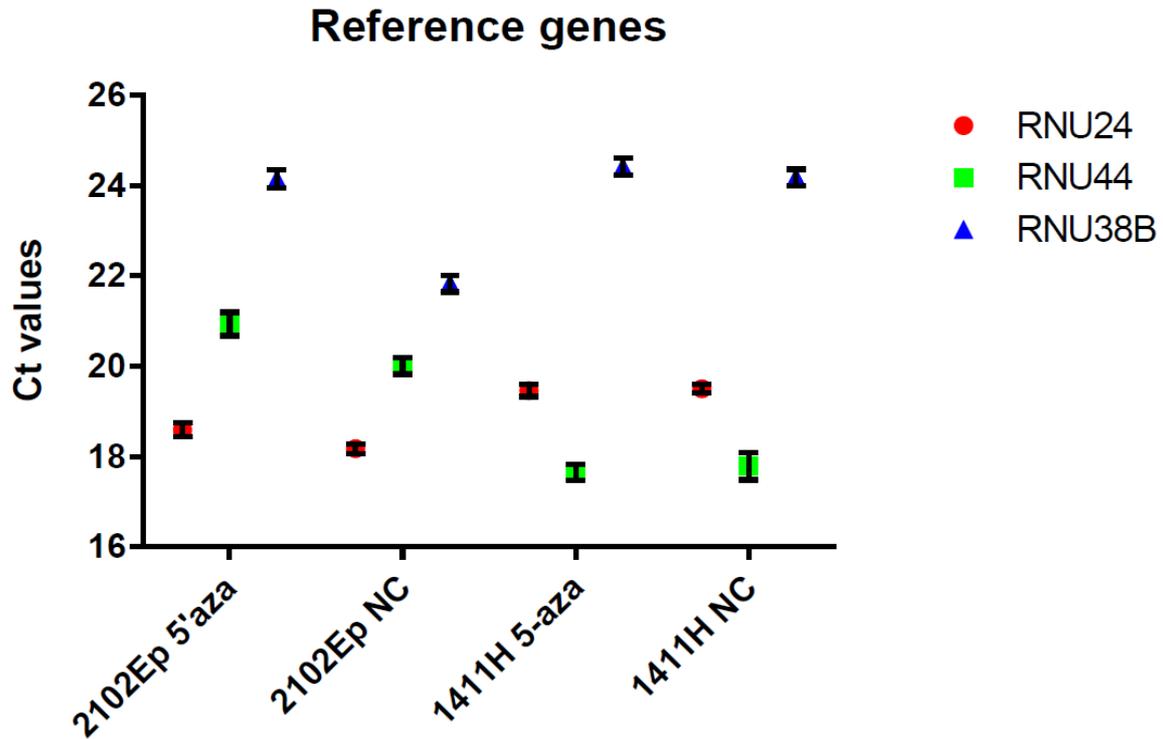
Figure 4.1: **Raw Ct values showing expression levels of three candidate housekeeping genes** (RNU24 in red, RNU44 in green and RNU38B in blue). The samples analysed were 2102Ep and 1411H cells in normal conditions and after treatment with 5-aza treatment for 4 days, are shown on the x axis. Data presented as mean ± SEM of n=3 biological replicates.

The potential effects of treatment with 5-aza on miR-99a-5p/-100-5p, miR-125b-5p and on a panel of non-differentially expressed miRNAs (miR-26b-3p, miR-500-3p and miR-137-3p) was then assessed (Figure 4.2). These non-differentially expressed miRNAs have previously been shown to have equal expression in both mGCT tissue samples and cell lines, and non-malignant control [gonadal controls and benign GCTs (i.e. teratomas)] (Palmer *et al*., 2010).
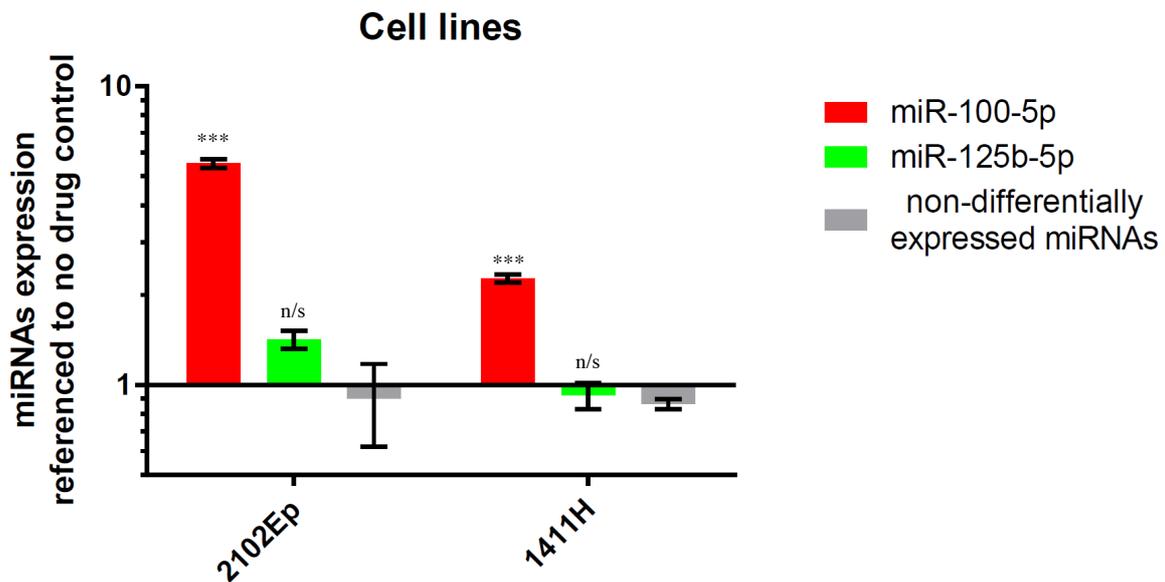
Figure 4.2: **Expression of miR-99a-5p/-100-5p (red bars), miR-125b-5p (green bars) in 2102Ep and 1411H cell lines following 5-aza treatment for 4 days.** Grey bars represent the average expression of non-differentially expressed miRNAs in mGCT cell lines under the same conditions. Expression was normalised to housekeeping genes and referenced the respective untreated control. Data presented as mean ± SEM of n=3 biological replicates. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

A significant upregulation of miR-99a-5p/-100-5p was seen after 5-aza treatment in both 2102Ep (*p*-value: 0.0001) and 1411H (*p*-value: 0.0001) cell lines. However, no significant difference in the expression of miR-125b-5p was detected in either 2102Ep or 1411H cells.

4.2.2 **Effects of multiple treatment with 5-azacytidine on 2102Ep cells**

To confirm whether transcription of miR-125b-5p was unchanged following treatment with 5-aza, the experiment was repeated with one cell line, 2102Ep, using a modified protocol. Briefly, 2102Ep cells were treated with 10 μM of 5-aza, but the media containing fresh 5-aza was replenished every day. Quantification of miRNAs expression quantification was then performed on day 1, 2, 3 and 4 of treatment.

RNU24 expression was analysed to ensure that there was no effect of the adjusted 5-aza treatment throughout the time-course as previously described in the section 4.2.1. RNU24 was shown to be consistent across all the samples at all timepoints ($\sigma = 0.4$). RNU24 was therefore utilised as the housekeeping gene (Figure 4.3).
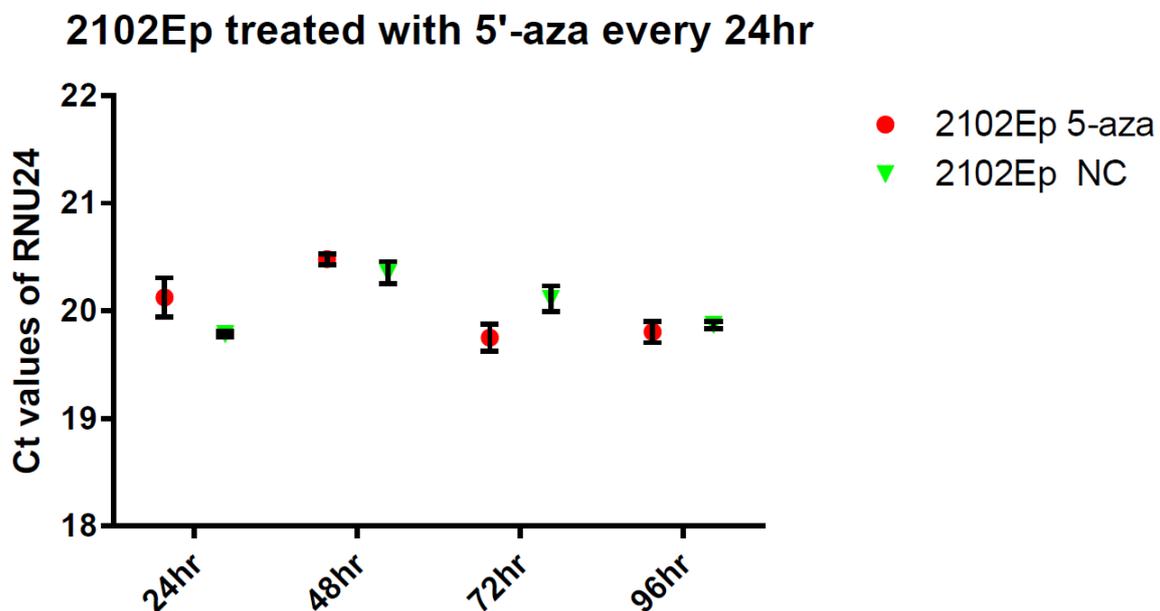
Figure 4.3: **Raw Ct values showing average expression levels of RNU24 housekeeping gene in normal conditions (in green) and after treatment (in red) with 5-aza treatment for 4 days.** Data presented as mean ± SEM of n=3 biological replicates.

The potential effects of treatment with 5-aza on miR-99a-5p/-100-5p, miR-125b-5p and the panel of non-differentially expressed miRNAs (miR-26b-3p, miR-500-3p and miR-137-3p) were then assessed (Figure 4.4) over a 4 days time-course.
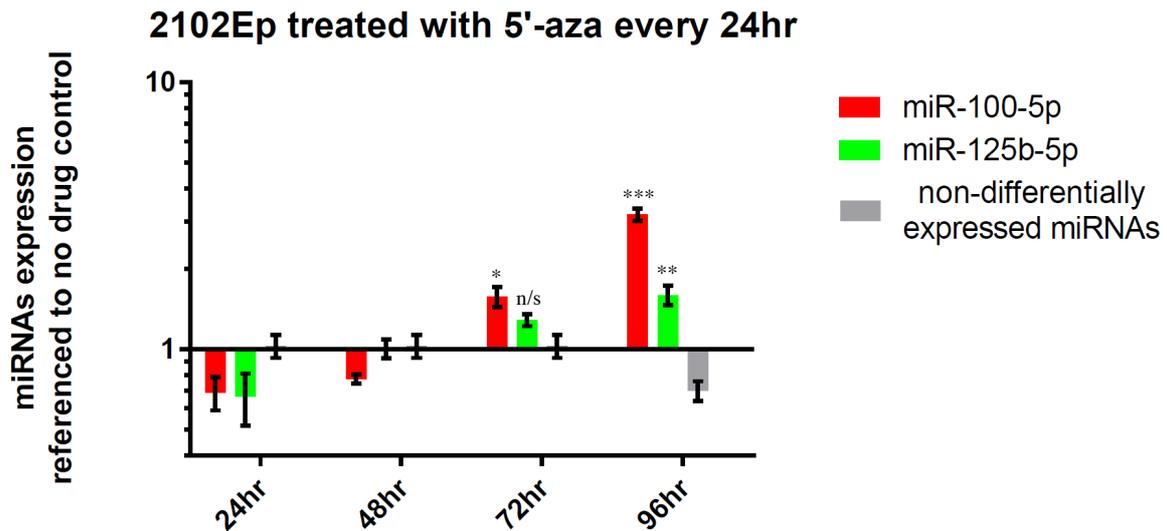
Figure 4.4: **Expression of miR-99a-5p/-100-5p (red bars), miR-125b-5p (green bars) in 2102Ep cells following 5-aza treatment for 4 days.** Grey bars represent the average expression of non-differentially expressed miRNAs in mGCTs cell lines under the same conditions. Expression was normalised to the respective day matched untreated control. Data presented as mean $\pm$ SEM of n=3 biological replicates. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

No significant difference in expression levels was detected for either miRNA at day 1 or 2 of treatment. A significant upregulation of miR-99a-5p/-100-5p after 5-aza treatment in 2102Ep compared with untreated control was seen at day 3 (*p*-value: 0.03), but no significant difference was detected for miR-125b-5p.

Differently from the previous experiment, a significant upregulation of both miR-99a-5p/-100-5p and miR-125b-5p was seen at d4 compared with untreated control (*p*-value: 0.0001 and 0.003, respectively). However, the levels of upregulation in this experiment for miR-99a-5p/-100-5p were lower than seen in the earlier experiment.

4.2.2.1 **Assessment of lincRNAs levels**

MiR-99a/-125b-2 and miR-100/-125b-1 are embedded in the introns of two long non-protein coding genes (lincRNAs), MIR99AHG (LIN00478) on chromosome 21, and MIR100HG on chromosome 11. Data presented in Chapter 3, (sections 3.2.7 to 3.2.8) show that there is a strong correlation between the expression of the mature miR-99a-5p, miR-100-5 and miR-125b-5p and their host lincRNAs. Consequently, the expression levels of these host genes were also assessed.

To identify which housekeeping genes were suitable for data normalisation, four housekeeping genes - *ACTβ, RPL13A, YWHAZ* and *HMBS* - were assessed to ensure that they were not affected by treatment with 5-aza. All HK genes had a low standard deviation between treated and untreated samples for each day (σ<1), and they had similar expression ratio across all the samples. These four HK genes were therefore utilised for normalisation purposes (Figure 4.5).
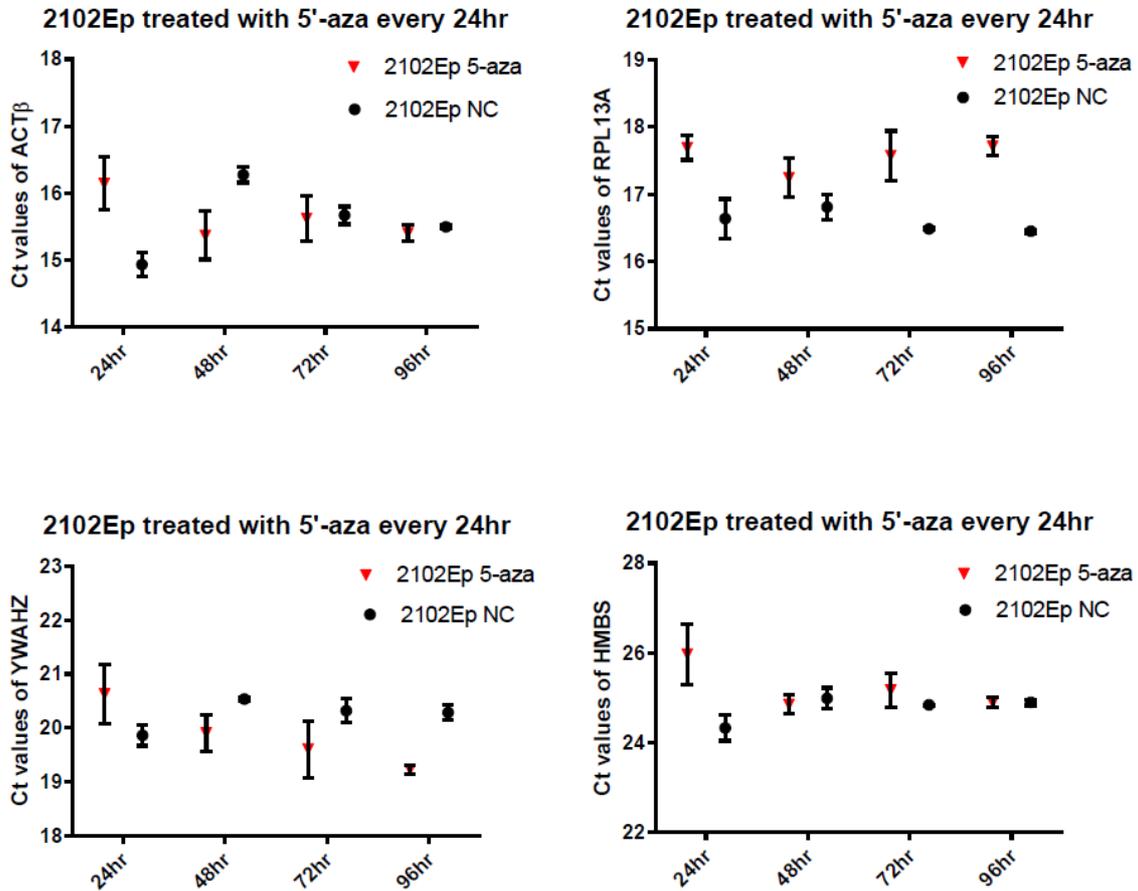
Figure 4.5: **Raw Ct values showing expression levels of four candidate housekeeping genes under normal conditions and after treatment of 2102Ep cells with 5-aza treatment for 4 days.** Data presented as mean ± SEM of n=3 biological replicates.

Analysis of the expression level of the host genes on chromosome 11 (MIR100HG) and chromosome 21 MIR99AHG (LINC00478) is shown in Figure 4.6.
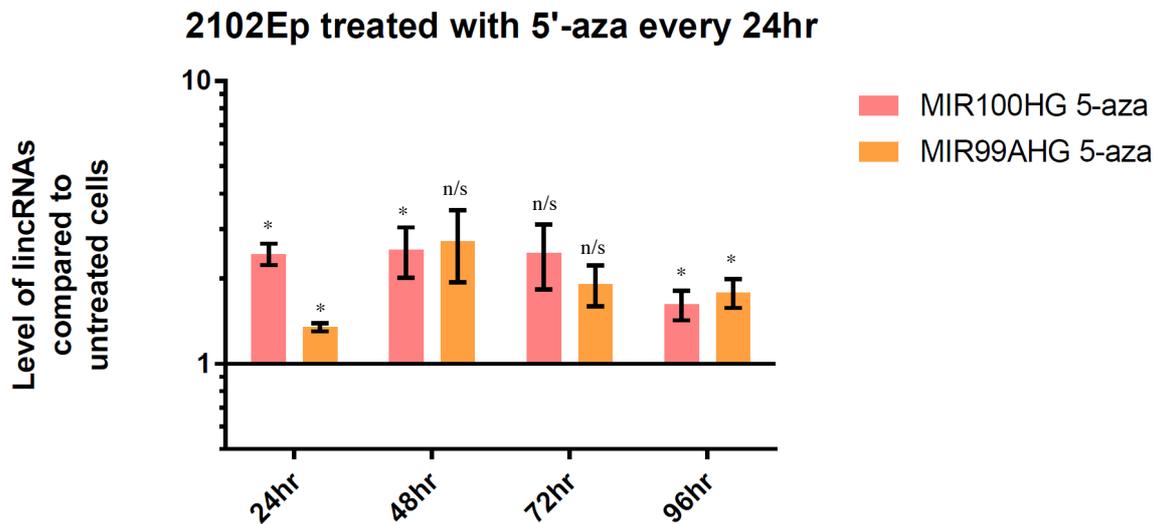
Figure 4.6: **Expression of MIR100HG and MIR99AHG (LINC00478) in 2102Ep treated with 5-aza.** Each day has been normalised individually with its respective untreated control. Data presented as mean ± SEM of n=3 biological replicates. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

Interestingly, levels of MIR100HG and MIR99AHG appeared significant elevated even after one day of 5-aza treatment (*p*-value: 0.0207 and 0.0162, respectively). Overall, MIR100HG and MIR99AHG levels were elevated across the 4 days, with significant upregulation for MIR100HG at day 2 and 4 of treatment (*p*-value: 0.0406 and 0.0326, respectively), and for MIR99AHG at day 4 (*p*-value: 0.0199).

In view of the changes in the level of expression of both miRNAs and their host genes after treatment with 5-aza, it was concluded that hypermethylation may potentially be the regulatory mechanism causing the downregulation of the expression of miR-99a-5p/-100-5p and miR-125b. Therefore, studies were carried out to identify more specifically the areas of this potential hypermethylation.

4.2.3 **Methylation profiles identified using bisulfite conversion**

Further experiments were done to identify the potential areas of hypermethylation that could be involved in miRNA regulation. Ideally, methylation profiles of genomic DNA from mGCTs cell lines would have been compared with normal gonadal cell lines. However, no cell lines derived from normal gonads are currently available, making the comparison between mGCT cell lines and appropriate controls challenging. Consequently, four cell lines derived from normal cervix tissue (NCX6, 95:14, 95:15 and 95:42), one primary fibroblast cell line (HFF-1) and one cell line derived from normal retinal pigmented epithelial cells (RPE) were used as controls. One control tissue derived from normal testis was also used together with four mGCT cell lines, namely 2102Ep, TCam2, 1411H, and GCT44.

4.2.3.1 **Primer design and validation**

DNA methylation can be studied by exploiting the ability of 5-methylcytosines to resist conversion by bisulfites to uracil as described in Figure 4.7.
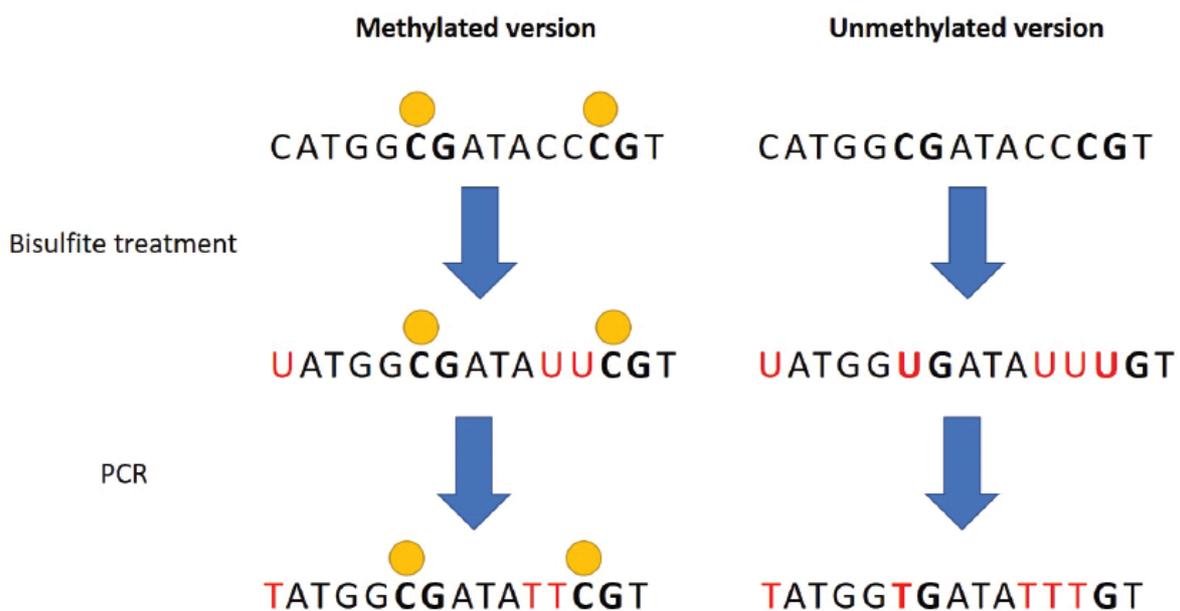
Figure 4.7: **Differential response to bisulfite conversion in methylated and unmethylated cytosines.** Unmethylated cytosines are converted to uracil by exposure to sodium bisulfite and can then be amplified to thymidine by PCR (right panel). The ability of 5-methylcytosines to resist conversion, which results in amplification by PCR as unchanged cytosines (left panel), allows evaluation of the thymidine *vs* cytosine relative ratio and can therefore be used therefore to quantify levels of methylation at each potential methylation site.

Little is known about the location of the potential areas of methylation in the genomic sequences upstream of the miR-99a/-125b-2 and miR-100/-125b-1. DNA methylation upstream of genes is most frequently found in areas which are rich in CpG (i.e. CpG islands) (McCabe, Brandes and Vertino, 2009; Jones, 2012), therefore I postulated that these may be areas of interest in my study.

Many bioinformatics tools are available to help identify CpG rich regions. For this analysis, I used MethPrimer, a free online tool that as well as predicting potential CpG islands in DNA sequences, designs bisulfite-converted DNA specific primers within or flanking the CpG island (Li and Dahiya, 2002) (Figure 4.8). Briefly, the sequence of interest is submitted via a menu, which also allows the definition of the minimal length required to define the CpG island (100 nt by default), as well as the CG percentage needed to identify a region as CG rich (50% by default). The program returns a visual representation of the characteristics of the sequence analysed, including the location of all CpG, the location of the potential CpG islands identified and primers for the analysis of such regions. The primers are designed by the program to avoid potential methylation sites, and therefore to be able to amplify both methylated and unmethylated templates and are matched in length and primer melting temperatures. Moreover, further refinement of the research parameters is also available, such as location, length and melting temperature. In this study, I also added the additional parameter of the melting temperature of the amplicon as one of the important criteria to identify primer sets. This was crucial as I assessed whether methylation was present using melting curve analysis, a technique first described by Ririe to differentiate between desired and undesired products of a PCR (Ririe, Rasmussen and Wittwer, 1997) and then adapted to become a method to quantify methylation (Ririe, Rasmussen and Wittwer, 1997; Wittwer *et al.*, 2003; Wojdacz and Dobrovic, 2007).

Briefly, the temperature at which dsDNA melts is determined by factors such as nucleotide sequence, length and GC/AT ratio. A methylated sequence of DNA, following bisulfite modification, will maintain a higher GC/AT ratio and so melting at a higher temperature than its unmethylated equivalent. Melting curve analysis can detect a single base difference but requires quite a bit of optimisation to give quantitative information on the composition of the amplicons. Rather than quantify the amount of methylation present, I used this approach to identify areas of interest based on differences in the melting curve of amplicons from mGCT cells DNA compared with DNA from normal cells.
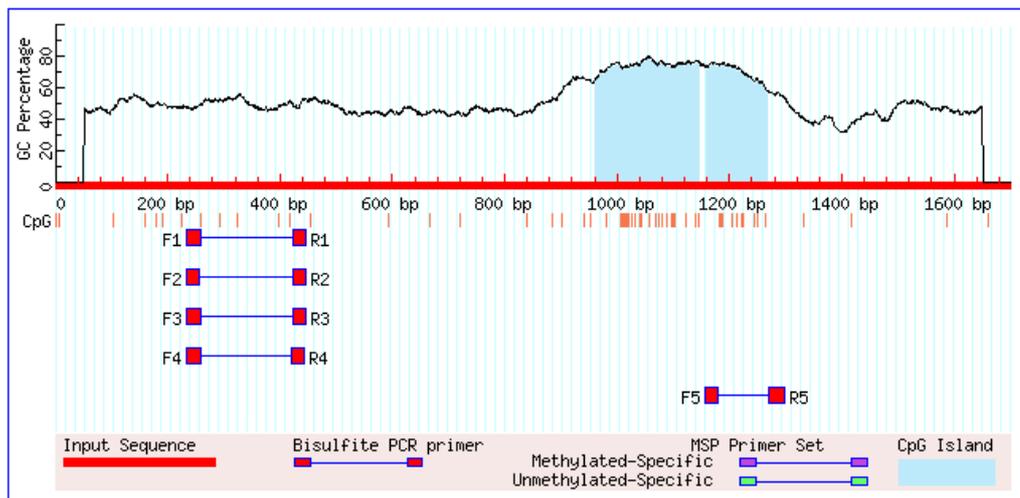
Figure 4.8: **Example of MethPrimer output.** The unconverted sequence of interest (max length 5Kb) is submitted, and CpG rich areas identified, using a 50% CG content as the cut-off. The sequence is then converted by the programme to that resulting from bisulfite conversion and primers are designed on the converted sequence in the area of interest.

For this approach to succeed, it was crucial that primers were designed to amplify regions (amplicons) that could be sufficiently different in their melting characteristic if they were methylated or not. Therefore, the melting temperature of amplicons derived from a 100% methylated DNA template versus that of amplicons derived from a 100% unmethylated template were calculated using the online program uMelt, which predicts the melting temperature of a PCR product based on its sequence. Primer pairs identified by MethPrimer were chosen if the difference between the predicted melting temperature of fully methylated

and fully unmethylated amplicon was >1°C wherever possible. Where necessary, primers were modified to fulfil this criterion.

The presence of CpG islands was searched for in an area up to 10Kb upstream of the transcriptional start of the lincRNAs transcripts on either chromosome. For reference, MIR100HG is located at the position nt 122,028,327-122,556,721 (GRCh38/hg38) on the minus strand of chromosome 11 while MIR99AHG (LINC00478) is located at the position nt 15,928,296-16,627,397 (GRCh38/hg38) on the plus strand of chromosome 21. Similarly, the presence of CpG islands was searched for up to 10Kb upstream of the transcriptional start of the individual microRNA. For reference miR-100 is located at the position nt 122,152,229-122,152,308, and miR-125b-1 is located at the position nt 122,099,757-122,099,844 on the minus strand of chromosome 11, whereas miR-99a is located at the position nt 16,539,089-16,539,169, and miR-125b-2 is located at the position nt 16,590,237-16,590,325 on the plus strand of chromosome 21.

Using the criteria listed above, I identified 11 areas on chromosome 11, and 5 areas on chromosome 21 that could be identified as CpG islands and were suitable for analysis by melting curve analysis. The characteristics of these areas are listed on Table 4.1 and Table 4.2.

PCR conditions were then optimised for each primer set identified by standard gradient PCR. Once the optimal annealing temperature was identified, Sybr green qPCR was run for each primer set using as templates the different cell lines (mGCTs *vs* non-malignant controls). The potential difference in the melting temperature of mGCT samples amplicons *vs* control amplicons was assessed to decide which areas may be differentially methylated. Where differences in melting curves and temperatures between mGCT and non-malignant control amplicons were detected, the qPCR product was purified and sent for Sanger sequencing. The chromatograms obtained by the sequencing process were then semi-quantitatively analysed using ImageJ as described in the Materials and Methods section 2.3.7, in order to determine methylation levels at each CpG sequenced. Finally established CpG methylation levels were compared between mGCT cell lines and non-malignant controls to detect any difference in methylation that may be significant. This process is summarised in Figure 4.9.

Chromosome 11							
Area	**Isla#01** **(MIR100HG.11.1)**		**Isla#02** **(MIR100HG.11.2)**		**Isla#03** **(MIR100HG.11.3)**		**Isla#04** **(MIR100HG.11.4)**
	Tm° CG	Tm° TG	Tm° CG	Tm° TG	Tm° CG	Tm° TG	Tm° CG	Tm° TG
10Kb upstream MIR100HG	77	76	83	81	77.5	77	88	83.5
	4 CpG		9 CpG		6 CpG		30 CpG	
	Length: 231 nt		Length: 262 nt		Length: 213 nt		Length: 425 nt	

Area	**Isla#05** **(miR-100.11.1)**		**Isla#06** **(miR-100.11.2)**	
	Tm° CG	Tm° TG	Tm° CG	Tm° TG
10Kb upstream miR-100	91.8	90	80.5	79.5
	7 CpG		6 CpG	
	Length: 440 nt		Length: 218 nt	

Area	**Isla#07** **(BLID.11.1)**	
BLID	Tm° CG	Tm° TG

	82	80.5						
	6 CpG							
	Length: 227 nt							
Area	**Isla#08** **(miR-125b-1.11.1)**		**Isla#09** **(miR-125b-1.11.2)**		**Isla#10** **(miR-125b-1.11.3)**		**Isla#11** **(miR-125b-1.11.4)**	
	Tm° CG	Tm° TG	Tm° CG	Tm° TG	Tm° CG	Tm° TG	Tm° CG	Tm° TG
5Kb upstream miR-125b-1	81.7	80.4	84.5	82	79	77	80.5	78.5
	6 CpG		8 CpG		14 CpG		17 CpG	
	Length: 374 nt		Length: 161 nt		Length: 344 nt		Length: 415 nt	

Table 4.1: **Areas identified by MethPrimer on chromosome 11.** For each area identified, the melting temperature of fully methylated amplicons (Tm° CG) and fully unmethylated amplicons (Tm° TG), the number of CpG identified and the length of the amplicon is reported. The name of the areas identified, correspond to the name of the primer pair and it will be used throughout this work.

Chromosome 21			
Area	**Isla#12** **(MIR99AHG.21.1)**		
10Kb upstream MIR99AHG	Tm° CG Tm° TG 85 83.5 5 CpG Length: 143 nt		
Area	**Isla#13** **(miR-99a.21.1)**	**Isla#14** **(miR-99a.21.2)**	**Isla#15** **(miR-99a.21.3)**
10Kb upstream miR-99a	Tm° CG Tm° TG 80 78 4 CpG Length: 170 nt	Tm° CG Tm° TG 89.6 88.7 3 CpG Length: 200 nt	Tm° CG Tm° TG 78 77 5 CpG Length: 401 nt
Area	**Isla#16** **(miR-125b-2.21.1)**		
5Kb upstream miR-125b-1	Tm° CG Tm° TG 82.5 81 15 CpG Length: 202 nt		

Table 4.2: **Areas identified by MethPrimer on chromosome 21.** For each area identified, the melting temperature of fully methylated amplicons (Tm° CG) and fully unmethylated amplicons (Tm° TG), the number of CpG identified and the length of the amplicon is reported. The name of the areas identified, correspond to the name of the primer pair and it will be used throughout this work.

Gradient PCR

100bp ladder 56.6 57.9 59.4 60.9 61.7

PCR optimisation:
Determination of an optimal annealing temperature for the PCR reactions for each CpG islands detected from MethPrimer

16 amplicons analysed from 10 samples 160 amplicons analysed overall

Melting curves

Melting curve:
Detection and comparison of Melting curves between mGCT cell line and a panel of normal cell line and tissue for every CpG island site analysed

16 amplicons analysed from 10 samples 160 amplicons analysed overall

Analysis of changes

Analysis of changes:
Screening of CpG island in mGCT cell line and a panel of normal cell line and tissue using sanger sequence of the PCR product of interest

2102Ep TTGGTTTCGGTTTTTTAGTAGTTTAAAGGG

Testis CCGGTTTTGGTTTTTTAGTAGTTTAAAGGG

GTCGGTCGTCGTGTGAAGTGCGGTTCGTTG

GTTGGTTGTTGTGTGAAGTGTGGTTTGTTG

6 amplicons analysed from 10 samples 60 amplicons analysed overall

Comparison between chromatograms

2102Ep   Testis

T T T C G   T T T T G

Initial comparison of putative CpG island:
Sequence Chromatogram analysis for each CpG island and comparison of mGCT cell line and a panel of normal cell line and tissue

6 amplicons analysed from 10 samples 60 amplicons analysed overall

Figure 4.9: **Workflow of the approach used to detect CpG islands and methylation status on chromosomes 11 and 21.**

A summary of the results obtained is shown in (Figure 4.10). Briefly, panel a) describes results for chromosome 11 and panel b) those for chromosome 21. The top row of both panels a) and b) shows the position of the methylation specific primers designed by MethPrimers (adjusted wherever needed) indicating the area of analysis. The following two rows indicate with a traffic light system whether differerences were seen between mGCT cell lines and normal controls, with green indicating clear differences, yellow uncertain differences and red no differences seen. The upper part of this results section refers to data obtained with the melting curves analysis, the lower to the results obtained from the analysis of the chromatograms obtained by Sanger sequencing. Overall, of all the potential regions identified as CpG islands by MethPrimers, 16 regions could be optimally amplified by PCR. Of these, seven showed obvious differences in the melting curves of products obtained from mGCT cell lines versus control cell lines, four more showed minor differences and where therefore considered uncertain, and five showed no differences. The 11 regions showing potential difference were further analysed by Sanger sequencing. For these regions, methylation at each individual potential methylation site was quantified by looking at relative C/T peak height in the chromatograms, and five regions showed us clear differences in CpG content between mGCTs and non-malignant controls. These five regions were used to quantify by pyrosequencing the amount of methylation at individual sites of relevance.
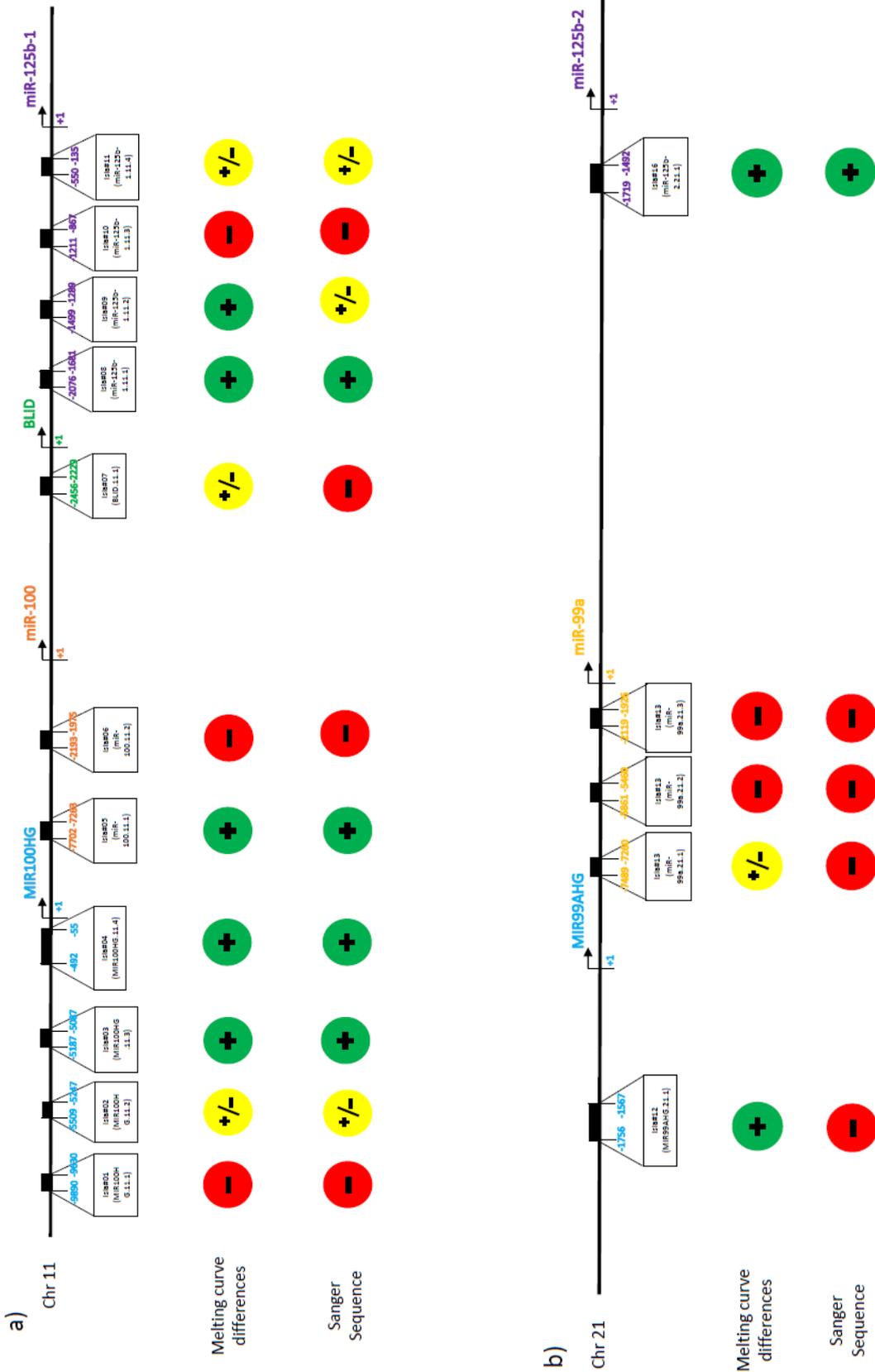
Figure 4.10**: Summary of the results obtained at each step of the process carried out to identify areas where methylation may occur.** Panel a) describes results for chromosome 11 and panel b) for chromosome 21. For both panels, the primer pair position is shown at the top, followed by differences in melting curves between mGCT and control DNA (middle) and results from the analysis of peak height in Sanger Sequencing chromatograms (bottom). Due to the length of the sequence examined (538,395 and 709,102bp on chromosome 11 and 21, respectively), the position of each primer pair is indicated relative to the gene or miRNA it is closest to. Each gene or miRNA and associated primer pairs is identified with a different colour. A red dot indicates that no differences were detected, a green dot indicates that differences were detected, and a yellow dot indicates uncertain results. The top panel represents chromosome 11 and in the lower panel is chromosome 21.

4.2.3.2 **Semi-quantitative analysis of CpG rich regions analysed on chromosome 11**

Following the process outlined above, semi-quantitative analysis of methylation was carried out for five regions (Figure 4.11).
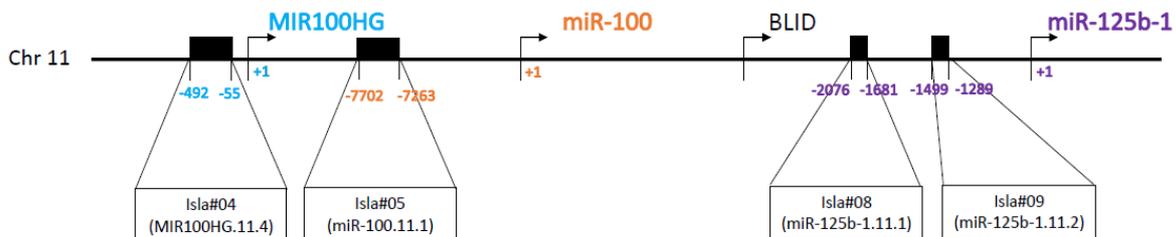
Figure 4.11: **CpG rich region selection in chromosome 11.** CpG rich islands analysed by semi-quantitative method and their location on chromosome 11 with reference to the miRNA they are upstream of – chromosome distribution is not to scale.

Four of these regions are located on chromosome 11 and were Isla#04 (MIR100HG.11.4) at position nt 122,557,193-122,556,721 on the reverse strand, in front of the lincRNA MIR100HG, Isla#05 (miR-100.11.1) at position nt 122,159,946-122,159,506 in front if miR-100 on the reverse strand, Isla#08 (miR-125b-1.11.1) at position nt 122,101,829-122,101,434

and Isla#09 (miR-125b-1.11.2) at position nt 122,101,148-122,100,987 on the reverse strand, both in front of miR-125-b-1 (Table 4.1).

In Figure 4.12, Figure 4.13 and Figure 4.14, the average percentage of CpG methylation of the mGCT cell lines and non-malignant controls for Isla#4, Isla#5 and Isla#8, are presented. This is assessed in a semi-quantitative manner from Sanger sequence chromatograms using the relative height of C/T peaks to determine the percentage of methylation in mGCT cell lines and non-malignant controls. Table 4.3, 4.4 and 4.5 show the percentage of CpG methylation for Isla#04, Isla#05, and Isla#08 in each of the mGCT cell lines and non-malignant controls. It should be noted that many returned sequences were of insufficient technical quality to be consistently quantified, despite repeat optimisations, which resulted in absence of semi-quantification for Isla#09.

Analysis of Isla#04 (Figure 4.12 and Table 4.3), initially suggested that this area may be regulated by methylation as there are many CpG sites (30 in total) which are significantly methylated in mGCTs cell lines compared with non-malignant controls (Figure 4.12).

Figure 4.12: **Cumulative levels of endogenous CpG DNA methylation in the area Isla#04 (MIR100HG.11.4).** The graph represents in each colour the average percentage of methylation per CpG site (30 CpGs in total) in all the mGCT cell lines, and the non-malignant controls. The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). Consequently, if each of the 30 CpG sites was fully methylated (100%), the cumulative value would be 3000. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

Observing the specific methylation contribution of each mGCT cell lines and non-malignant control is important for an adequate interpretation of the data (Table 4.3). Isla#4 is highly methylated in 2012Ep cells and is unmethylated in normal cervix cell lines and normal testis tissue DNA. However, there appear to be a similar level and pattern of methylation on DNA from the cell line RPE (from CpG9 to CpG21), suggesting that this methylation pattern may not be specific to mGCT. Unfortunately, technical issues prevented further semi-quantitative analysis, and thus interpretation of these results in necessarily limited.

Isla#04 (MIR100HG.11.4)	Isla#04 CpG 1	Isla#04 CpG 2	Isla#04 CpG 3	Isla#04 CpG 4	Isla#04 CpG 5	Isla#04 CpG 6	Isla#04 CpG 7	Isla#04 CpG 8	Isla#04 CpG 9	Isla#04 CpG 10	Isla#04 CpG 11	Isla#04 CpG 12	Isla#04 CpG 13	Isla#04 CpG 14	Isla#04 CpG 15
TCam2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
2102Ep	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	≥75%	≥75%	25-75%	25-75%	0%	≥75%	25-75%	≥75%	25-75%
GCT44	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1411H	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Normal testis	n/a	n/a	n/a	25-75%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
NCX6 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	n/a
95:14 cervix	n/a	n/a	n/a	25-75%	25-75%	25-75%	0%	0%	0%	0%	0%	0%	0%	0%	0%
95:15 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
95:42 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
RPE	0%	0%	0%	0%	0%	0%	0%	0%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	25-75%
HFF-1	0%	0%	0%	0%	0%	0%	0%	≥75%	≥75%	≥75%	n/a	n/a	n/a	n/a	n/a

Isla#04 (MIR100HG.11.4)	Isla#04 CpG 16	Isla#04 CpG 17	Isla#04 CpG 18	Isla#04 CpG 19	Isla#04 CpG 20	Isla#04 CpG 21	Isla#04 CpG 22	Isla#04 CpG 23	Isla#04 CpG 24	Isla#04 CpG 25	Isla#04 CpG 26	Isla#04 CpG 27	Isla#04 CpG 28	Isla#04 CpG 29	Isla#04 CpG 30
TCam2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
2102Ep	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	≤25%	≥75%	25-75%	25-75%	25-75%	n/a	n/a	n/a	n/a
GCT44	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1411H	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Normal testis	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	n/a	n/a	n/a	n/a
NCX6 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
95:14 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
95:15 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
95:42 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
RPE	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
HFF-1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table 4.3: **Semi-quantitative assessment of percentage of CpG methylation.** The table show the area upstream of the lincRNA MIR100HG [Isla#04 (MIR100HG.11.4)] in mGCT cell lines and a panel of control samples, including both tissue and cell lines. Methylation above 75% is colour coded with red, methylation levels between 25 and 75% are shown in orange and low (≤25%) or no methylation in pale yellow. A large proportion of the Sanger sequencing performed was not able to be quantitated, despite multiple attempts to improve sequence quality, and are therefore labelled n/a.

Analysis of Isla#05, (Figure 4.13 and Table 4.4) upstream of miR-100, showed full methylation of all the CpG sites (7 in total) in mGCT cell lines compared with non-malignant controls (Figure 4.13).

Figure 4.13: **Cumulative levels of endogenous CpG DNA methylation in the area Isla#5 (miR-100.11.1).** The graph represents in each colour the average percentage of methylation per CpG site (seven CpGs in total) in all the mGCTs cell lines, and the non-malignant controls. The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). If each of the seven CpG sites were fully methylated (100%), the cumulative value would be 700. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

Isla#05 was fully methylated in 2012Ep cells, but unfortunately, no other mGCT cell line could be analysed for this region due to technical issues (Table 4.4). This region appeared to be generally unmethylated in normal cell lines, with occasional low to medium methylated CpGs and could therefore potentially be regulated by methylation in mGCTs. DNA from testicular tissue showed a somewhat more methylated region, but overall a much lower level of methylation compared with 2102Ep cells. Taken together, the data for this region point to a potential difference in the level of methylation in this region in mGCT derived cell lines versus normal cell lines and tissue.

Isla#05 (miR-100.11.1)	Isla#05 CpG1	Isla#05 CpG2	Isla#05 CpG3	Isla#05 CpG4	Isla#05 CpG5	Isla#05 CpG6	Isla#05 CpG7
TCam2	n/a	n/a	n/a	n/a	n/a	n/a	n/a
2102Ep	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%
GCT44	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1411H	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Normal testis	10%	25-75%	25-75%	25-75%	25-75%	≥75%	25-75%
NCX6 cervix	0%	0%	0%	0%	0%	0%	10%
95:14 cervix	n/a	n/a	n/a	n/a	n/a	n/a	n/a
95:15 cervix	0%	0%	25-75%	0%	0%	0%	0%
95:42 cervix	0%	0%	≤25%	≤25%	25-75%	0%	25-75%

RPE	n/a	n/a	n/a	n/a	n/a	n/a	n/a
HFF-1	0%	0%	0%	0%	0%	0%	≤25%

Table 4.4: **Semi-quantitative assessment of percentage of CpG methylation.** The table show the area upstream of miR-100 [Isla#05 (miR-100.11.1)] in mGCTs cell lines and a panel of control samples, including both tissue and cell lines. Methylation above 75% is colour coded with red, methylation levels between 25 and 75% are shown in orange and low (≤25%) or no methylation in pale yellow. A large proportion of the Sanger sequencing performed could not be quantitated, despite multiple attempts to improve sequence quality, and are therefore labelled n/a.

Analysis of Isla#08, (Figure 4.14 and Table 4.5) upstream of miR-125b-1, showed significant methylation of the CpG sites (six in total) in mGCTs cell lines compared with non-malignant controls (Figure 4.14).
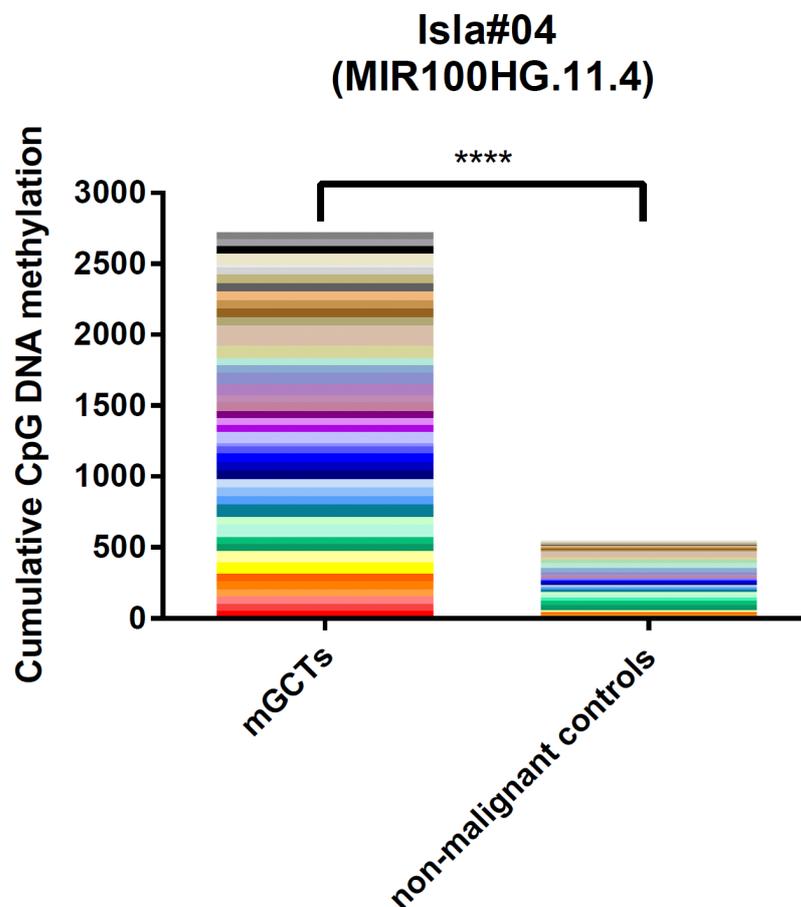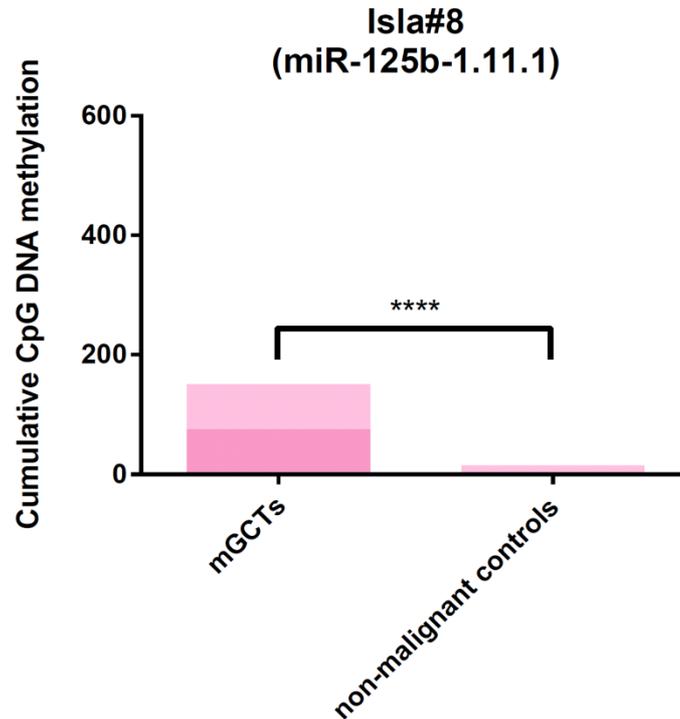

Figure 4.14: **Cumulative levels of endogenous CpG DNA methylation in the area Isla#8 (miR-125b-1.11.1).** The graph represents in each colour the average percentage of methylation per CpG site (six CpGs in total) in all the mGCT cell lines and the non-malignant controls. The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). If each of the six CpG sites were fully methylated (100%), the cumulative value would be 600. * $p < 0.05$, **$p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

When individual contributions were analysed, I observed a 100% methylation in three mGCT cell lines at the Isla#8 CpG4 and Isla#8 CpG5, which could not be seen in any of the four normal cell lines and tissue DNAs analysed (Table 4.5). However, no methylation was observed at these CpGs in the GCT44 cell line, suggesting perhaps that this cell line may be regulated differently at this locus. Taken together, the data suggest that there may be two CpG

that are utilised in mGCT to regulate this area in a manner different from normal cells/tissue, but that this may not be a ubiquitous mechanism.

Isla#8 (miR-125b-1.11.1)	Isla#07 CpG1	Isla#07 CpG2	Isla#07 CpG3	Isla#07 CpG4	Isla#07 CpG5	Isla#07 CpG6
TCam2	n/a	0%	0%	≥75%	≥75%	0%
2102Ep	n/a	0%	0%	≥75%	≥75%	0%
GCT44	n/a	0%	0%	0%	0%	n/a
1411H	n/a	0%	0%	≥75%	≥75%	n/a
Normal testis	n/a	0%	0%	≤25%	≤25%	n/a
NCX6 cervix	n/a	0%	0%	0%	0%	n/a
95:14 cervix	n/a	n/a	n/a	n/a	n/a	n/a
95:15 cervix	n/a	n/a	n/a	n/a	n/a	n/a
95:42 cervix	n/a	0%	0%	0%	25-75%	0%
RPE	n/a	0%	0%	0%	0%	0%
HFF-1	n/a	0%	0%	0%	0%	0%

Table 4.5: **Semi-quantitative assessment of percentage of CpG methylation.** The table show the area upstream of miR-125b-1 [Isla#08 (miR-125b-1.11.1)] in mGCT cell lines and a panel of control samples, including both tissue and cell lines. Methylation above 75% is colour coded with red, methylation levels between 25 and 75% are shown in orange and low (≤25%) or no methylation in pale yellow. A large proportion of the Sanger sequencing performed could not be quantitated, despite multiple attempts were made to improve sequence quality, and are therefore labelled n/a.

4.2.3.3 Semi-quantitative analysis of CpG rich regions analysed on chromosome 21

Only one region was selected on chromosome 21, Isla#16 (miR-125b.21.1) at position 16,582,169-16,582,372 on the forward strand in front on miR-125b-2 (Table 4.2).
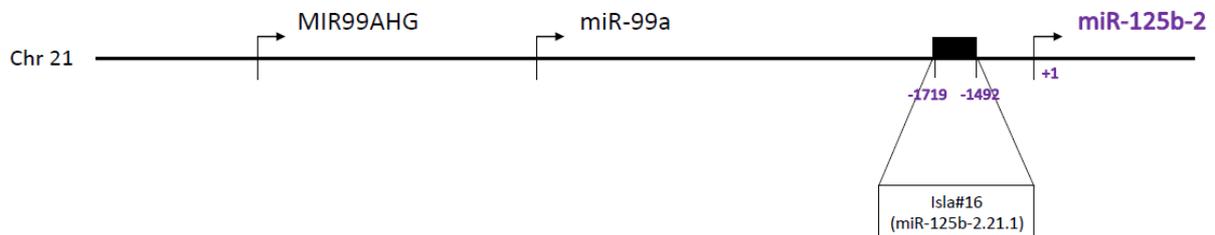
Figure 4.15: **CpG rich region selection in chromosome 21.** CpG rich islands analysed by semi-quantitative method and their location on chromosome 11 with reference to the miRNA they are upstream of – chromosome distribution is not to scale.

Analysis of Isla#16, (Figure 4.16 and Table 4.6) upstream of miR-125b-2, assessed in a semi-quantitative manner from Sanger sequence chromatograms using the relative height of C/T peaks to determine the percentage of methylation, suggests that this area may be regulated by methylation as there are many CpG sites (14 in total) which are significantly methylated in mGCT cell lines compared with non-malignant controls (Figure 4.16).
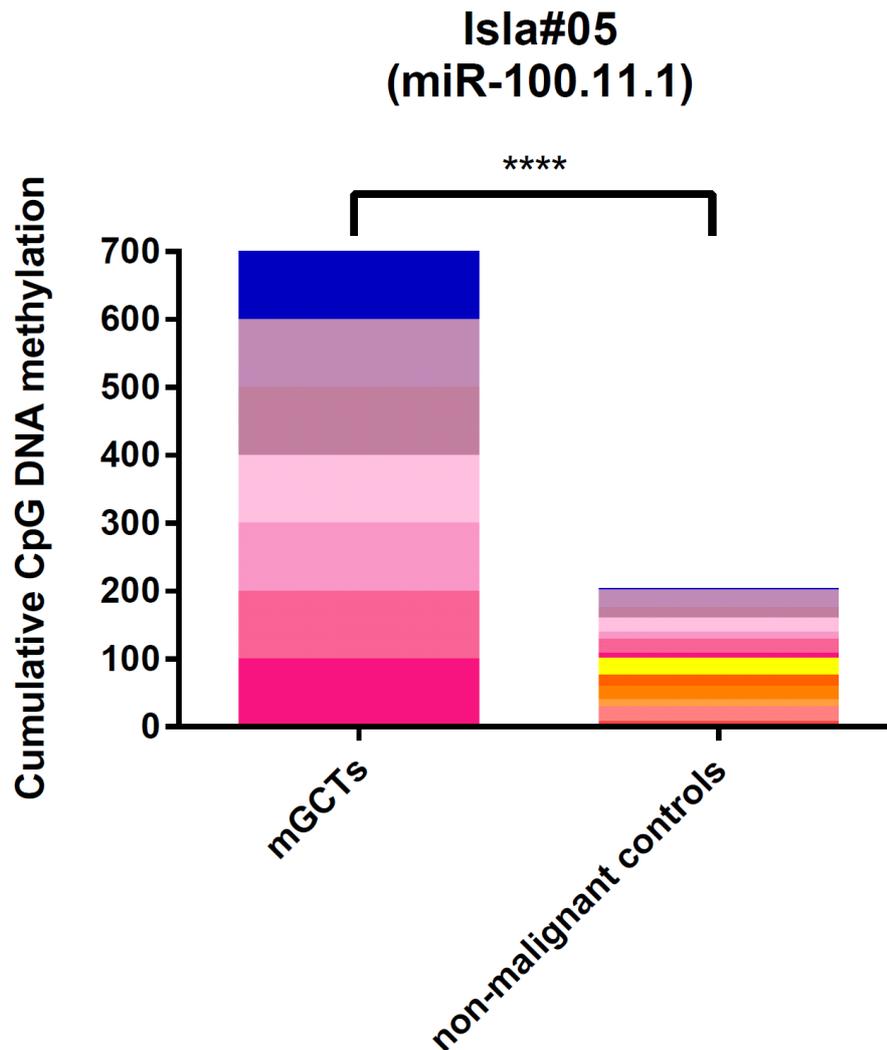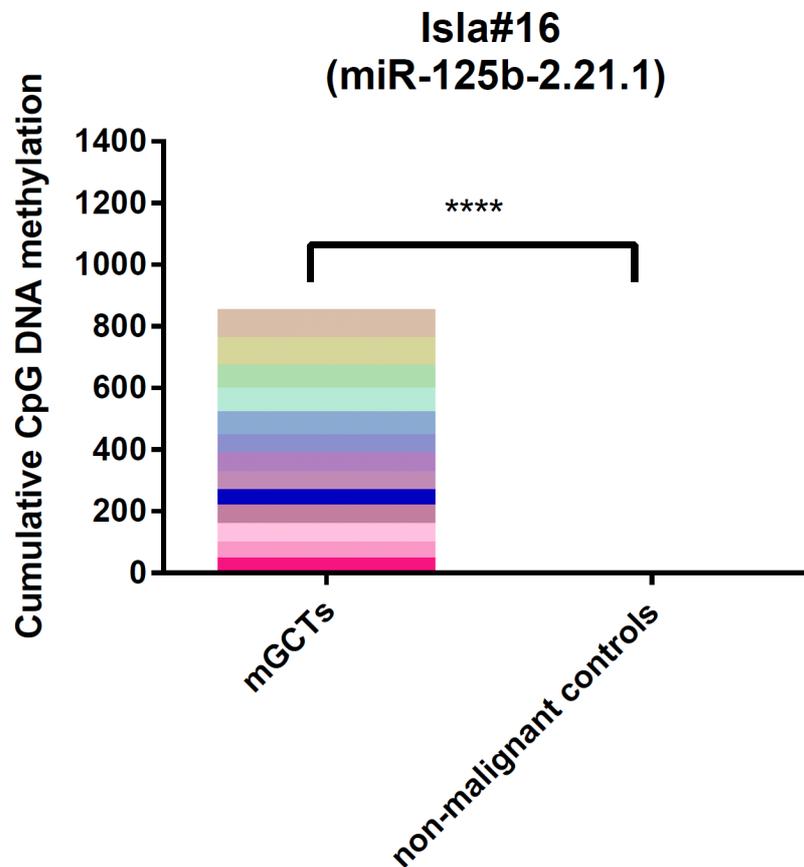
Figure 4.16: **Cumulative levels of endogenous CpG DNA methylation in the area Isla#16 (miR-125b-2.21.1).** The graph represents in each colour the average percentage of methylation per CpG site (14 CpGs in total) in all the mGCTs cell lines, and the non-malignant controls. The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). If each of the 14 CpG site were fully methylated (100%), the cumulative value would be 1400. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

Table 4.6 shows in detail the percentage of CpG methylation for each mGCTs cell lines and non-malignant control, for Isla#16. All normal cell lines showed no methylation in this region, whereas three mGCT cell lines showed medium to high levels of methylation for the same CpGs. This suggested that this area may be targeted by methylation to regulate expression on miR-125b-2 in mGCT cell lines in a manner that is different from normal cells/tissue.

Isla#16 (miR-125b-2.21.1)	Isla#16 CpG 1	Isla#16 CpG 2	Isla#16 CpG 3	Isla#16 CpG 4	Isla#16 CpG 5	Isla#16 CpG 6	Isla#16 CpG 7	Isla#16 CpG 8	Isla#16 CpG 9	Isla#16 CpG 10	Isla#16 CpG 11	Isla#16 CpG 12	Isla#16 CpG 13	Isla#16 CpG 14
TCam2	n/a	n/a	n/a	n/a	25-75%	25-75%	25-75%	25-75%	25-75%	n/a	n/a	n/a	n/a	n/a
2102Ep	25-75%	25-75%	0%	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	25-75%	≥75%	≥75%
GCT44	n/a	n/a	n/a	n/a	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%	≥75%
1411H	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Normal testis	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
NCX6 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
95:14 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
95:15 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
95:42 cervix	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
RPE	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

HFF-1	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 4.6: **Semi-quantitative assessment of percentage of CpG methylation.** The table show the area upstream of miR-125b-1 [Isla#16 (miR-125b-2.21.1)] in mGCTs cell lines and a panel of control samples, including both tissue and cell lines. Methylation above 75% is colour coded with red, methylation levels between 25 and 75% are shown in orange and low (≤25%) or no methylation in pale yellow. A large proportion of the Sanger sequencing performed could not be quantitated, despite multiple attempts to improve sequence quality, and are therefore labelled n/a.

Taken together, the data presented in sections 4.2.3.2 and 4.2.3.3 suggests a differential methylation in mGCT cell lines compared with normal cells/tissues.

4.2.4 Quantitative analysis of methylation by pyrosequencing

The semi-quantitative analysis of methylation in putative CpG islands on both chromosomes 11 and 21 points to differential methylation in mGCT cell lines that could contribute to the downregulation of the miRNAs miR-99a-5p, miR-100-5p and miR-125b-5p. However, the data obtained by the semi-quantitative method relies on the quality of chromatograms, and it was frequently difficult to derive consistent data, which resulted in an incomplete analysis. The issues experienced in this kind of analysis can be overcome by utilising a quantitative method such as pyrosequencing.

What distinguishes pyrosequencing from other traditional DNA sequencing methods such as Sanger sequencing, is the use of a chemical light reaction to detect the incorporation of dNTPs into the synthesising strand. The light generated is directly proportional to the amount of incorporated nucleotide and can be used to indirectly quantify incorporation. Moreover, pyrosequencing uses a sequencing-by-synthesis system in which nucleotides are dispensed one at a time, incorporated into the extending strand and degraded prior to the next nucleotide dispensation, making the pyrosequencing DNA sequence more accurate as well as quantitative. There are various applications for pyrosequencing, including SNP identification and quantification. In methylation analysis, the relative peak height of T versus C at a particular

CpG site in bisulfite converted DNA is measured to give percent methylation. (Harrington *et al.*, 2013). However, because of the nature of the methodology, only small regions up to 60 nt can be analysed at a time and knowledge of areas of interest are key for this methodology to be successful. The work carried out with the screening and semi-quantitative methods enabled the selection of such areas of interest and made possible the utilisation of this quantitative method.

4.2.4.1 Experimental design and optimisation

Optimisation of pyrosequencing is not trivial; therefore, this technology is used primarily in already well characterised assays for screening purposes, and less commonly in investigative assays. A workflow of the steps used for this study is described in Figure 4.17.

To prepare the template for pyrosequencing, the same combination of primers used for the semi-quantitative work were utilised, but one of the primers was substituted with a biotinylated one, so that the template could be immobilised for sequencing.

PCR optimisation

PCR optimisation:

PCR amplification of every sample (mGCT and control) using the new reverse biotinylated primer and the normal forward primer

5 amplicons analysed from 10 samples 50 amplicons analysed overall

Assay design

Sequencing primer design:

Sequencing primer design making sure to not include any CpG in it and with a product length of 60nt

assay1 sequencing primer →
TTTATTTTTAGTTTGATGAGGAAAGCGATAATTTATTAGGTATAGAATTAGGGGTTGTTTGCG

assay2 sequencing primer →
ATGTTTCGGATTTTTTGAAATATTGGTTAAAGGTTTTAAAGAATTCGTAGCGGGAGGTGCG

TAGTTCGTATTTAGTTTTTTTTTATTTCGTGTTTATTAAATGATAGTTTGGTG
← biotinylated reverse PCR primer

Designed at least 2 primers sequences per each template 18 assays in total

Example of histogram generated before the run

Pyrosequencing preparation:

For each sequencing primer, a new assay has to be made using the PyroMark Q24 software. Appropriate controls (such as a control for a correct C to T conversion in area of non CpG island) need to be included.

1 or 2 assays per each template were run in 10 samples, for a total of 20 assays per sample 100/200 runs

Efficient C to T conversion

Pyrosequencing run:

All the assay were run on a Q24 PyroMark machine and analysed by PyroMark Q24 software. When run, if a red alert sign happened, a new run had to be prepared.

Reliable data for at least 1 assay for each template was obtained in at least 6 to 8 samples, 30 to 40 runs in total

Figure 4.17: **Workflow for the detection of CpG methylation using pyrosequencing.**

The PCR products obtained after amplification with the biotinylated primer were then run on a gel, to make sure that the amplification was successful, and purified.

The sequencing assays were then designed. Briefly, sequencing primers are designed complementary to the DNA strand containing the biotinylated primer. For methylation analysis, these sequencing primers are designed upstream of the CpGs to analyse and depending on the template composition, each assay may cover one or more CpGs of interest. The number of primers designed depends primarily on the sequence of the template, in particular on the distribution of the CpG. As only short sequences up to 60 nt can be quantified in each assay, CpG that are close together may only require one assay and hence a sequencing primer, whereas separate single CpG may need to be quantified on their own assay and primer.

For each of the five CpG islands analysed by pyrosequencing several sequencing primers were utilised. Details are listed in Table 4.7:

Island analysed	# of assay designed	# of CpG analysed	# of sample
Isla#4 (MIR100HG.11.4)	7	11	4 mGCT 2 non-malignant control
Isla#5 (miR-100.11.1)	3	5	4 mGCT 3 non-malignant control
Isla#8 (miR-125b-1.11.1)	4	4	4 mGCT 4 non-malignant control
Isla#9 (miR-125b-1.11.2)	2	5	4 mGCT 4 non-malignant control

Isla#16 (miR-125b-2.21.1)	2	12	4 mGCT 4 non-malignant control

Table 4.7: **Details of the pyrosequencing assays.** The table shows in detail, the islands analysed, the number of assays designed per island, the number of CpG analysed and the number of samples used per assay.

Finally, specific assays were designed for each sequencing primer using the PyroMark Q24 software. As well as investigating the level of C and T incorporation at a specific CpG site, the assays have to fulfil other requirements, primarily the presence of a series of appropriate controls including assay controls (e.g. normalisation of single peak and repeated nucleotides peak height, baseline light emission, to mention a few) and target specific controls (e.g. for methylation studies bisulfite conversion efficiency). The programme helps with the design, but further optimisation is generally required. Once the assay and relative conditions are set up, the assays can be run multiple times: the output will consist of percent methylation readouts calculated automatically using the relative C/T peaks and referenced to the inserted controls (Figure 13).

4.2.4.2 **Quantification of Methylation by Pyrosequencing**

For each of the five CpG islands, where differential methylation was either seen or suggested by the semi-quantitative data analysis, pyrosequencing was performed. The CpG islands studied were: Isla#04 (MIR100HG.11.4) at position nt 122,557,193-122,556,721 on the reverse strand, in front of the lincRNA MIR100HG, Isla#05 (miR-100.11.1) at position nt 122,159,946-122,159,506 in front if miR-100 on the reverse strand, Isla#08 (miR-125b-1.11.1) at position nt 122,101,829-122,101,434 and Isla#09 (miR-125b-1.11.2) at position nt 122,101,148-122,100,987 on the reverse strand, both in front of miR-125-b-1, all located on chromosome 11 (Table 4.1). Isla#16 (miR-125b-2.21.1) at position nt 16,582,169-16,582,372 on the forward strand in front on miR-125b-2 on chromosome 21 (Table 4.2).

Figure 4.18 shows a diagram summarising the quantitative data obtained after analysis of the Isla#04 (MIR100HG.11.4). The data were initially collected using seven individual assays that

overall cover 30 CpGs, however, due to technical issues due to the pyrosequencing technique, data from only 11 CpG sites are reliable in four mGCTs cell lines and 2 non-malignant controls (namely normal tissue testis and HFF-1).

Figure 4.18: **Schematic representing levels of methylation on Isla#04 (MIR100HG.11.4) located on chromosome 11 upstream of the lincRNA MIR100HG.** CpGs are indicated as circles across the area. Black circles indicate full methylation, white circles indicate no methylation. A half white and half black circle indicates partial methylation (< 50%) and three quarters black circle indicates the CpG is mostly methylated (>50%). The four GCT cell lines, and the normal testis sample represent data from each individual listed sample, whereas normal cell lines shows the percentage of methylation established averaging data from five individual normal cell lines.

The data showed a region of full methylation spanning CpGs #4.6 to #4.16 in three of the mGCT cell lines. Such region of methylation is not seen in the remaining mGCT cell line (TCam2) nor in normal testicular tissue. Partial methylation in this region was seen in the

normal cell line control, primarily in the HFF-1 cell lines, whereas no methylation was detected.

Data obtained after analysis of the Isla#05 (miR-100.11.1), 7702bp to 7263bp upstream of miR-100 transcription start site is shown in Figure 4.19. The data were collected using three individual assays that overall cover seven CpGs. As for Isla#04, technical issues made possible the analysis of five CpG sites in four mGCTs cell lines and three non-malignant control (normal tissue testis, HFF-1 and NCX6) (Figure 4.19).

Figure 4.19: **Schematic representing levels of methylation on Isla#05 (miR-100.11.1) located on chromosome 11 upstream of miR-100.** CpGs are indicated as circles across the area. Black circles indicate full methylation, white circles indicate no methylation. A half white and half black circle indicates partial methylation (< 50%) and three quarters black circle indicates the CpG is mostly methylated (>50%). The four GCT cell lines, and the normal testis sample represent data from each

individual listed sample, whereas normal cell lines shows the percentage of methylation established averaging data from five individual normal cell lines.

Analysis of this region confirmed the full methylation of CpGs #5.3 to #5.7 in 2012Ep cells seen by Sanger sequencing. The same region appears to be partially methylated in 1411H and GCT44 cells and more heavily methylated in TCam2 cells. Partial methylation is also seen in normal testis tissue, whereas the normal cell lines have little methylation in this area.

Figure 4.20, shows a diagram summarising the quantitative data obtained after analysis of the two CpG islands identified on chromosome 11 upstream of miR-125b-1, Isla#08 (miR-125b-1.11.1) and Isla#09 (miR-125b-1.11.2) which are located 1-2Kb upstream of miR-125b-1 transcription start site), The data were collected using different amplicons and for each amplicons individual assays were designed to respectively cover six plus eight CpGs. From the Isla#8 (miR-125b-1.11.1), data from four CpGs sites were obtained from four mGCTs and four non-malignant controls (normal tissue testis, HFF-1, NCX6 and RPE). From the Isla#9 (miR-125b-1.11.2), data from five CpGs sites were obtained from four mGCTs cell lines and four non-malignant control (normal tissue testis, HFF-1, NCX6 and RPE).
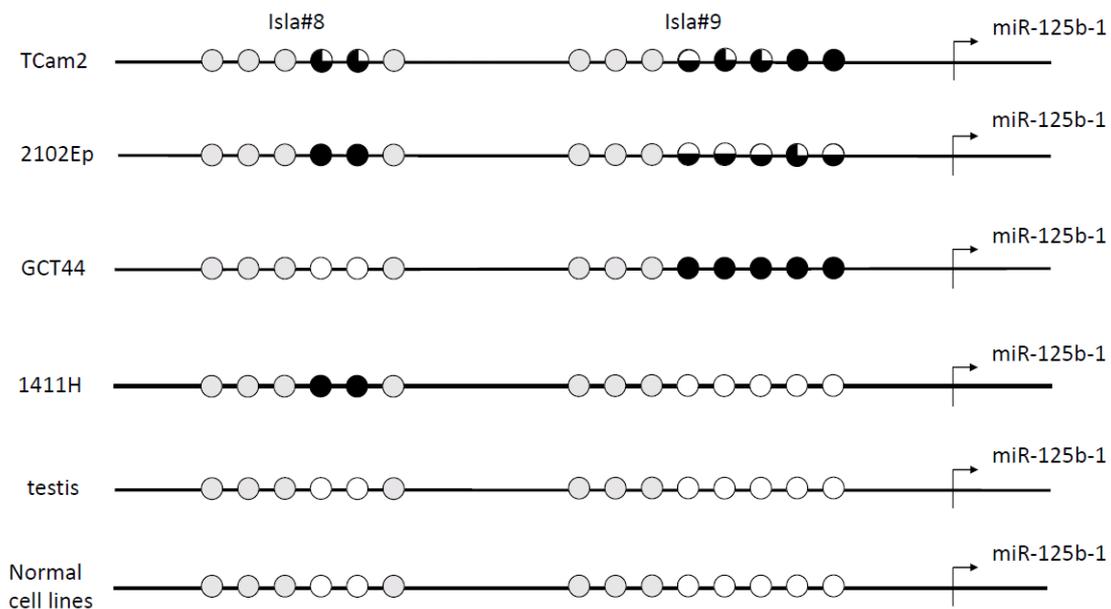
Figure 4.20: **Schematic representing levels of methylation on Isla#08 (miR-125b.11.1) and Isla#09 (miR-125b.11.2) located on chromosome 11 upstream of miR-125-b.** CpGs are indicated as circles across the area. Black circles indicate full methylation, white circles indicate no methylation. A half white and half black circle indicates partial methylation (< 50%) and three quarters black circle indicates the CpG is mostly methylated (>50%). The four GCT cell lines, and the normal testis sample represent data from each individual listed sample, whereas normal cell lines shows the percentage of methylation established averaging data from five individual normal cell lines.

The data showed significant or full methylation of CpGs #8.4 to #8.5 in three of the mGCT cell lines. Methylation is not seen in the remaining mGCT cell line (GCT44) nor in normal testicular tissue and cell lines. Partial methylation of CpGs #9.4 to #9.8 was seen in three of the mGCT cell lines, whereas there was no methylation on the fourth mGCT cell line (1411H). Interestingly, 1411H DNA was fully methylated on the CpGs #8.4 to #8.5, whereas GCT44 DNA, which was not methylated at these CpGs was fully methylated at CpGs #9.4 to #9.8. No methylation was detected in normal testicular tissue and cell lines.

Finally, data obtained after analysis of the Isla#16 (miR-125-b-2.21.1), located 1-2Kb upstream of miR-125-b transcription start site on chromosome 21 is shown in Figure 4.21. The data were collected using two individual assays that overall cover 12 CpGs. As for the other datasets, the data were derived from four mGCTs cell lines and four non-malignant control (normal tissue testis, HFF-1, NCX6 and RPE).
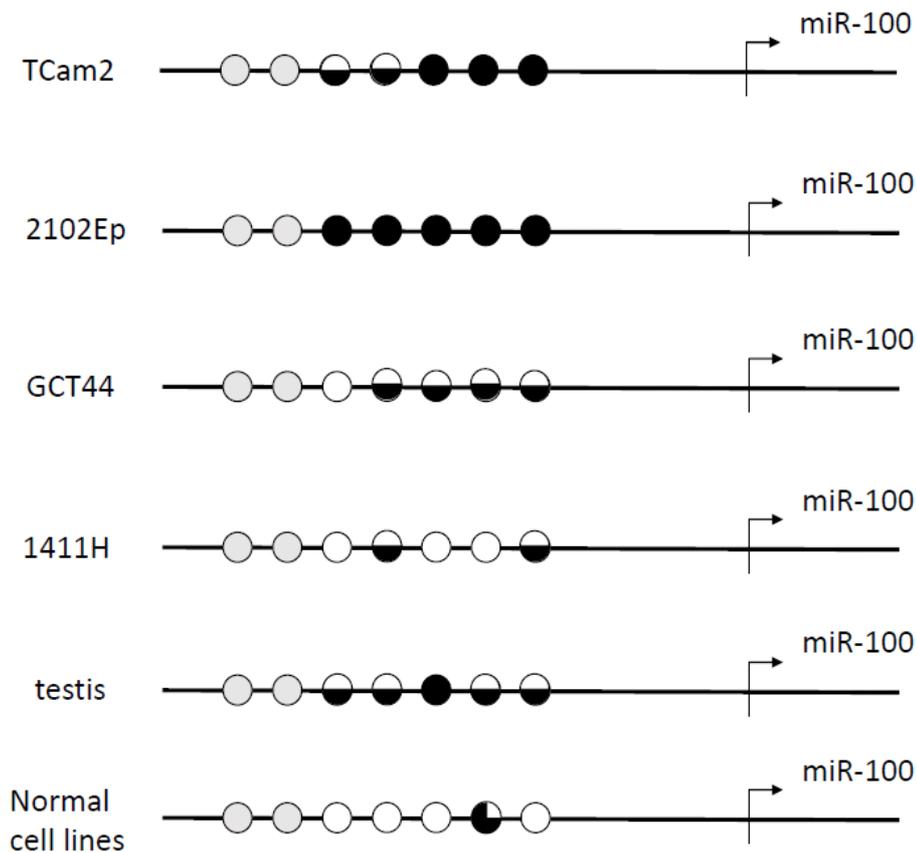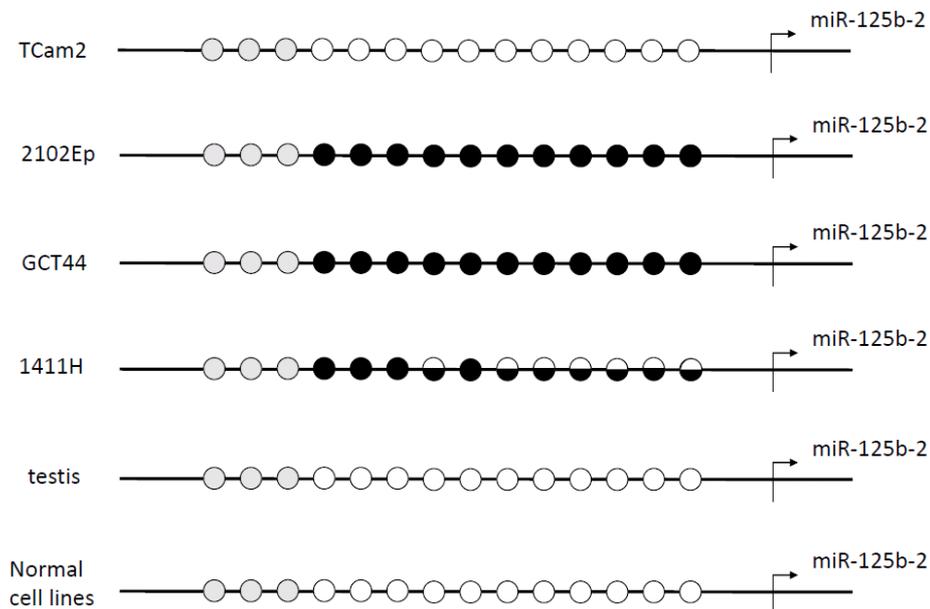

Figure 4.21: **Schematic representing levels of methylation on Isla#16 (miR-125b-2.21.1) located on chromosome 21 upstream of miR-125-b-2.** CpGs are indicated as circles across the area. Black circles indicate full methylation, white circles indicate no methylation. A half white and half black circle indicates partial methylation (< 50%) and three quarters black circle indicates the CpG is mostly methylated (>50%). The four GCT cell lines, and the normal testis sample represent data from each individual listed sample, whereas normal cell lines shows the percentage of methylation established averaging data from five individual normal cell lines.

The data showed full methylation of the region spanning CpGs #14.4 to #14.14 in three of the mGCT cell lines. Methylation is not seen in the remaining mGCT cell line (TCam2) nor in normal testicular tissue and cell lines.

4.2.5 **Analysis of methylation by pyrosequencing: summary**

The individual contribution of a single methylation site can be important for transcription, but in some cases the overall methylation of a region is more important than methylation at a specific site. The data shown in 4.2.4.2 suggests a certain degree of variability in the methylation profiles of the mGCT cell lines studied. Therefore, overall methylation of the region encoding for the miRNAs of interest was also assessed. To assess overall methylation in the samples studied, all mGCT cell lines and controls were split into two groups and individual methylation contributions added. Figure 4.22 shows cumulative CpG methylation in mGCT cell lines versus normal controls, both loci together (panel a) and each individual chromosome (panels b and c for chromosomes 11 and 21, respectively). The data shows a significant difference in cumulative methylation for mGCT vs normal controls both overall and when each individual chromosome is analysed ($p<2$x10^{-16} for all comparisons).
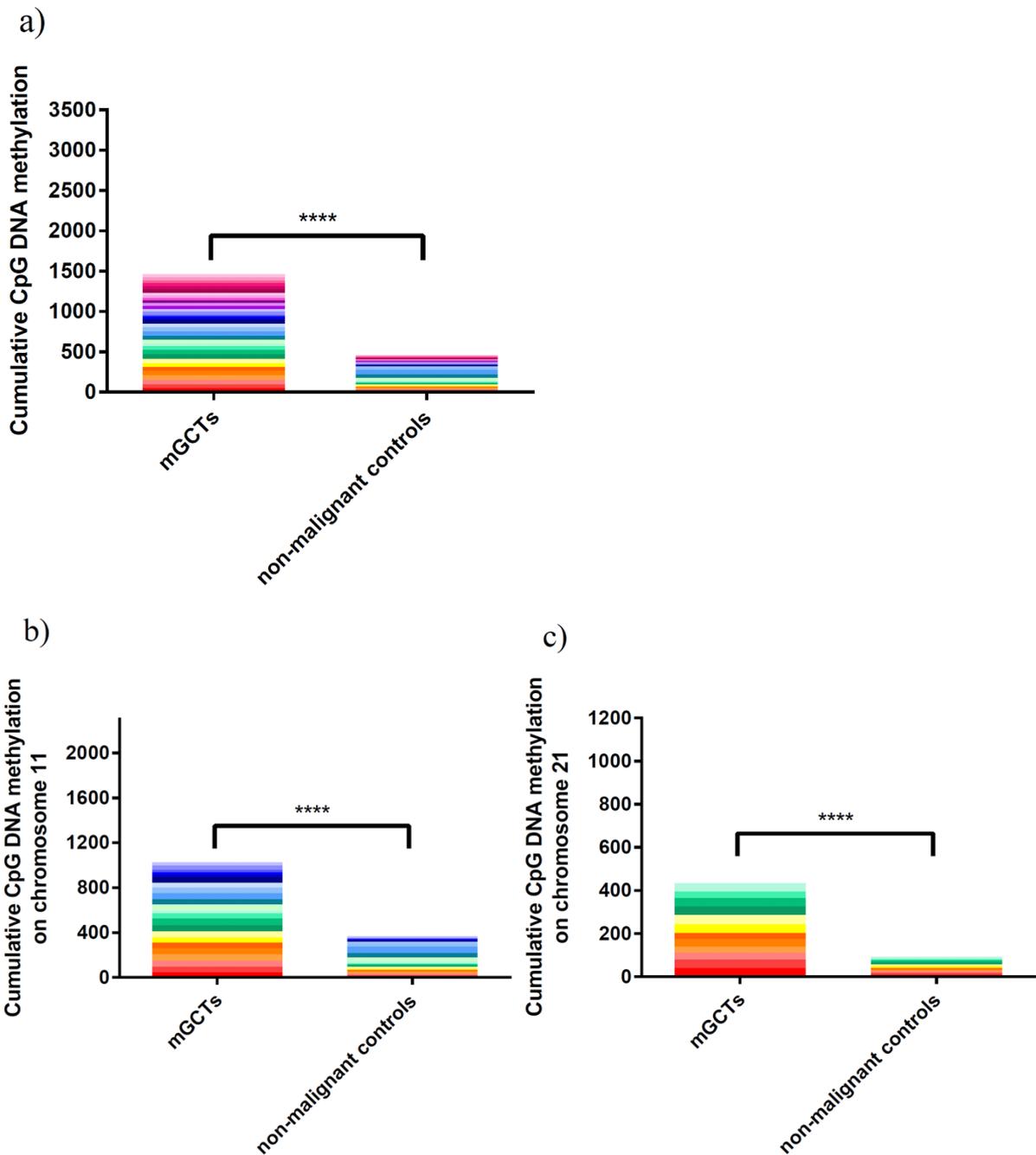
Figure 4.22: **Cumulative levels of endogenous CpG DNA methylation in mGCTs vs non-malignant control.** The graph represents in each colour the average percentage of methylation per CpG site in a) both chromosomes (11 and 21 for a total of 35 CpGs sites), b) only chromosome 11 (for a total of 23 CpGs sites) and c) only chromosome 21 (for a total of 12 CpGs sites). The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). If each CpG site was

fully methylated (100%), the cumulative value would be a) 3500, b) 2300 and c) 1200. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$.

Individual mGCT cell lines were analysed (Figure 4.23), both chromosomes together (panel a) and as individual chromosomes (panels b and c). The data showed a significant difference in cumulative methylation between 2102Ep, 1411H and GCT44 cells compared with testis tissue and normal cell lines. TCam2 cells showed overall a lower level of methylation at the loci analysed. This is due to the fact that in some areas, such as Isla#08 (miR-125b-1.11.1) and Isla#09 (miR-125b-1.11.2), the methylation pattern on TCam2 cells was in line with mGCT cells, but in others such as Isla#16 (miR-125b-2.21.1), the same methylation pattern was seen as in normal controls.
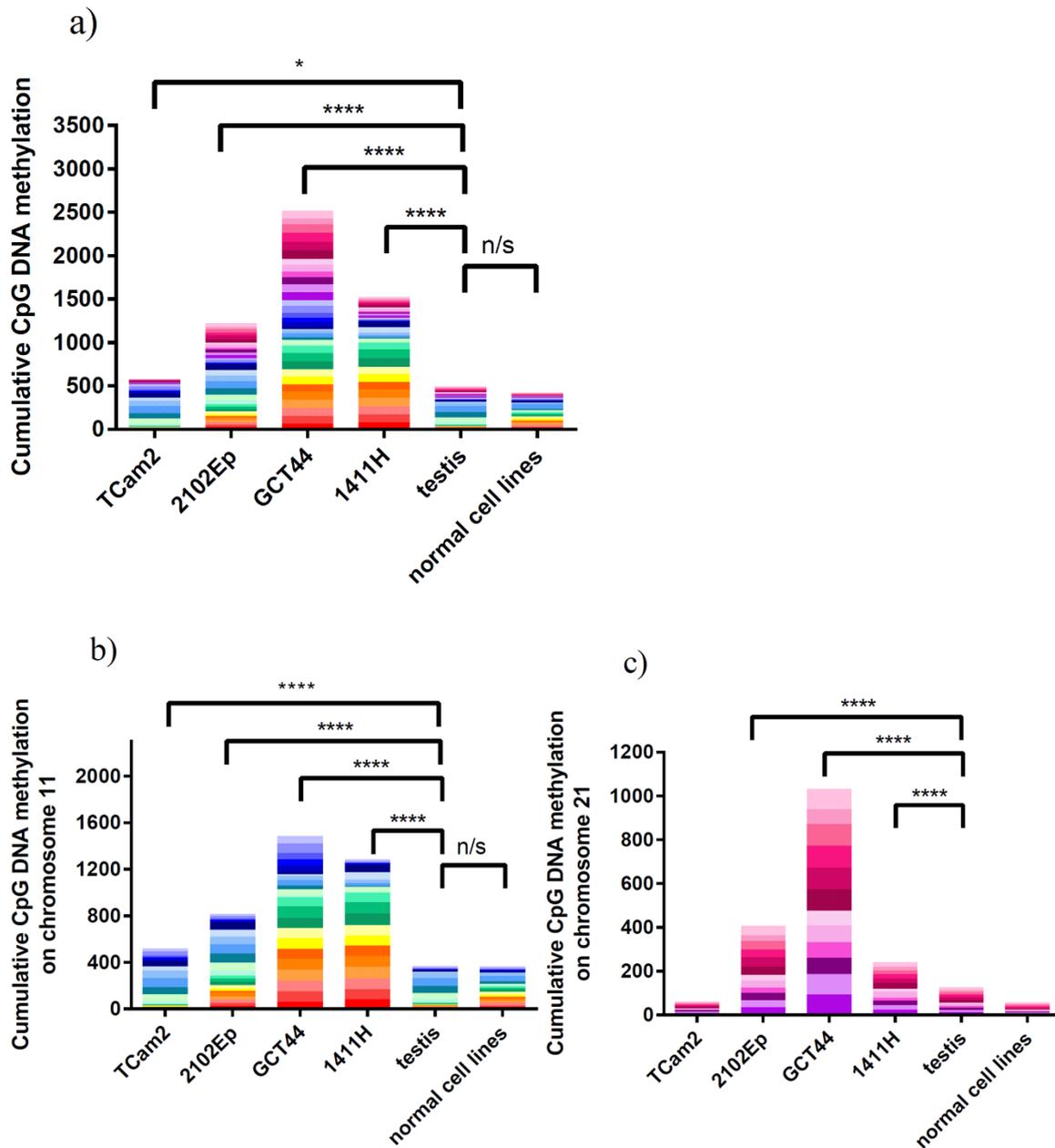
Figure 4.23: **Cumulative levels of endogenous CpG DNA methylation.** The graph represents in each colour the average percentage of methylation per CpG site in a) both chromosomes (11 and 21 for a total of 35 CpGs sites), b) only chromosome 11 (for a total of 23 CpGs sites) and c) only chromosome 21 (for a total of 12 CpGs sites). The percentage of methylation at each CpG site can range from 0 (no methylation), to 100% (fully methylated). If each CpG site was fully methylated (100%), the cumulative value would be a) 3500, b) 2300 and c) 1200. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

Together, the data presented here show a clear differential methylation between mGCT cell lines and normal controls in areas (CpG islands) that are likely to be of importance in the transcriptional regulation of miRNAs, providing an insight on the potential mechanisms for the downregulation of these transcripts in mGCTs.

4.3   Discussion

While it is clear that dysregulated miRNA expression contributes to the initiation and progression of human cancers, the mechanism underlying it is not yet fully understood. MiRNA dysregulation, like the dysregulation of oncogenes and tumour suppressor genes, can be caused by multiple mechanisms such as deletion, amplification, mutation, transcriptional dysregulation, epigenetic changes, etc (Iorio and Croce, 2012). As shown in Chapter 3, the downregulation of miR-99a-5p/-100-5p and miR-125b-5p in mGCTs cells is not genetically driven, occurs at early stages of their biogenesis and is not associated with the miRNA processing and maturation processes.

Aberrant expression of miRNAs located near CpG island, has suggested that methylation may have a role in tumour progression resulting from miRNA dysregulation (Brueckner *et al*., 2007). Methylation mechanisms are highly dynamic during germ cell development and aberrant DNA methylation has been suggested to be the cause of the mGCT pathogenesis. This is thought to be linked to potential errors in methylation control occurring during germ cell development (Lind, Skotheim and Lothe, 2007; Cheung *et al*., 2009). It is possible that dysregulation of miRNA expression may be associated with these potential errors in methylation control that specifically occur in mGCTs.

Importantly, silencing of miR-125b-5p was shown to be related to promoter hypermethylation in invasive breast cancer and in TGCT (Y. Zhang *et al*., 2011a; Batool *et al*., 2018) and in colorectal cancer (Chen and Xu, 2015), while miR-100 has been shown to be downregulated due to hypermethylation of his host gene MIR100HG (Chen *et al*., 2014).

Taken together, it seems plausible that aberrant methylation in the promoter region or in areas upstream to the miRNAs of interest may be the mechanism by which the expression of miRNA miR-99a-5p/-100-5p and miR-125b-5p is downregulated in mGCT.

To test this hypothesis, mGCT cell lines were treated with 5-azacytidine (5-aza). This is a well-established method to remove the inhibitory effects of methylation on gene transcription. When 5-aza, which is a cytidine nucleotide analogue, is incorporated into DNA instead of cytidine, methylation of the newly synthesised DNA is inhibited, as 5-aza blocks the activity of DNA

methyltransferases (Stresemann and Lyko, 2008). Therefore, transcription of genes downregulated by methylation would increase following treatment with this compound.

Initially, cells in culture were treated with 10 μM of 5-aza in the media for four days without any media change. Changes in miR-99a-5p/-100-5p expression were clearly noticeable, though a lesser effect was seen on expression of miR-125b-5p. This experimental set up was further optimised by changing the media daily, therefore replenishing the 5-aza every 24 hrs. With this condition, a significant change in expression was seen for both miR-125b-5p and miR-99a-5p/-100-5p. Different cell lines respond to 5-aza treatment differently, and others have described the need of replenishing the media frequently. For examples studies with embryonal carcinoma cell lines found that the best results were obtained by treating the cells for 3 days with drug replenishment each day (Albany *et al.*, 2017).

The changes detected in miR-99a-5p/-100-5p and miR-125b-5p transcription levels were significant after 3 and 4 days of 5-aza exposure. Interestingly, changes in expression detected for the host gene lincRNAs, MIR100HG (host gene of miR-100 and miR-125b-1) and MIR99AHG (host gene of miR-99a and miR-125b-2), were also significant only at 4 days of treatment, confirming the correlation already seen between the miRNAs and their host genes and described in Chapter 3.

These experiments suggested that methylation had a role in the regulation of miR-99a-5p/-100-5p and miR-125b-5p expression. Attempts to find information that would help identify areas of potential methylation were carried out by accessing data produced by other groups but publicly available. For instance, online data on genome wide DNA methylation profiles of germ cell tumours were analysed. These studies used the array Infinium HumanMethylation450 BeadChip which covers 450K methylation sites located mainly in promoter areas. Because long non-protein coding genes and miRNA do not have well characterised promoters, it was not possible to find any information about areas of potential interest in the database (Rijlaarsdam *et al.*, 2015).

Because of this lack of information about the specific location of the promoter regions of miRNAs, DNA sequences up to 10Kb upstream the starting site of the host genes (MIR100HG

and MIR99AHG) were analysed to detect any significant enrichment in CpG dinucleotides. The majority of methylation involved in gene silencing occurs where there is a high frequency of CpG dinucleotides clustered together, in areas known as CpG islands. The accepted definition of a CpG island is a region of DNA greater than 200 bp, with guanine/cytosine content above 0.5. MethPrimer, an online tool designed to identify CpG islands, was utilised to identify areas that fitted the definition of CpG islands (Li and Dahiya, 2002). Several areas of interest containing high density of CpG dinucleotides and showing the characteristics of a CpG island were identified in this manner. Once the areas of interest were identified, primers that would enable the analysis of several CpG were designed.

Bisulfite conversion-based PCR is to date the most commonly used technique for methylation mapping. Primers spanning CpG rich areas that could amplify both methylated and unmethylated bisulfite converted DNA, were designed and tested. A series of optimisation steps designed to help identify potential areas on methylation were carried out. The first step assessed whether methylation was present using melting curve analysis, a method used to quantify methylation (Ririe, Rasmussen and Wittwer, 1997; Wittwer *et al.*, 2003; Wojdacz and Dobrovic, 2007). This method exploits the fact that the temperature at which dsDNA melts is determined by factors such as nucleotide sequence, length and GC/AT ratio. Following bisulfite modification, a methylated sequence of DNA will maintain a higher GC/AT ratio and will have a melting temperature higher than its unmethylated equivalent. The analysis of the melting curve of an amplified products can give therefore information on the methylation status of a specific area (Lorente *et al.*, 2008).

Of all the PCR products analysed, five products corresponding to a region on chromosome 11 and one PCR fragment amplified from chromosome 21 had very different melting profiles when the product amplified from DNA derived from normal cells was compared with that amplified from the DNA of GCT cells. Three products amplified from chromosome 11 and one product amplified from chromosome 21 gave uncertain results. These 10 PCR products were therefore purified and sequenced to look for evidence of potential methylation. When a difference in CG or TG composition was detected, the chromatogram was analysed semi-quantitatively to determine the CG percentage composition between samples.

On chromosome 11, four areas (one in front of MIR100HG, one in front of miR-100 and two in front of miR-125b-1) had differential methylation in mGCT DNA compared with non-malignant controld, and three areas gave uncertain results. On chromosome 21, only one area (just upstream of miR-125b-2) showed differences in CpG abundance between mGCTs and non-malignant controls.

This initial screening gave strong evidence supporting the relationship between miRNA downregulation and a methylation mechanism of repression in mGCT. It also showed that the two chromosomal sites where the clusters of miRNAs were found are very different in terms of methylation density. However, the semi-quantitative nature of this analysis raised some uncertainty, and therefore a quantitative method was then utilised to measure precisely the level of methylation that could be found in these areas.

With the use of pyrosequencing, quantitative data on the level of methylation present at each CpG site identified was acquired and compared between PCR fragments amplified from mGCT cell lines, a normal testis tissue sample and a panel of normal control cell line. This analysis did not find consistent differences at specific CpG sites, but rather identified an overall different methylation profile between cancer and normal DNA across all the areas analysed. When overall methylation is analysed across these areas by plotting cumulative CpG DNA methylation, a strikingly significant difference was seen between malignant and non-malignant samples (Figure 5.22 panel A). This difference existed whether both chromosomal loci were analysed together (Figure 5.22 panel A) or separately (Figure 5.22 panel B and C for chromosome 11 and 21 respectively). Furthermore, when malignant cell lines were separated and looked at independently, again a striking difference was seen between each malignant cell line and the average of the non-malignant samples. Again, these differences were highly significant when both chromosomal loci were studied together (Figure 5.23 panel A) or separately (Figure 5.23 panel B and C for chromosome 11 and 21 respectively).

Interestingly, not all mGCT cell lines showed the same degree of methylation: in particular the embryonal carcinoma cell line (2102Ep) and yolk sac tumour cell lines (1411H and GCT44) were overall heavily methylated at the loci analysed, whereas the seminoma cell line (TCam2)

was more heavily methylated than normal controls, but to a much lesser extent than the other mGCT cell lines. In particular, at some loci, the methylation pattern on TCam2 cells resembled that of normal cells, where in others it resembled that of the other mGCT cell lines. Again, this was true whether both chromosomal loci are examined together or separately. These results are consistent with several studies that have investigated global methylation of paediatric GCTs, and have identified substantial difference between GCT subtypes, with seminomas being described as having overall a less methylated genome compared with non-seminomas (Smiraglia *et al.*, 2002; Smith-Sørensen *et al.*, 2002; Honorio *et al.*, 2003; Wermann *et al.*, 2010; Jeyapalan *et al.*, 2011; Amatruda *et al.*, 2013b; Noor *et al.*, 2016; Shen *et al.*, 2018). This different pattern of methylation is thought to derive from the fact that mGCTs originate from germ cells at different developmental stages and inherit their methylation profile from their ancestors (Sievers *et al.*, 2005). The zygote is the only totipotent cell of the organism because it will give rise to all cells of the embryo. However, male and female gametes in mammals originate from embryonic cells marked with their own landscape of DNA methylation and histone modifications. For the zygote to be totipotent, these pre-existing epigenetic marks must be reset during germ-cell specification and new ones established. This 'reset' process is termed epigenetic reprogramming (Chaganti and Houldsworth, 2000; Bastien, Smallwood and Kelsey, 2011), and takes place during migration and maturation of the primordial germ cells (PGCs). It occurs in two stages: a first phase of global DNA CpG demethylation takes place during the early phases of migration. However, specific areas like imprinted regions remain methylated until the PGCs arrive in the developing gonads where imprinting is subsequently gradually erased. After these maturing gonadal germ cells reach mitotic (male) or meiotic (female) arrest, de novo methylation is initiated, and uniparental sex specific imprinting is acquired.

Regarding partial methylation of TCam2 cells, it has been shown that certain genes can be downregulated to the same extent even though only partial methylation of a regulatory region has occurred. For example, SOX17 is underexpressed in non-seminomatous tumour cells to the same level that it is in yolk sac tumour cells, despite the level of methylation in the promoter region of embryonal carcinoma and teratoma cells being not as high as in a yolk sac tumour cell line (Noor *et al.*, 2016). Moreover, as the methylation mediated suppression of transcription appears to be linked to overall methylation of an area rather than to single CpG,

it cannot be excluded that adjacent areas that were not identified as methylated in the screening process described here may be methylated only in TCam2 cells and have the same suppressive function and ultimately mediate downregulation of the expression of the miRNAs. It is also possible that more than a mechanism is responsible for the downregulation of the expression of miRNAs in these seminoma cells.

In this context, it is important to mention the role of the genomic redundancy of miR-99a-5p/-100-5p and miR-125b-5p. Genomic redundancy is a mechanism that the genome has evolved as a strategy to ensure that important functions are regulated (Krakauer and Plotkin, 2002). Different strategies for the downregulation of the two chromosomes might have evolved as an extra control for their regulation. One different method of downregulation could derive from transcription factor binding upstream miRNAs or lincRNAs as many evidence have suggested this method to control the expression of miRNAs (Fazi *et al*., 2005; Fontana *et al*., 2007; Rosa *et al*., 2007).

Next, bioinformatic studies to look for transcription factor binding sites (TFBS) were performed in proximity to the miRNAs, in order to see if transcription factors could putatively bind to the methylated areas. Only one TFBS on chromosome 11 was found, upstream of miR-100, which binds a factor called FOXJ2, while five TFBS were found on chromosome 21 for the following transcription factors: CUX1 (one binding site in front of MIR99AHG), POUF3F2 (one binding site in front of miR-99a), Oct1 (two binding sites upstream miR-99a and miR-125b-2) and FOXJ2 (one binding site in front of miR-99a). Further work to assess and confirm the potential role of transcription factors on the regulation of the expression of miR-99a-5p/-100-5p and miR-125b-5p will be required to elucidate this further.

To conclude, the data presented in this chapter supports methylation as a mechanism of down-regulation of gene expression of miR-99a-5p/-100-5p and miR-125b-5p, in the studied mGCT cells. As no individual consistent sites of methylation were identified, it is unlikely that methylation regulated expression of the miRNAs *per se* is utilising individual or specific CpG sites. The overall mechanism appears instead to be linked to the presence of methylation over large parts of a CpG island. It is possible that this widespread methylation may give rise to a

repressive environment via the recruitment of histone modifiers and the consequent presence

repressive heterochromatin, but further work is required to confirm this potential mechanism.

Chapter 5   Investigation into the downstream consequences of miR-99-5p/miR-100-5p and miR-125b-5p downregulation

5.1   Introduction

MiR-99a-5p/-100-5p and miR-125b-5p have been identified as the most downregulated miRNAs in mGCTs. Data presented in Chapter 3 has shown that these two miRNAs are located in two clusters on chromosome 11 and 21. Despite being in two locations, the expression of these miRNAs is coordinated. Similarly, the processes that have resulted in the downregulation of these miRNAs have affected both chromosomal loci. In particular, it was shown in Chapter 4 that methylation at GpC sites, and consequent chromatin condensation, are responsible for the downregulation of these miRNAs. Having studied the upstream mechanisms which are responsible for the under-expression of miR-99a-5p/-100-5p and miR-125b-5p, in this Chapter the consequences of such deregulation will be explored, with particular focus to mechanisms that may be of relevance in mGCTs cancer progression.

MiR-99a-5p/-100-5p and miR-125b-5p have been shown to regulate different aspect of cancer progression and different biological function in the cells in different settings. For example in breast cancer, low expression of miR-125b-5p increases cell motility while in metastatic colorectal cancer tissues (CRC), the levels of miR-100-5p and miR-125b-5p were significantly lower than in CRC tissues without metastases (Akhavantabasi *et al*., 2012; Fujino *et al*., 2017). MiR-99a-5p/-100-5p have also been shown to be involved in cell cycle progression in renal

cell carcinoma via targeting of mTOR and miR-125b-5p plays an important role in regulating the proliferation of glioma stem cells by targeting E2F2 (Cui *et al*., 2012; N. Wu *et al*., 2012).

Due to the important role that miR-99a-5p/-100-5p and miR-125b-5p have in cancer, we wanted to evaluate whether the replenishment of these putative tumour-suppressor miRNAs in mGCT cell lines had phenotypic effects such as cancer cell invasion, migration and proliferation *in vitro*.

MiRNA replenishment using mimics is the most common way of miRNAs reintroduction in a biological system. Many studies have looked at different ways to optimise delivery by assessing transfection conditions as well as mimic concentration and composition. Single strand mimics are much less efficient than double strand mimics, as their purpose is to exactly mimic the function of mature miRNAs. Proteins involved in the processing of miRNAs such as DICER and AGO, only recognise double stranded miRNAs. Duplex unwinding is followed by RISC complex activation and subsequent target recognition (Chorn *et al*., 2012). Therefore, if the mimic is presented as a single strand, it will have lower probabilities of being recognised from the maturation complex.

Many studies have focused on optimising the concentration of mimic to be used in transfections. It was shown that specific effects on the target mRNA could be achieved using low concentrations of mimics (1 nM) (Jin *et al*., 2015) while at high concentrations (up to 100 nM) high molecular weight RNA complexes were forming intracellularly in the transfected cells. Therefore, the use of low concentrations of mimics was found to be more effective.

MiRNAs replenishment has become of increasing interest in the clinical field, as they may be used as potential therapeutic strategies. In 2014, the first clinical trial involving miRNA-replacement therapy for cancer in humans commenced. Phase I trials, utilising a nanoparticle-derived minicell delivery system to replenish levels of miR-16 in mesothelioma and non-small cell lung cancer have now been completed and Phase II trials are ready to start (van Zandwijk *et al*., 2017b).

In this chapter, miRNAs replenishment in mGCTs cell lines will be performed to look at any phenotypic change that they may induce. Following the restoration of miR-99a-5p/-100-5p and

miR-125b-5p, global mRNA levels will be assessed using microarray technology. Systematic interrogation using the bioinformatic tool *Sylamer* will then be utilised, to identify specific protein-coding targets altered following replenishment of the miRNAs. *Sylamer* will also be used to detect miRNA specific target genes which will be consequently used for pathway analysis to fully comprehend the biological relevance of miR-99a-5p/-100-5p and miR-125b-5p in mGCTs progression (van Dongen, Abreu-Goodger and Enright, 2008b).

5.2   Results

As discussed in Chapters 3 and 4, the downregulation of miR-99a-5p/-100-5p and miR-125b-5p is consistent across all mGCT subtypes, and mGCT cell lines, and is therefore a characteristic feature in these tumours (Figure in the first result chapter). Consequently, it is important to understand the biological changes which are mediated by their deregulation in mGCTs.

5.2.1   Relative contribution of each miRNA to the overall downregulation

As both families of miRNAs share almost identical seed region and are therefore predicted to interact with many common targets, it was important to determine the relative miRNA contribution to the overall downregulation mediated by each miRNA family. To this end, each miRNA was initially quantified using qRT-PCR. Subsequently, the RNU24-normalised Ct values derived from miR-99a-5p/-100-5p and miR-125b-5p was equated to 100% and a proportion for each miRNA, was assigned based on Ct values (Figure 5.1).
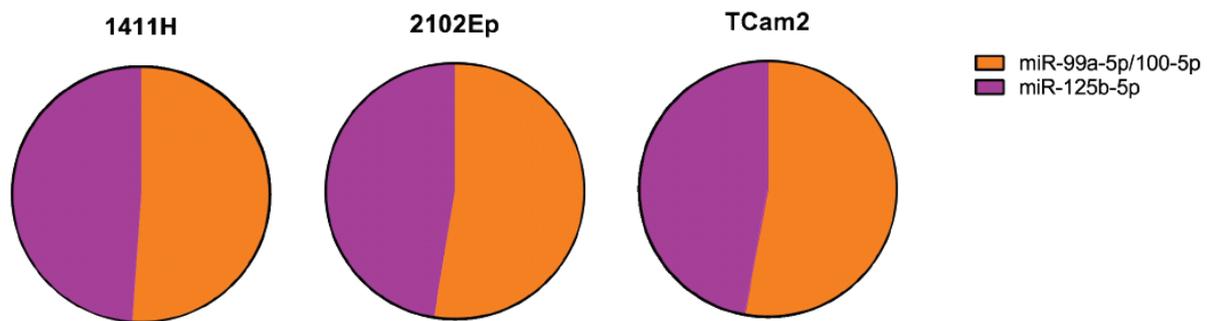
Figure 5.1: **Expression of each downregulated miRNA in 3 representative mGCTs cell lines.** In purple is the contribution derived from miR-125b-5p, while in orange is the contribution derived from miR-99a-5p/-100-5p.

The data shows that miR-99a-5p/-100-5p and miR-125-5p are expressed in equal levels in 1411H cells (miR-99a-5p/-100-5p: 51%, miR-125b-5p: 49%), whereas, in 2102Ep and TCam2

cells, miR-99a-5p/-100-5p have a slightly higher expression (52.6% and 52.8%, respectively) compared with miR-125b-5p (47.4% and 47.2%, respectively).

5.2.2 **Replenishment of miRNA with synthetic mimics**

A common approach to determine the effects of a downregulated miRNA is to reintroduce it and analyse the consequence of this renewed expression. Such experiments give insight on the pathways the miRNA is involved in, this may vary in different contexts, as a single miRNA can target numerous transcripts for degradation and the specific target have been shown to vary in different tissues (Preusse, Theis and Mueller, 2016). To further study the impact that these miRNAs have on mGCTs progression, the reintroduction of the miR-99a-5p/-100-5p and miR-125b-5p in representative mGCT cell lines was achieved by transfecting the cells with synthetic miRNA mimics. As miR-99a-5p and miR-100-5p mature sequences are almost identical (there is only a nucleotide mismatch and not in the seed region), it is impossible to distinguish one from the other by qPCR methods (as shown in Chapter 3.2.1.1). Therefore, a single mimic was utilised to replenish the members of this miRNA family, namely the miR-100-5p mimic. Furthermore, data presented in Chapter 3 points strongly to the fact that both families of miRNA are co-regulated in these tumours, and together with the fact that miR-99a-5p/-100-5p have equal contribution in the overall downregulation as miR-125b-5p, the reintroduction of both miRNA mimics at the same time (called Combo in the experimental setup) was studied primarily, as well as separately.

5.2.2.1 **Optimisation of miRNA mimics transfection**

To optimise the experimental conditions of the reintroduction of miR-100-5p and miR-125b-5p, transfection with different concentrations of combined miR-100-5p and miR-125b-5p mimics were tested. Furthermore, single mimics were also used to quantify any additive or synergistic effects. This work was initially performed on three mGCT cell lines; 2102Ep (originally derived from an embryonal carcinoma or EC), 1411H (originally derived from a yolk sac tumour or YST) and TCam2 (originally derived from a seminoma or Sem).

For each cell line, therefore three different concentrations of mimics were tested both singly (miR-100-5p or miR-125b-5p) or in combination (miR-100-5p + miR-125b-5p - identified from now on throughout the text as Combo) (Table 5.1).

Transfection	16.7 nM total concentration	33.3 nM total concentration	66.6 nM total concentration
Single mimic (either miR-100-5p or miR-125b-5p)	16.7 nM	33.3 nM	66.6 nM
Combo (miR-100-5p + miR-125b-5p)	8.3 nM + 8.3 nM	16.7 nM + 16.7 nM	33.3 nM + 33.3 nM
MNC (Mimic Non-targeting Control)	16.7 nM	33.3 nM	66.6 nM

Table 5.1: **Detailed description of the conditions used for mimics' transfection.**

Cells were seeded at different densities ($1x10^5$ cells/well for 1411H and TCam2 cells, and $7.5x10^5$ cells/well for 2012Ep cells, all in technical triplicate in a 6 well plate) to compensate for their growth rates, so that all would reach the same level of confluence by day 7.

Cells were transfected with the relevant mimic singly or in combination at day 0 (d0) and then counted on day 7 (d7) to assess any phenotypic effect on cell growth. In both 1411H and TCam2 cells a reduction of the total number of cells as compared with the Mimic Non-targeting Control (MNC) was seen at day 7 after transfection of the Combo at a final concentration of 16.7 nM and 33.3 nM (Figure 5.2). 2102Ep cells were transfected with Combo combinations up to 66.6 nM, however no significant effect on cell growth was seen in any of the tests (data shown in the Appendix 1).
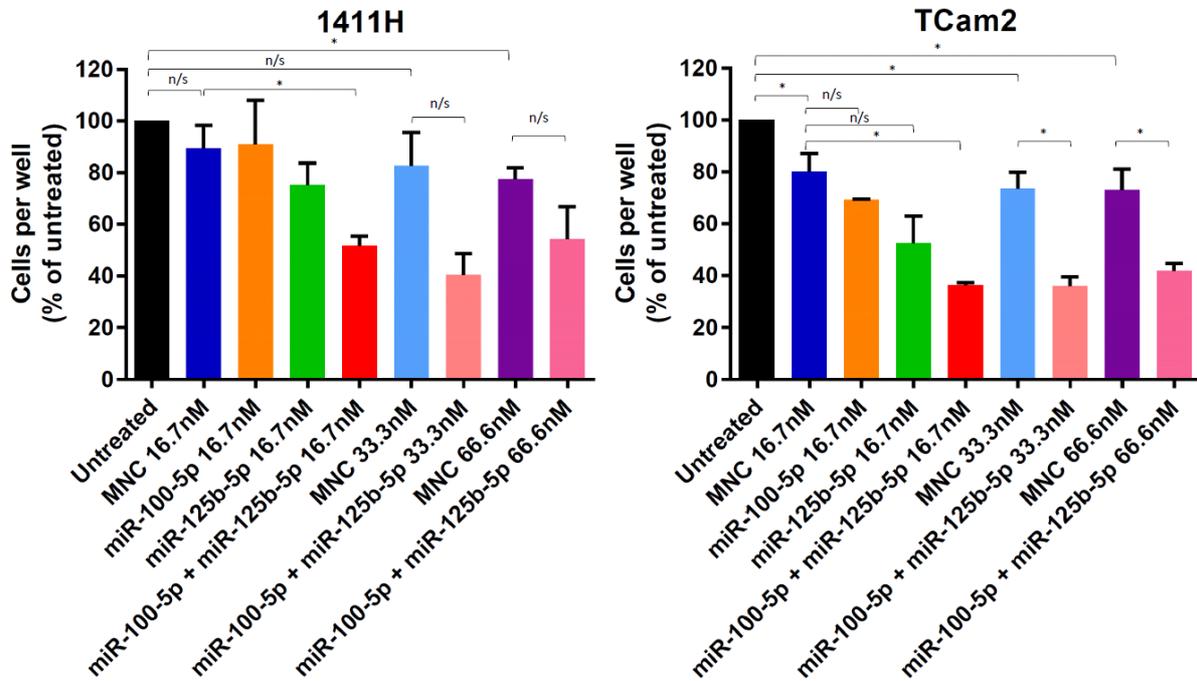
Figure 5.2: **Average cell number per well at day 7 after transfection of mimics in 1411H and TCam2.** Cell number is expressed as percentage of the untreated control. Transfection using a single mimic was performed only at 16.7 nM, whereas MNC treated cells and cell receiving a combination of mimics (Combo) were transfected at 16.7 nM, 33.3 nM and 66.6 nM. The MNC matched the total amount of mimic used in the Combo. Data presented as mean ± SEM of n=3 biological replicates. * $p<0.05$, **$p<0.01$, *** $p<0.005$ and **** $p<0.001$. For statistical analysis, Student's t-test was used to calculate significance.

Representative images of 1411H and TCam2 cells at day 7 after transfection are shown in Figure 5.3, illustrating the results reported in Figure 5.2. Transfection with MNC at the lowest dose of mimic tested (i.e. 16.7 nM) indicated that the transfection itself had some intrinsic effect on cell viability, as the average number of cells in MNC treated wells decreased to almost 80% of the untreated wells in TCam2. While the reduction of cell number following transfection with MNC was statistically significantly for TCam2 (*p*-value: 0.05), changes in the number of 1411H cells did not reach statistical significance compared with the untreated control. This difference is likely due to the increased sensitivity of TCam2 cells to the lipofectamine reagent used for the transfection. There was no significant decrease in cell numbers in 1411H cells transfected with a single 16.7 nM miRNA mimic as compared with

cells transfected with the MNC. In TCam2, while transfection with miR-100-5p did not result in a decrease in the total number of cells at day 7, replenishing miR-125b-5p had an effect on cell proliferation, but this was not statistically significant. The transfection of both mimics (miR-100-5p + miR-125b-5p) at a combined concentration of 16.7 nM resulted in a significant decrease in cell number compared with the MNC treated cells in both cell lines (*p*-value*:* 0.008 and 0.01 for 1411H and TCam2, respectively). No additional phenotypic effect on cell growth was seen using the combination of mimics Combo at higher doses, therefore 16.7 nM was chosen as the optimal concentration of mimics to be used for subsequent replenishing experiments.
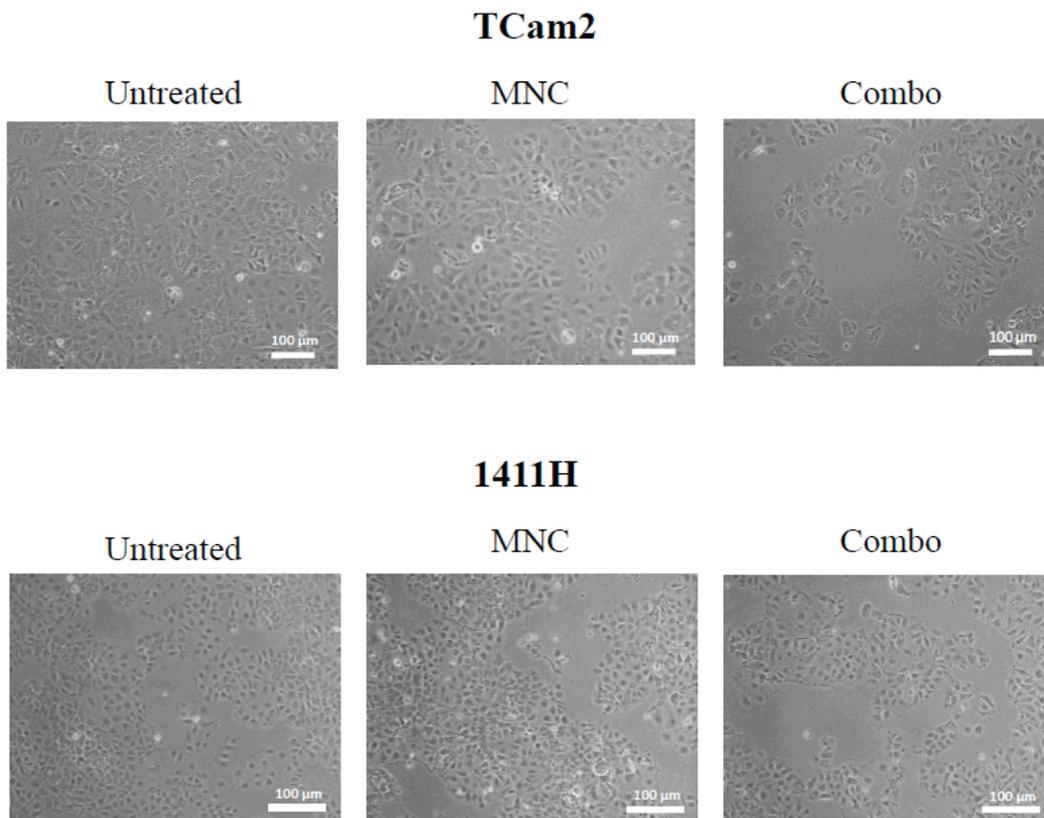
Figure 5.3: **Images of 1411H and TCam2 cell lines taken at day7**. Untreated cells were grown for 7 days in standard conditions while MNC or the Combo (miR-100-5p + miR-125b-5p) were transfected with 16.7 nM of the appropriate mimic/combination.

5.2.3 **Effects of miRNA miR-100-5p and miR-125b-5p replenishment**

5.2.3.1 **Characterisation of cells where miR-100-5p and miR-125b-5p are replenished**

Having shown that replenishing miR-100-5p and miR-125b-5p in combination resulted in a statistically significant decrease in the number of cells per well after seven days, a detailed assessment of the growth of the cells after transfection of the mimics was carried out. As no significant phenotypic effect was observed in cells transfected with the mimic individually, only the combination of mimics was utilised for these experiments.

Three cell lines, 1411H, TCam2 and 2102Ep, were transfected with MNC or the Combo at a final concentration of 16.7 nM and their growth was followed at d2, d4 and d7 after transfection (Figure 5.4 panels a, c and e for 1411H, TCam2 and 2102Ep, respectively). To inform on the level of the replenishment and its kinetics, intracellular miR-100-5p and miR-125b-5p were quantified at d2, d4 and d7 after transfection (Figure 5.4 panels b, d and f). Although replenishment in 2102Ep cells did not give rise to a phenotypic effect on growth, even at the highest concentrations of mimics used (66.6 nM), this cell line was included in this experiments to further characterise this apparent lack of response.
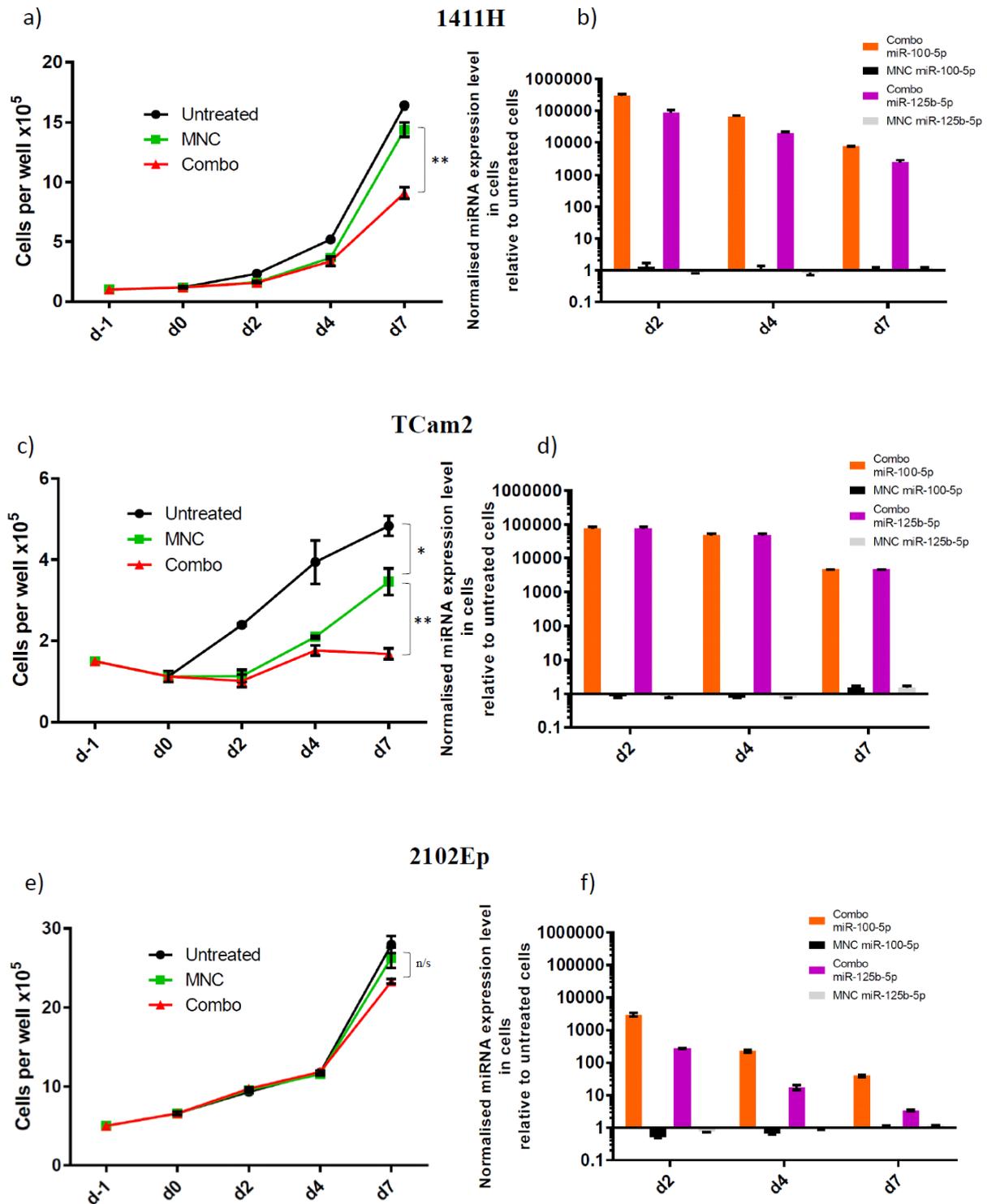
Figure 5.4: **Growth curves over 7 days of 1411H (panel a), TCam2 (panel c), and 2102Ep (panel e).** Growth was assessed after replenishing miR-100-5p and miR-125b-5p, and the levels of intracellular mimics found at the same timepoints (1411H, TCam2 and 2102Ep in panel b, panel d and panel f

respectively). Data presented as mean ± SEM of n=3 biological replicates. * p<0.05, **p<0.01, *** p<0.005 and **** p<0.001. For statistical analysis, Student's t-test was used to calculate significance.

Consistently with the optimisation experiment, these experiments showed that transfection resulted in a sight, but not statistically significant reduction in the viability of the cells, with the exception of TCam2 cells where the effect was statistically significant at day 7 (Figure 5.4, panel c). Replenishing miR-100-5p and miR-125b-5p in TCam2 and 1411H cell lines resulted in a statistically significant decrease in the number of cells at day 7, as previously observed, whereas no difference was seen again for 2102Ep cells (Figure 5.4, panel e). Interestingly, the growth effect could only be seen at day 7, in 1411H cells, suggesting that only a minority of cells are affected by the treatment, whereas in TCam2 cells a reduction in total cell number per well can be seen also at day 4. This is probably due to the toxic effect of the lipofectamine reagent, as the number of cells in MNC and Combo were the same. The level of replenishment of the mimics in 1411H compared with untreated control cells (Figure 5.4, panel b), appears to be higher at day 2 than in TCam2 (Figure 5.4, panel d), but no changes were seen in the phenotypic effect among these cell lines. This is in line with the lower level of pri-miR-100-5p and pri-miR-125b-5p seen in 1411H cells compared with TCam2 as shown in Chapter 3 section 3.2.4.2, which resulted in lower levels of the mature form of these miRNAs (Chapter 3 section 3.2.5.1). Consistently with the absence of a phenotypic effect on cell growth, in 2012Ep cells the replenishment of the miRNAs was less efficient and levels of mimics at 2 days post transfection were 100 to 200-fold lower than in the other two cell lines. Moreover, the level of intracellular mimics appeared to decrease more quickly in 2012Ep cells, suggesting that perhaps the turnover of the mimics is faster in these cells.

All together this data further confirmed the initial observation that replenishing miR-100-5p and miR-125b-5p results in a reduced growth in both 1411H and TCam2 cell lines. It also suggests that these effects are proportional to the level of mimics reintroduced in the cell line, as 2012Ep cells, where the levels of replenished were much lower, show no phenotypic effect on cell growth.

5.2.3.2 **Identification of expression changes by microarray**

Having shown that replenishing miR-100-5p and miR-125b-5p in combination resulted in a statistically significant decrease in the number of cells after seven days, we then investigated whether this phenotypic effect was associated with a change in gene expression. To start with, gene expression changes were investigated at day 7 post replenishment of the miRNA, to help identify the pathways that had been affected, rather than the individual genes.

To this aim, three cell lines, 1411H, TCam2 and 2102Ep, were transfected with MNC or the Combo at a final concentration of 16.7 nM and cells were collected at day 7, followed by RNA extraction. Quality and concentration of RNA extracted from cells was assessed with SpectroStar (BMG Labtech, Aylesbury, UK) and a Bioanalyser (Agilent Technologies, Cheadle, UK). Matching replicates were chosen on the basis of their quality and further used for gene expression microarray experiments. Library preparation and the hybridisation to microarrays, HumanHT-12 v4 Expression BeadChip (Illumina, Chesterford, UK), as well as data collection were performed at Cambridge Genomic Services, University of Cambridge. Ultimately, arrays using material obtained from treatment of TCam2 and 1411H cell lines were run with 3 biological replicates of untreated cells (Untreated), 3 biological replicates of cells treated with the Mimic Non-targeting Control (MNC) and 3 biological replicates of cells treated with miR-100-5p and miR-125b-5p together (Combo). For the 2102Ep cell line only 3 biological replicates of MNC and Combo were used (no untreated cells).

All raw data obtained passed quality control checks and was put through the standard pipeline of data processing and normalisation. Initially, the normalised microarray expression data were utilised to generate volcano plots for all samples. A volcano plot (Figure 5.5 panels a, c and e) is used to quickly identify changes in large data sets composed of replicate data. It combines a measure of statistical significance from a statistical test with the magnitude of the change, enabling quick visual identification of those data-points (i.e. genes) that display large magnitude changes that are also statistically significant. Degree of significance can be made more obvious by choosing a different colour for genes with significance above a certain threshold, as in the data presented here where differential expression with an adjusted -log10(p-value)>0.05 is marked in red. As seen in Figure 5.5 a number of significantly differentially

expressed genes were identified at day 7 in both 1411H and TCam2 cells (221 and 2,938 differentially expressed genes, respectively). More genes were identified as significantly differentially expressed genes in TCam2 cells than in 1411H cells when MNC and Combo where compared, however more changes were also identified in the comparison between MNC and untreated samples for this cell line, perhaps suggesting a technical rather than a biological reason for this difference. Consistent with the lack of effects on proliferation seen in 2012Ep cells following miRNA replenishment, no significantly differentially expressed genes were identified by comparing expression in cells treated with either MNC or Combo.

1411H d7

a) Untreated vs MNC

b) MNC vs combo

TCam2 d7

c) Untreated vs MNC
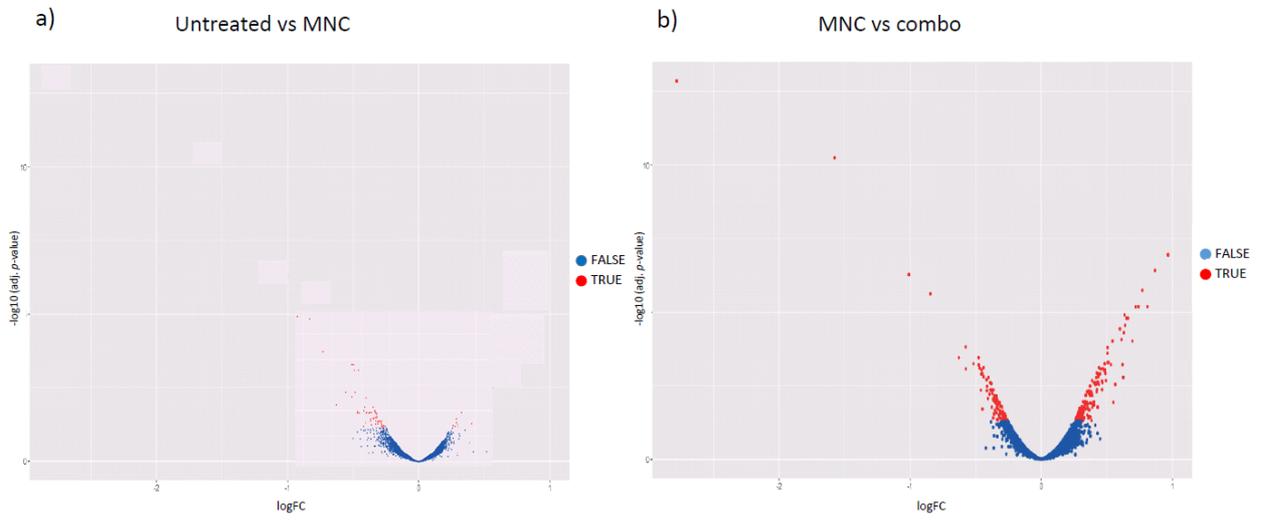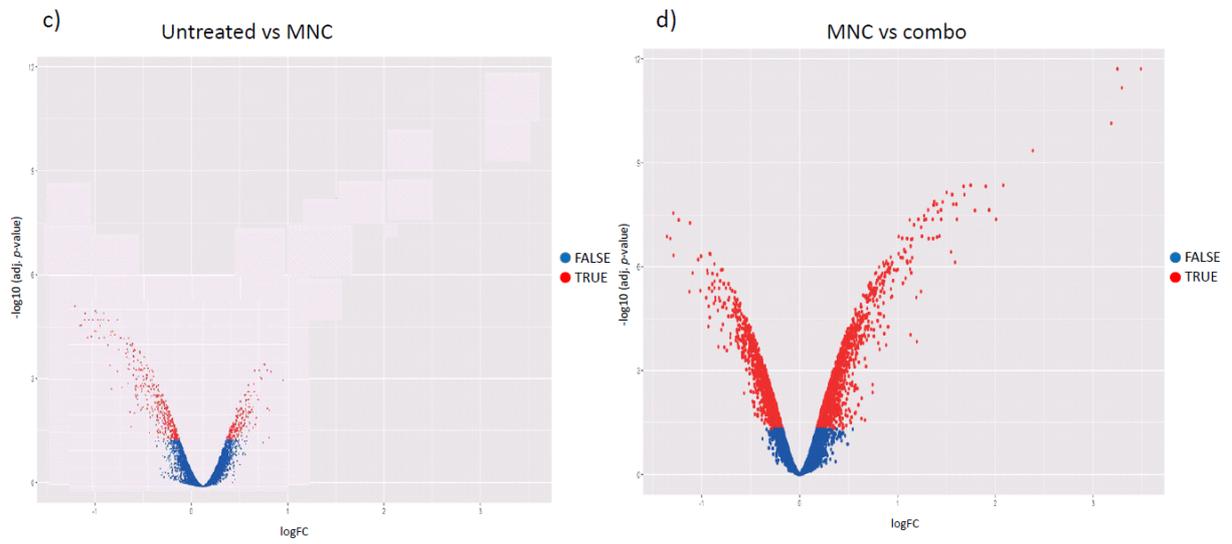
d) MNC vs combo

e)   2102Ep

MNC vs combo
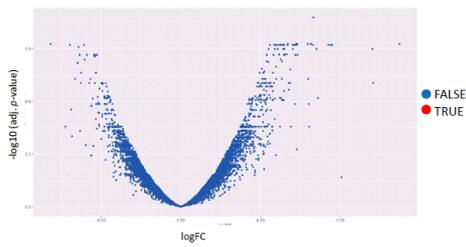
Figure 5.5: **Volcano plots showing differential gene expression**. a) 1411H cells untreated vs treated with MNC, b) 1411H cells treated with MNC vs treated with Combo, c) TCam2 cells untreated vs treated with MNC, d) TCam2 cells treated with MNC vs treated with Combo, and e) 2102Ep cells treated with MNC vs treated with Combo. Combo is equimolar combination of mimics for miR-100-5p and miR-125b-5p. RNA samples were extracted 7 days after mimic transfection. Level of change (fold change in log2 scale) is shown on the x axis while the y axis denotes significance (p-value) in $\log_{10}$.

Subsequent hierarchical clustering analysis was carried out with the normalised microarray expression data. This type of clustering is based on the core idea of objects being more related to nearby objects than to objects further away and are therefore based on distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster. At different distances, different clusters will form, which can then be represented using a dendrogram, which illustrates the hierarchy of clusters that merge with each other. In a dendrogram, the y-axis marks the distance at which the clusters merge, while the samples are placed along the x-axis such that the clusters don't mix.

This analysis is useful to get an overview of the levels of changes and their significance (Figure 5.6). Two dendrograms per cell line were generated by this analysis, one containing the control untreated group (untreated) vs the sample treated with Mimic Non-targeting Control (MNC) and a second containing the MNC vs the Combo treated group. The untreated vs MNC control dendrogram showed some differences, particularly for TCam2 cells, however more obvious

differences were seen in both cell lines when the MNC was compared with the Combo treated group.

For the 2102Ep cell line only one dendrogram was produced, the one containing the MNC vs the Combo treated group. As expected, no obvious differences were visible between the samples from 2102Ep cells treated with the non-targeting control (MNC) vs the Combo treated group.
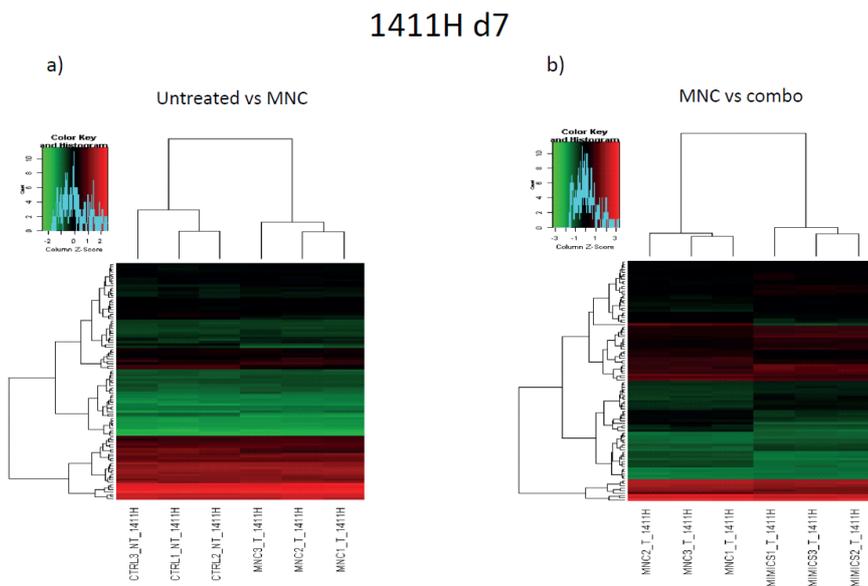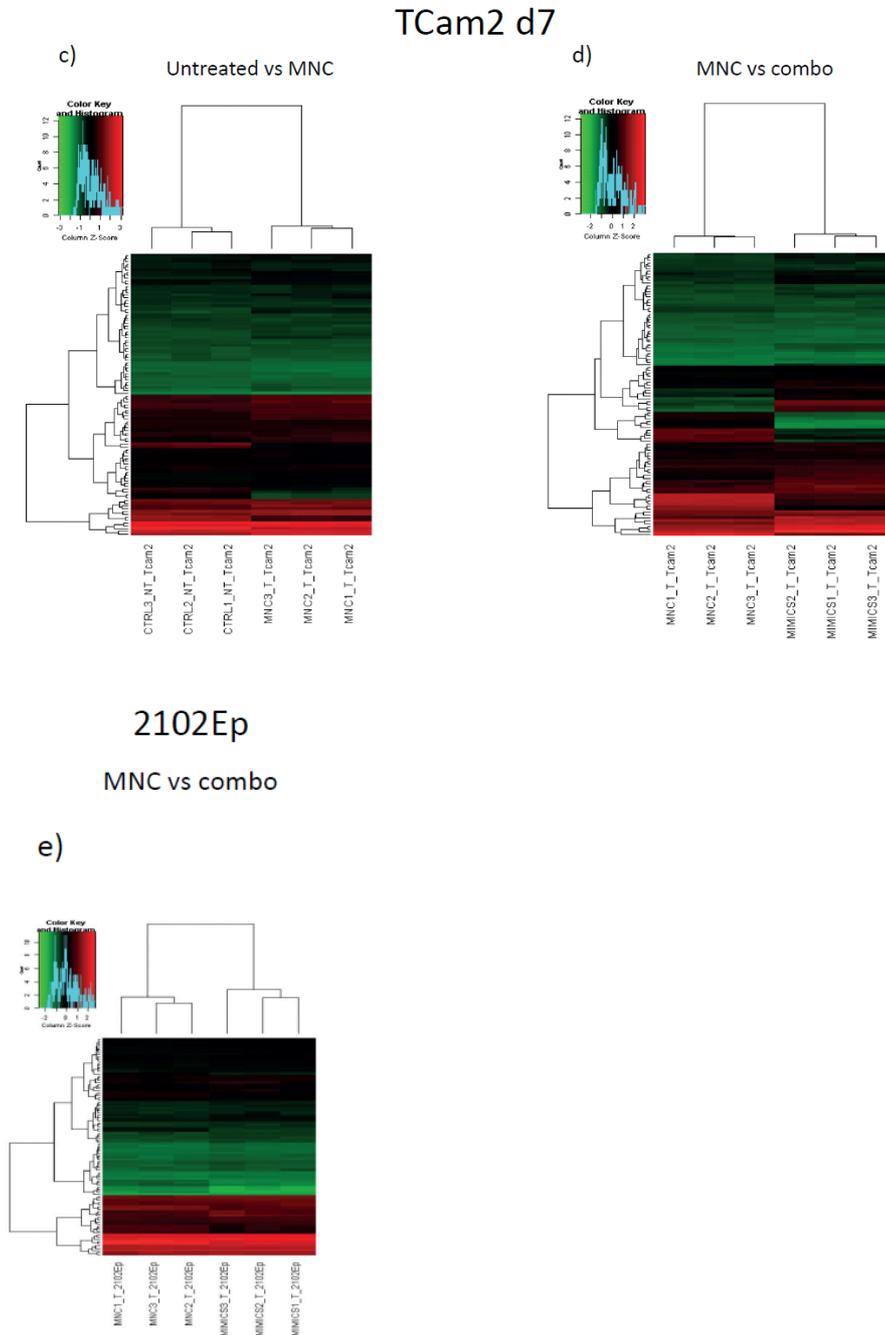
Figure 5.6: **Heatmaps showing hierarchical clustering**. a) 1411H cells untreated vs treated with MNC, b) 1411H cells treated with MNC vs treated with Combo, c) TCam2 cells untreated vs treated with MNC, d) TCam2 cells treated with MNC vs treated with Combo, and e) 2102Ep cells treated with MNC vs treated with Combo. Combo is equimolar combination of mimics for miR-100-5p and miR-125b-5p.

RNA samples were extracted 7 days after mimic transfection. The heatmaps display level of expression for both groups in the comparison as a colour map for the top differentially expressed 50 genes.

5.2.4 Analysis of expression using Sylamer

Although the preliminary analysis gave us an indication that differential expression of genes was occurring after treatment with mimics replenishing miR-100-5p and miR-125b-5p compared with a MNC control, this analysis was not sufficient to give insight on the genes specifically targeted by the miRNAs. This is due to the fact that the primary targets of the miRNA are likely to themselves have further downstream targets which would be differentially expressed by day 7 after replenishment. Discriminating between these two classes of genes (i.e. directly affected vs indirectly affected) is impossible with conventional analysis.

Therefore, a different type of analysis, using a programme called *Sylamer* was performed. *Sylamer* is a system for finding significantly over or under-represented 'words' in sequences according to a sorted gene list. An example of such a 'word' could be the complementary region to the seed sequence of a miRNA (Seed Complementary Region or SCR). *Sylamer* is hence typically used to find significant enrichment or depletion of microRNA or siRNA seed sequences from microarray expression data by selectively looking for the sequence complementary to the miRNA seed region in the 3'UTR region of the mRNA transcript. Results are plotted in terms of a significance landscape plot, which displays significance profiles for each individual all 'word' studied across the sorted genelist. The most significant 'words' are highlighted, alternatively specific 'words' can also be looked at. In a successful replenishment experiment where the mimic has had an effect on its targets, the replenished miRNA would be expected to be pulled out as most significant. Importantly, the programme can also be used to produce a list of all the genes differentially expressed that contain such 'word', ranked according to how significant the differential expression is. This would allow the identification of the genes directly affected by the replenishment of miR-100-5p and miR-125b-5p and, consequently, the identification of pathways deregulated due to the under-expression of these miRNAs.

Analysis with *Sylamer* was carried out on mRNA gene lists ranked according to differential gene expression obtained from the comparisons between untreated samples vs MNC and MNC vs Combo treated cells. Significance landscape plots for 1411H, TCam2 and 2102Ep cells produced from differential gene lists from MNC vs Combo treated cells are shown in Figure 5.7, 5.8 and 5.9, respectively.

These *Sylamer* plots, constructed from day 7 post replenishment expression data, consider 'words' corresponding to six SCR sequences relative to miR-99a-5p/-100-5p and miR-125b-5p, namely sequences corresponding to three seed hexamers (1-6 nt, 2-7 nt, 3-8 nt), two heptamers (1-7 nt, 2-8 nt), and one octamer (1-8 nt) for each miRNA. Both miRNAs are represented on the plot, and in total six hexamers, four heptamers and 2 octamers curves are shown per graph per cell line.

a)
hexamers

Selected words
ACGGGT (miR-99a-5p/-100-5p)
CGGGTA (miR-99a-5p/-100-5p)
TACGGG (miR-99a-5p/-100-5p)
CTCAGG (miR-125b-5p)
TCAGG (miR-125b-5p)
CAGGGA (miR-125b-5p)

b)
heptamers

Selected words
ACGGGTA (miR-99a-5p/-100-5p)
TACGGGT (miR-99a-5p/-100-5p)
CTCAGGG (miR-125b-5p)
TCAGGGA (miR-125b-5p)

c)
octamers

Selected words
TACGGGTA (miR-99a-5p/-100-5p)
CTCAGGGA (miR-125b-5p)

Figure 5.7: **Sylamer plots derived from analysis of differential expression data for 1411H treated with MNC vs Combo.** a) significance landscape plots were produced searching for 6-mers (a), 7-mers (b) and 8-mers (c). Highlighted are the significance curves for the seed complementary regions (SCR) for miR-99a-5p/-100-5p and miR-125b-5p. The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated. The y-axis represents the SCR abundance in the 3'UTR of the genes.

a) hexamers

b) heptamers

c) optamers
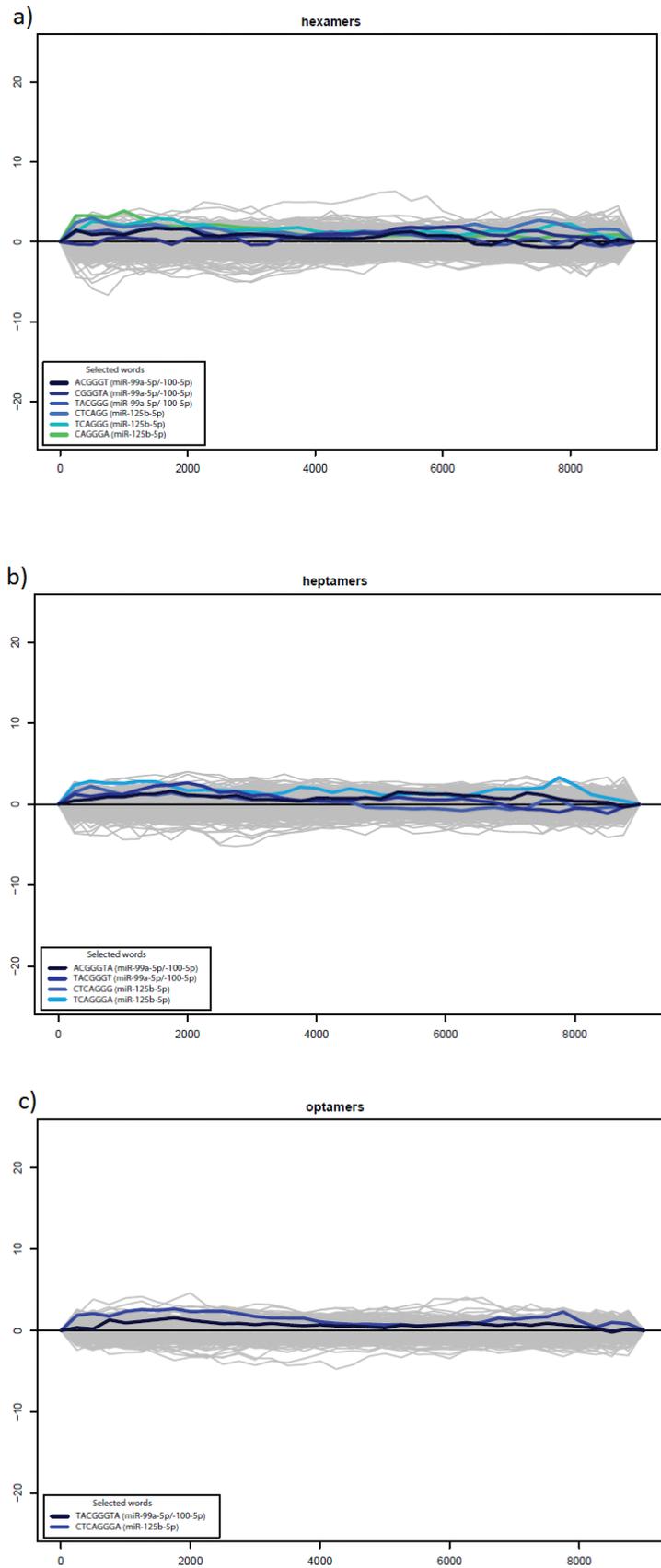
Figure 5.8: **Sylamer plots derived from analysis of differential expression data for TCam2 treated with MNC vs Combo.** a) significance landscape plots were produced searching for 6-mers (a), 7-mers (b) and 8-mers (c). Highlighted are the significance curves for the seed complementary regions (SCR) for miR-99a-5p/-100-5p and miR-125b-5p. The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated. The y-axis represents the SCR abundance in the 3'UTR of the genes.

a) hexamers

Selected words
ACGGGT (miR-99a-5p/-100-5p)
CGGGTA (miR-99a-5p/-100-5p)
TACGGG (miR-99a-5p/-100-5p)
CTCAGG (miR-125b-5p)
TCAGGG (miR-125b-5p)
CAGGGA (miR-125b-5p)

b) heptamers

Selected words
ACGGGTA (miR-99a-5p/-100-5p)
TACGGGT (miR-99a-5p/-100-5p)
CTCAGGG (miR-125b-5p)
TCAGGGA (miR-125b-5p)

c) octamers

Selected words
TACGGGTA (miR-99a-5p/-100-5p)
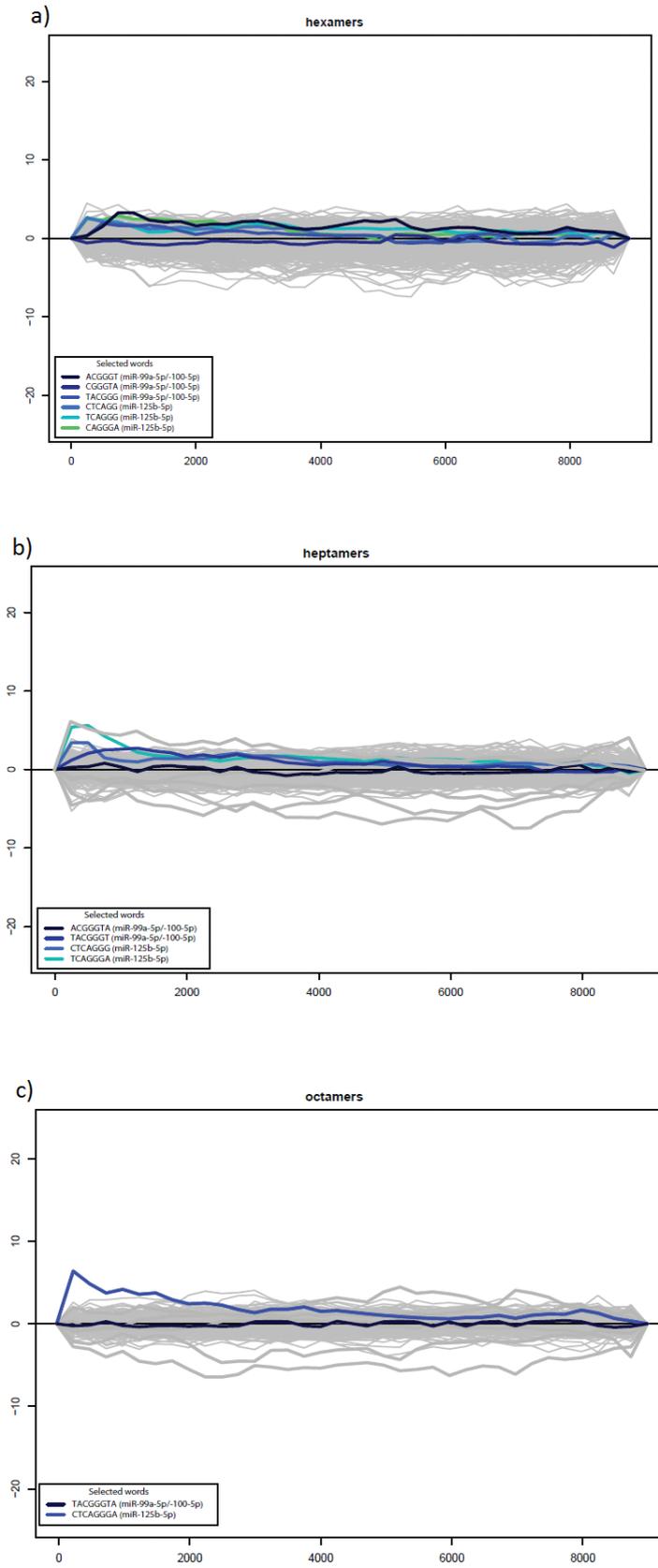CTCAGGGA (miR-125b-5p)
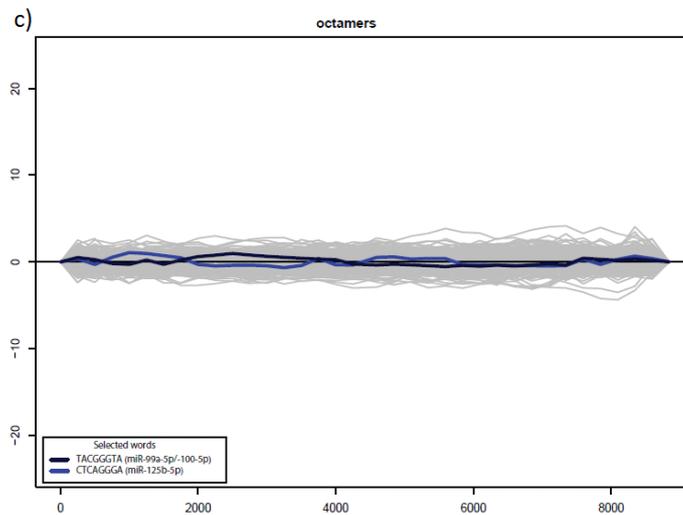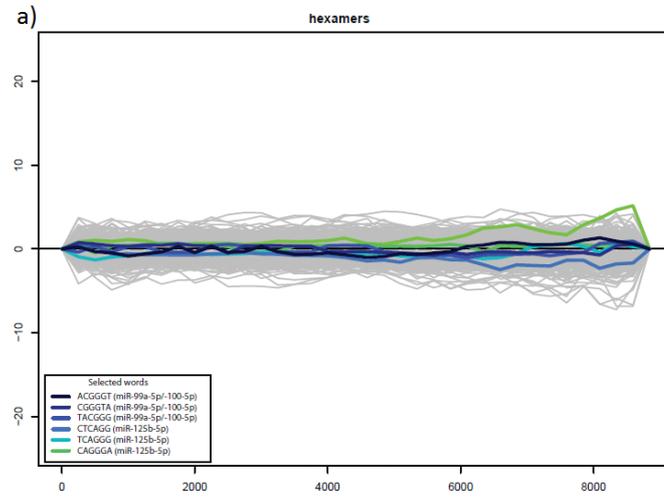
Figure 5.9: **Sylamer plots derived from analysis of differential expression data for 2102Ep treated with MNC vs Combo.** a) significance landscape plots were produced searching for 6-mers (a), 7-mers (b) and 8-mers (c). Highlighted are the significance curves for the seed complementary regions (SCR) for miR-99a-5p/-100-5p and miR-125b-5p. The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated. The y-axis represents the SCR abundance in the 3'UTR of the genes.

This analysis showed inconsistent results for the various cell lines and seeds. This is unusual for true differentially expressed miRNAs where the SCR element is expected to be the most significant. Moreover, plots of overlapping SCR elements typically follow a coordinated pattern (van Dongen, Abreu-Goodger and Enright, 2008c), which was not seen in this instance. The effects of all SCR elements of interest (i.e. six hexamers, four heptamers and 2 octamers) were therefore considered simultaneously (in a process called Single-Sum Significance Sylamer analysis or SSS) and the graphical representation of the data is shown in Figure 5.10 panel a for 1411H, panel b for TCam2 and panel c for 2102Ep. For this analysis genes were grouped together in bins of 200 genes ranked from the most down-regulated (to the left) to the most up-regulated (to the right). An arbitrary cut off was chosen depending on the mean of the two peaks generated for miR-100-5p and miR-125b-5p.

a)

1411H d7 Single−Sum Significance Sylamer

b)

TCam2 d7 Single−Sum Significance Sylamer
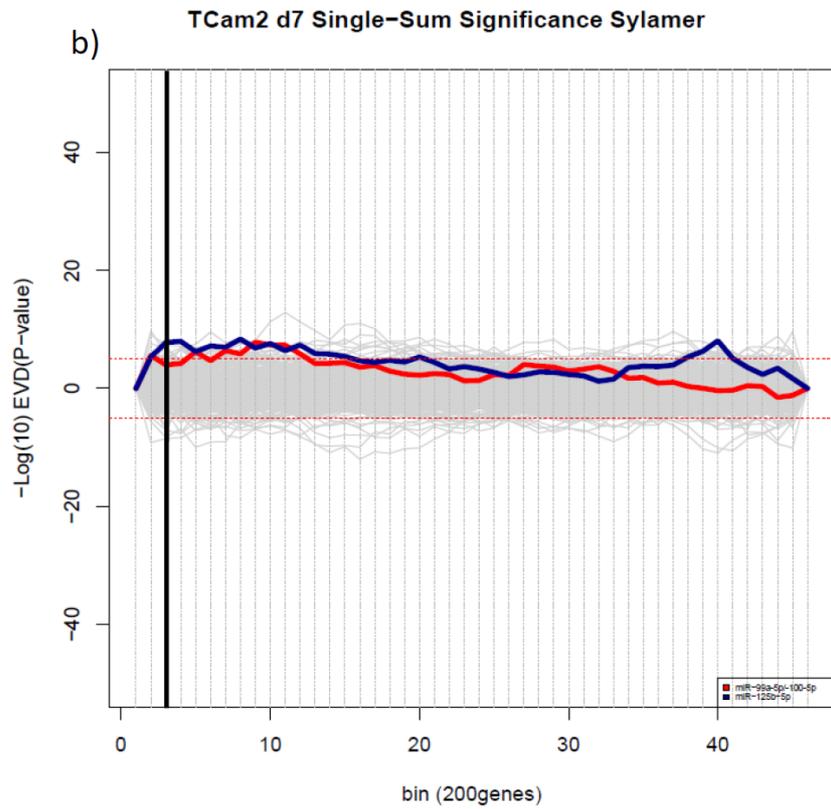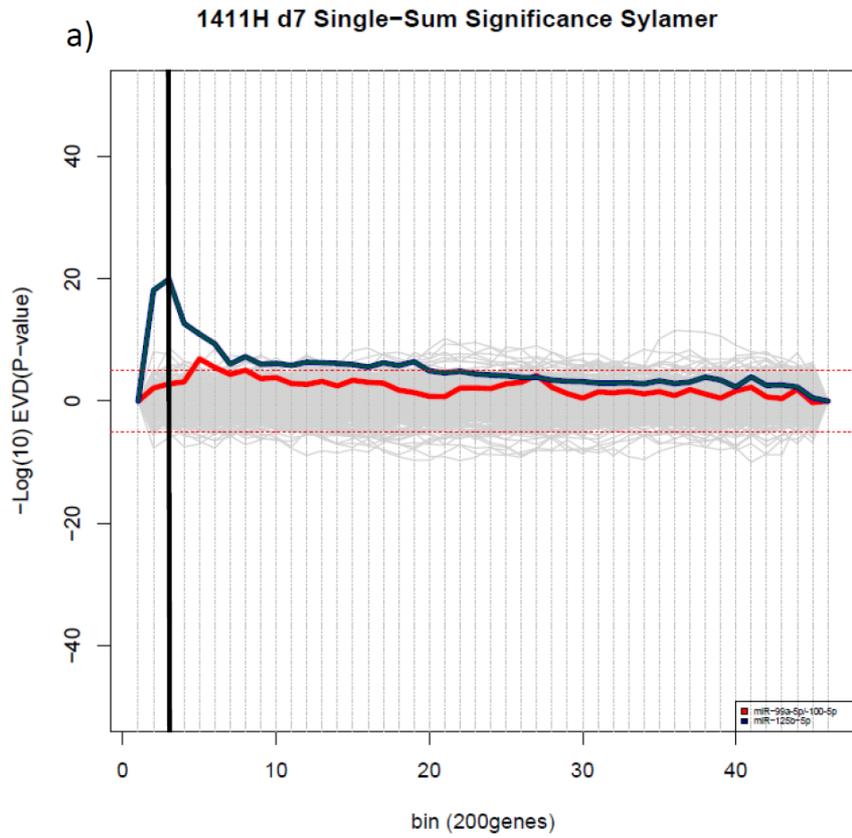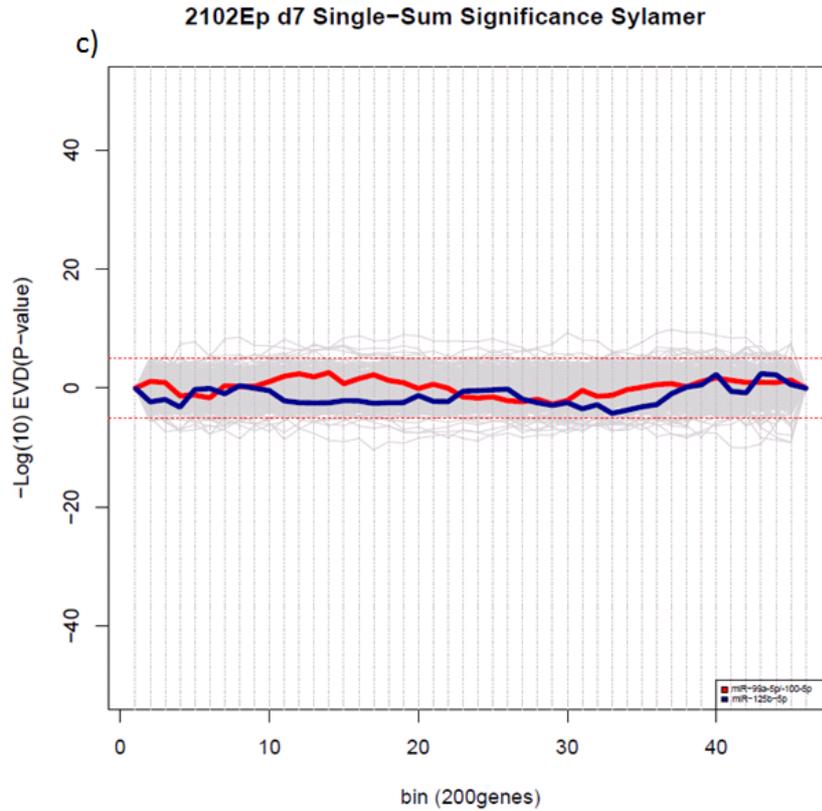
Figure 5.10: **SSS plots derived from analysis of differential expression data.** The graphs show (a) 1411H, (b) TCam2 and (c) 2102Ep cell lines treated with MNC vs Combo. Highlighted are the significance curves for miR-99a-5p/-100-5p (red line) and miR-125b-5p (blue line). The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated, each bin corresponds to 200 genes. The black vertical line corresp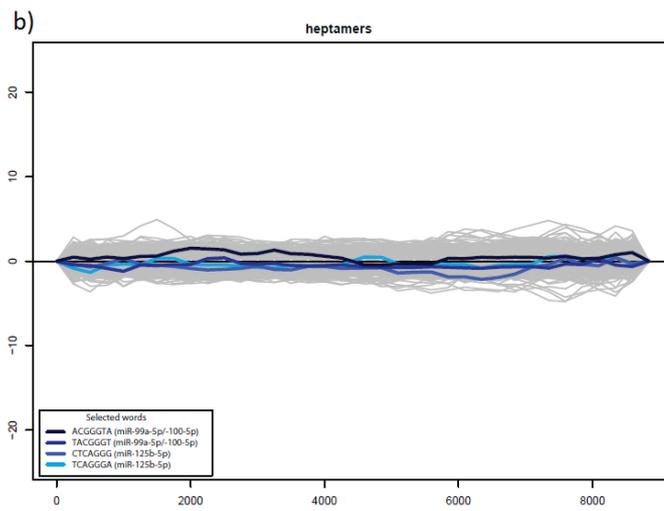onds to the assigned cut-off and the circle on its left underlines the genes that were used for further pathway analysis.

5.2.5 Pathway analysis following miRNA replenishment

From the Significant Single-Sum Sylamer plots, an appropriate cut off for each cell line was determined, and the most differentially expressed genes identified as containing the miR-99a-5p/-100-5p and miR-125b-5p seed complementary region above the threshold were utilised to determine the pathways regulated by these miRNAs. 400 genes were identified in 1411H cells and 400 genes were identified in TCam2 cells as for both cell lines the threshold was set at the second 200 genes bin in the Single-Sum Significance Sylamer graph. As well as these genes,

pathways analysis of the most differentially expressed genes as determined by the initial analysis of the microarray data were also carried out to investigate the downstream effect of the miRNA targeted genes. Genes differentially expressed in TCam2 treated cells showing 0.5 fold or larger expression changes and with a minimum p-value of 0.05 were included in this analysis. For genes differentially expressed in 1411H treated cells a 0.3 fold or larger expression change, with a minimum p-value of 0.05 were considered. Different thresholds were set in the two cell lines to ensure that a sufficient number of genes were included in the analysis. This was required as a lower number of gene changes were seen in the experiments with 1411H cells as shown in Figure 5.5 panel b) and panel d). A single gene list containing both the genes identified by the *Sylame*r analysis and the genes identified by the differential expression analysis was generated. Gene duplication, due to genes being identified in both analyses, was eliminated before further pathway analysis was performed. As no significant enrichment of genes containing miR-99a-5p/-100-5p and miR-125b-5p complementary seed regions was found by *Sylamer* analysis and very few differentially expressed genes were identified for 2102Ep cells, no further analysis was carried out for these samples.

Pathway analysis using KEGG (Kyoto Encyclopedia of Genes and Genomes), gene onthology (GO) and reactome were carried out. The analysis was executed as follows. Firstly, each database was used individually and only pathways enriched for three or more genes were taken into account. Secondly, these resultant pathways found were compared among the three different bioinformatical systems. Finally, only the pathways found to be commonly enriched among the three approaches were further investigated. Significance of a particular pathway was then utilised to graphically represent the results obtained for 1411H (Figure 5.11 panel a) and TCam2 (Figure 5.11 panel b) cells.

a)

1411H d7

b)

TCam2 d7
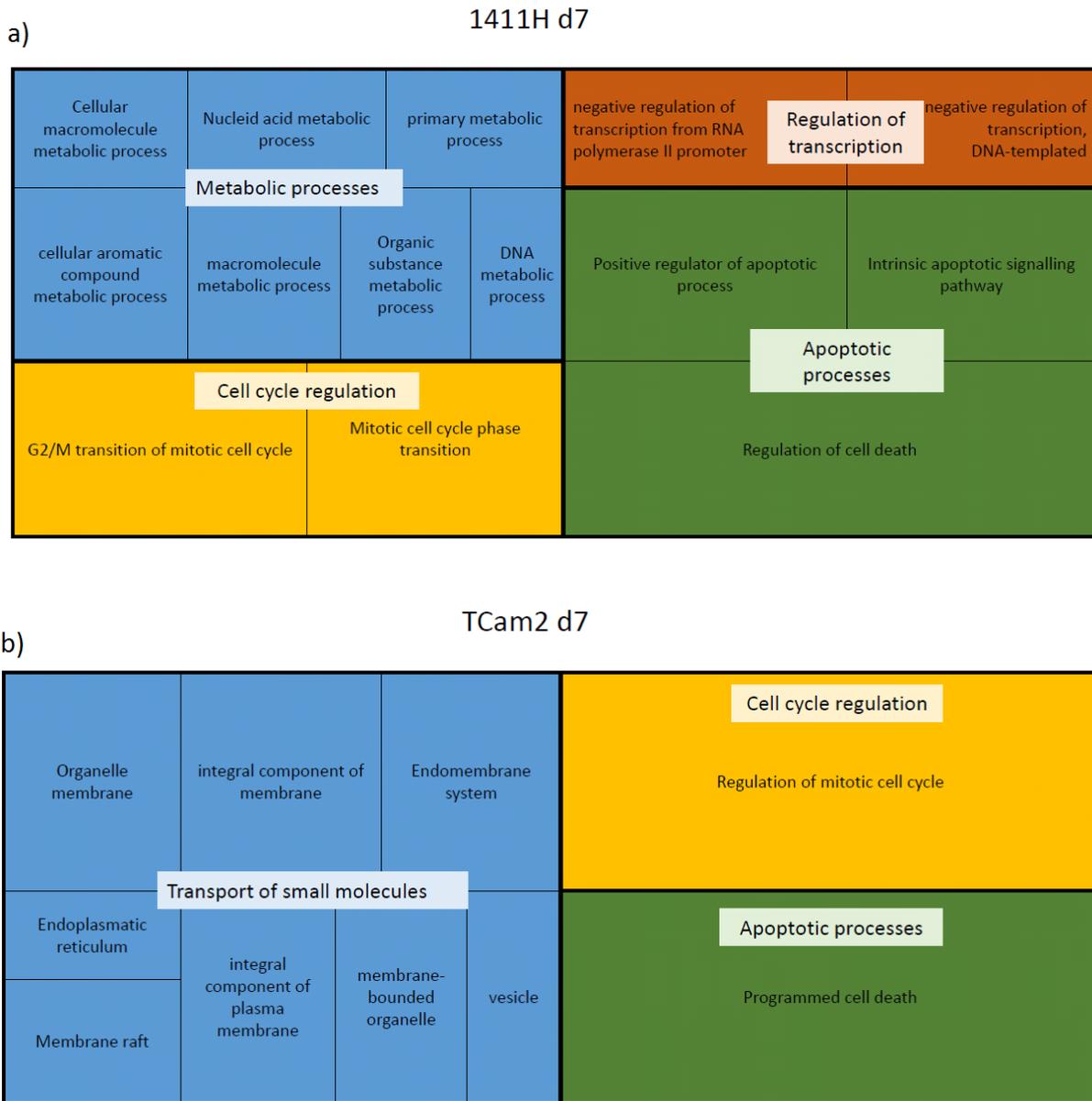
Figure 5.11: **Pathway analysis.** The most significant pathways affected by the replenishment of miR-100-5p and miR-125b-5p and identified in 1411H (panel a) and TCam2 (panel b) cells are shown. The size of each quadrant is relative to the significance of the pathway analysed.

The analysis of the pathways affected at day 7 post miRNA replenishment illustrates the potential mechanisms involved in the phenotypic effect seen in 1411H and TCam2 cells. For both cell lines a quarter of the genes differentially expressed at day 7 following replenishment of the miRNA code for proteins involved in the regulation of cell cycle and apoptotic process which is consistent with the observed phenotypic changes on cell growth seen. However, the vast majority of differentially expressed genes appeared to be involved in cell line specific pathways. For instance, a high percentage of genes are involved in metabolic processes in 1411H cells, while in TCam2 the most deregulated genes were involved in the transport of small molecules. Moreover, in the TCam2 cell line, genes involved in the PIK3-AKT pathway were highly represented (p-value: 0.0000129) in the list of most downregulated differentially expressed genes.

Ultimately, this data outlines the common changes in the regulation of cell cycle occurring within both 1411H and Tcam2 cell lines that are likely to be responsible for the phenotypic effects observed. Due to the time elapsed from the replenishment, direct targets affected by the mimics showed changes of low significance and magnitude, which would be difficult to validate experimentally by RT-qPCR and more importantly, by Western blot. Therefore, further studies aimed to assess earlier events were then performed.

5.2.6 Differential gene expression at earlier times after miRNA replenishment

Having shown (Figure 5.4) a phenotypic effect after replenishing miR-100-5p and miR-125b-5p in mGCTs cell lines and seeing an association between the changes in gene expression detected 7 days post replenishment and the regulation of the cell cycle, earlier changes in gene expression were studied. It has been shown by other groups that changes in the regulation of miRNA targeted transcripts are often seen more significantly two days post replenishment (Bhutra *et al.*, 2014). For this reason, differential gene expression was investigated by microarray with samples collected two days after replenishment with either a mimic non-targeting control or the combination of miR-100-5p and miR-125b-5p mimics (Combo) in the same conditions used previously. The experiments were carried out in 1411H and TCam2 cell lines alone, as the cell line 2102Ep cell line was excluded from these investigations due to the

low differential effects seen seven days post replenishment and the lack of a detectable phenotypic effect. As the differences seen between untreated cells and cells treated with MNC at day 7 did not warrant the inclusion of these samples, untreated controls were also not included in these further investigations.

All raw data obtained passed quality control checks and were put through the standard pipeline of data processing and normalisation, as previously described. The normalised microarray expression data were then utilised to generate volcano plots for all samples (Figure 5.12). A number of significantly differentially expressed genes were identified at day 2 in both 1411H and TCam2 cells. As seen at day 7, more genes were identified as significantly differentially expressed in TCam2 cells (1,738 genes at p-value: 0.05) cells than in 1411H cells (99 genes at p-value: 0.05) cells when MNC and Combo where compared. Hierarchical clustering analysis was also performed and heatmaps obtained are shown in Figure 5.13. One dendrogram per cell line was generated by this analysis, containing the samples treated with Mimic Non-targeting Control (MNC) in three biological replicates vs the Combo treated group, again in three biological replicates. As per day 7, obvious differences were seen in both cell lines at day 2 post treatment, which are consistent with a mimic mediated effect on the cells at this early timepoint. These analyses gave an overview of the levels of changes and their significance, but as per the data obtained at day 7, it was important to determine the specific targets modulated by the mimics, as secondary targets (i.e. genes regulated by the miRNAs target genes themselves) are potentially differentially modulated even at the early timepoint.
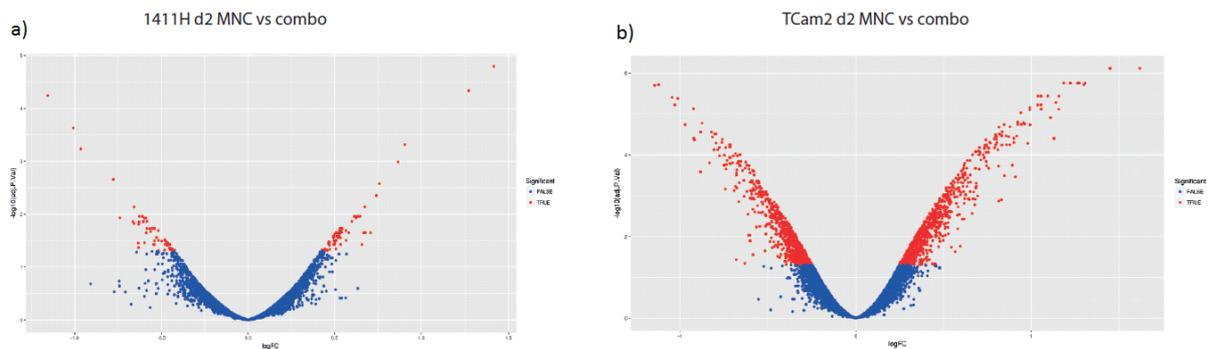
Figure 5.12: **Volcano plots showing differential gene expression**. The graphs show in a) 1411H cells treated with MNC vs treated with Combo, and in b) TCam2 cells treated with MNC vs treated with Combo. Combo is equimolar combination of mimics for miR-100-5p and miR-125b-5p. RNA samples were extracted 2 days after mimic transfection. Level of change (fold change in log2 scale) is shown on the x axis while the y axis denotes significance (p-value) in $\log_{10}$.
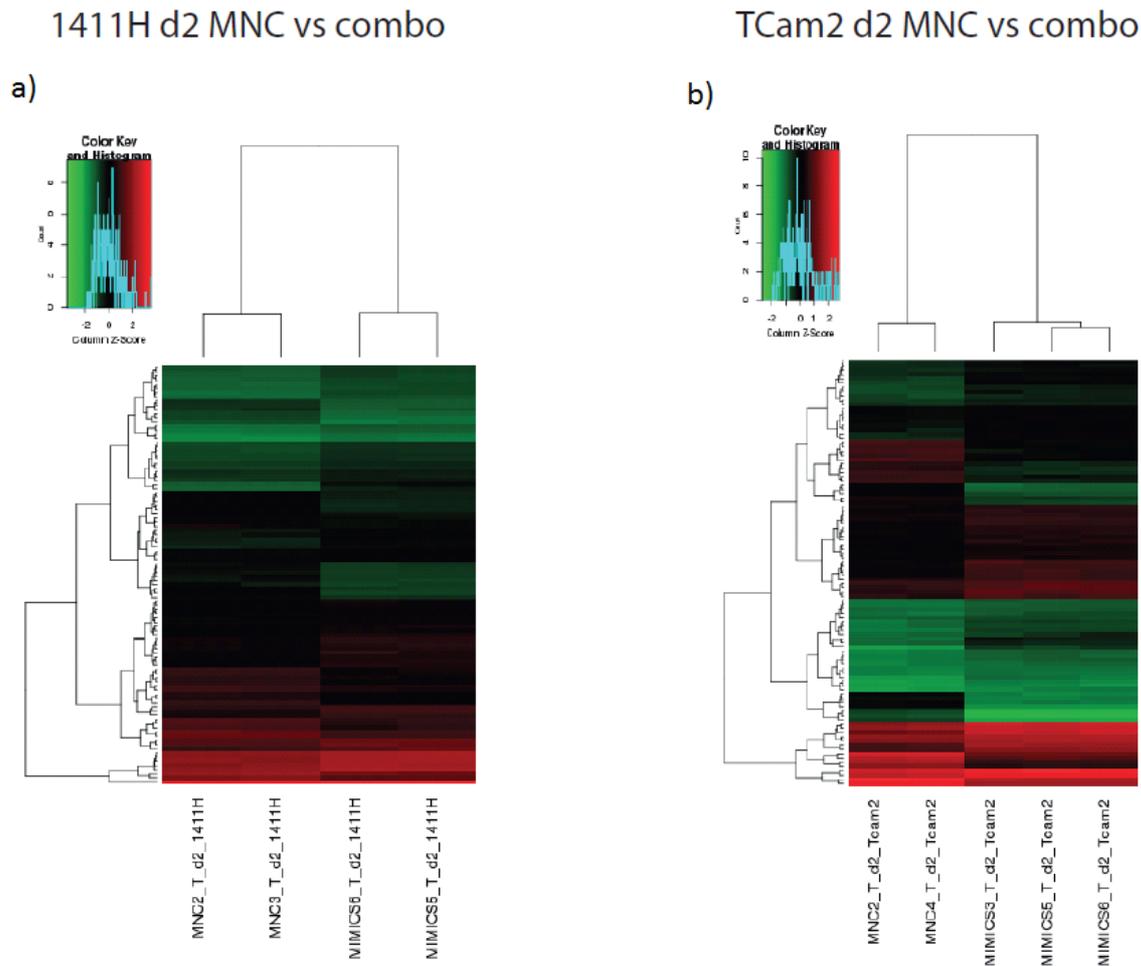
Figure 5.13: **Heatmaps showing hierarchical clustering.** The graphs show in a) 1411H cells treated with MNC vs treated with Combo, b) TCam2 cells treated with MNC vs treated with Combo. Combo is equimolar combination of mimics for miR-100-5p and miR-125b-5p. RNA samples were extracted two days after mimic transfection. The heatmaps display level of expression for both groups in the comparison as a colour map for the top differentially expressed 50 genes.

5.2.7 **MiRNA target genes at early timepoints after replenishment of miRNAs**

As mentioned in 5.2.3.2, the preliminary analysis was useful to demonstrate that differential expression of genes occurred after treatment with mimics replenishing miR-100-5p and miR-125b-5p compared with a MNC control two days after transfection, but further analysis had to be carried out to give insight on the genes specifically targeted by the miRNAs. Analysis of the differentially expressed genes was therefore further analysed using *Sylamer* and the significance landscape plots obtained for 1411H and TCam2 are shown in Figure 5.14 and Figure 5.15 respectively.

Consistent with the analysis of the data generated at day 7 post-transfection, these Sylamer plots originated from day 2 post replenishment expression data, consider 'words' corresponding to six Seed Complementary Regions sequences relative to miR-99a-5p/-100-5p and miR-125b-5p, namely sequences corresponding to three seed hexamers (1-6 nt, 2-7 nt, 3-8 nt), two heptamers (1-7 nt, 2-8 nt), and one octamer (1-8 nt) for each miRNA. Both miRNAs are represented on the plot, and all together six hexamers, four heptamers and 2 octamers curves are shown per graph per cell line.

a)

b)

c)
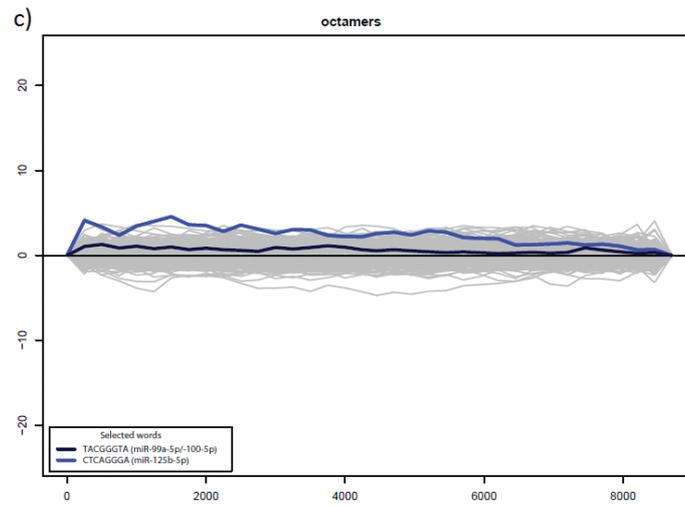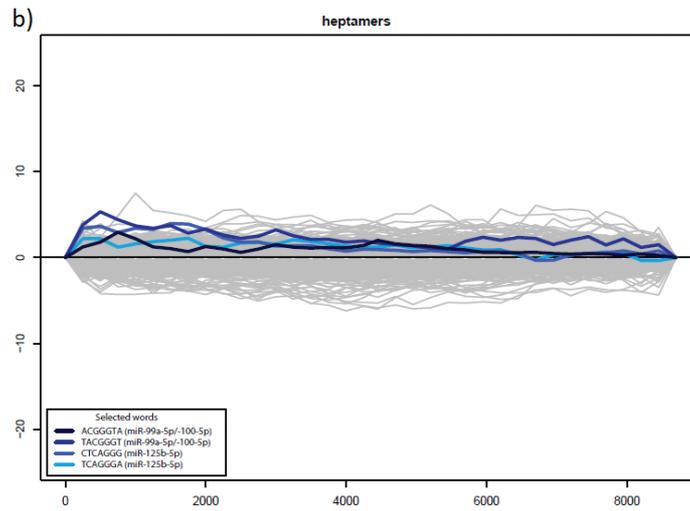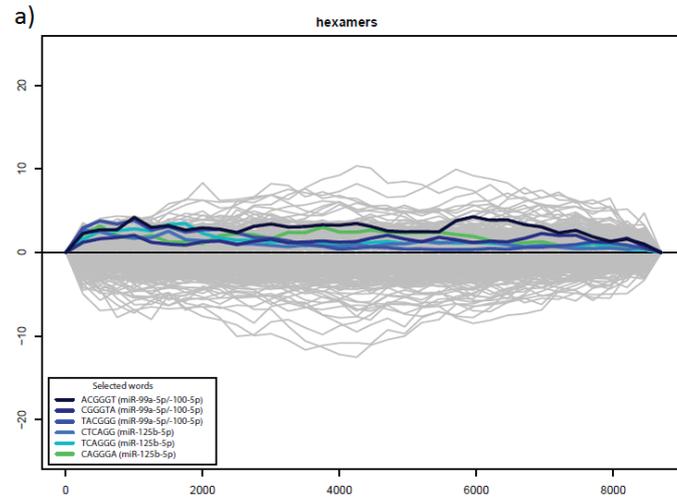
Figure 5.14: **Sylamer plots derived from analysis of differential expression data for 1411H treated with MNC vs Combo.** a) significance landscape plots were produced searching for 6-mers (a), 7-mers (b) and 8-mers (c). Highlighted are the significance curves for the seed complementary regions (SCR) for miR-99a-5p/-100-5p and miR-125b-5p. The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated. The y-axis represents the SCR abundance in the 3'UTR of the genes.

a) hexamers

b) heptamers

c) octamers
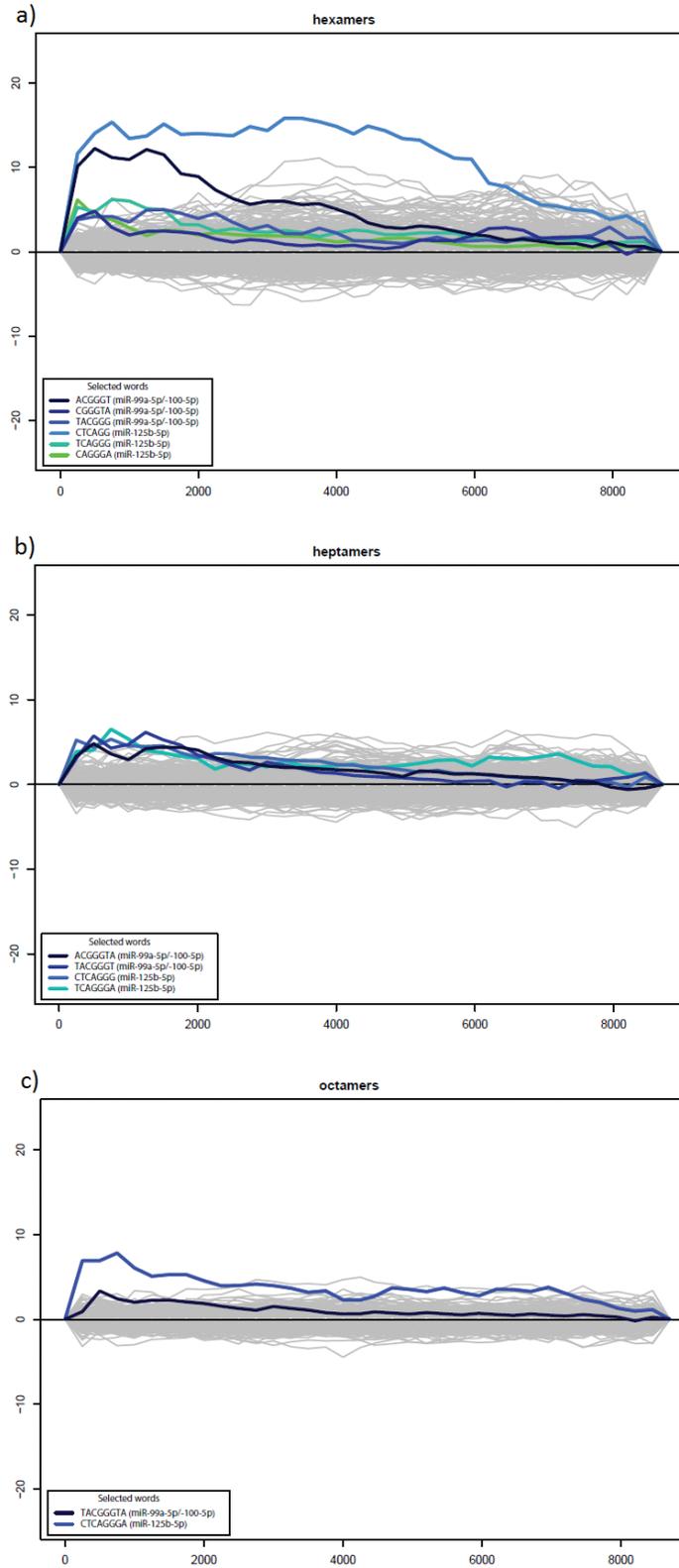
Figure 5.15: **Sylamer plots derived from analysis of differential expression data for TCam2 treated with MNC vs Combo.** a) significance landscape plots were produced searching for 6-mers (a), 7-mers (b) and 8-mers (c). Highlighted are the significance curves for the seed complementary regions (SCR) for miR-99a-5p/-100-5p and miR-125b-5p. The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated. The y-axis represents the SCR abundance in the 3'UTR of the genes.

Sylamer analysis showed significant curves for SCRs relative to miR-99a-5p/-100-5p and miR-125b-5p in plots derived from the differential expression gene list created by the analysis of the samples treated with MNC vs samples treated with Combo in both cell lines. Moreover, plots of overlapping SCR elements followed a coordinated pattern as expected for truly differentially expressed miRNAs. The effects of all SCR elements of interest (i.e. six hexamers, four heptamers and 2 octamers) were also considered simultaneously by Single-Sum Significance Sylamer analysis and the graphical representation of the data is shown in Figure 5.16 panel a) for 1411H, and panel b) for TCam2. For this analysis genes were grouped together in bins of 200 genes showing similar differential expression. The mRNA gene list was ranked left to right from most down-regulated to most up-regulated.
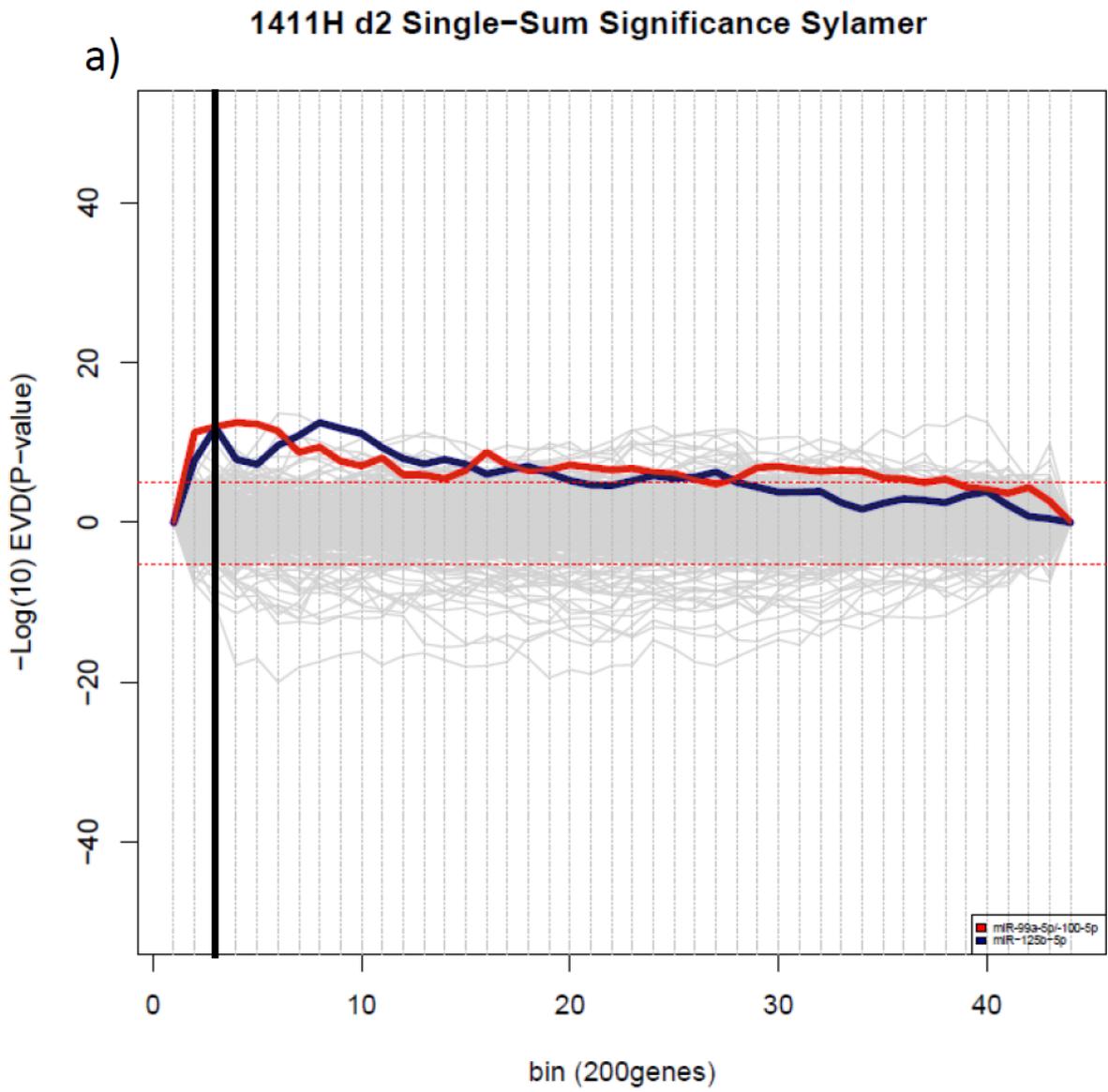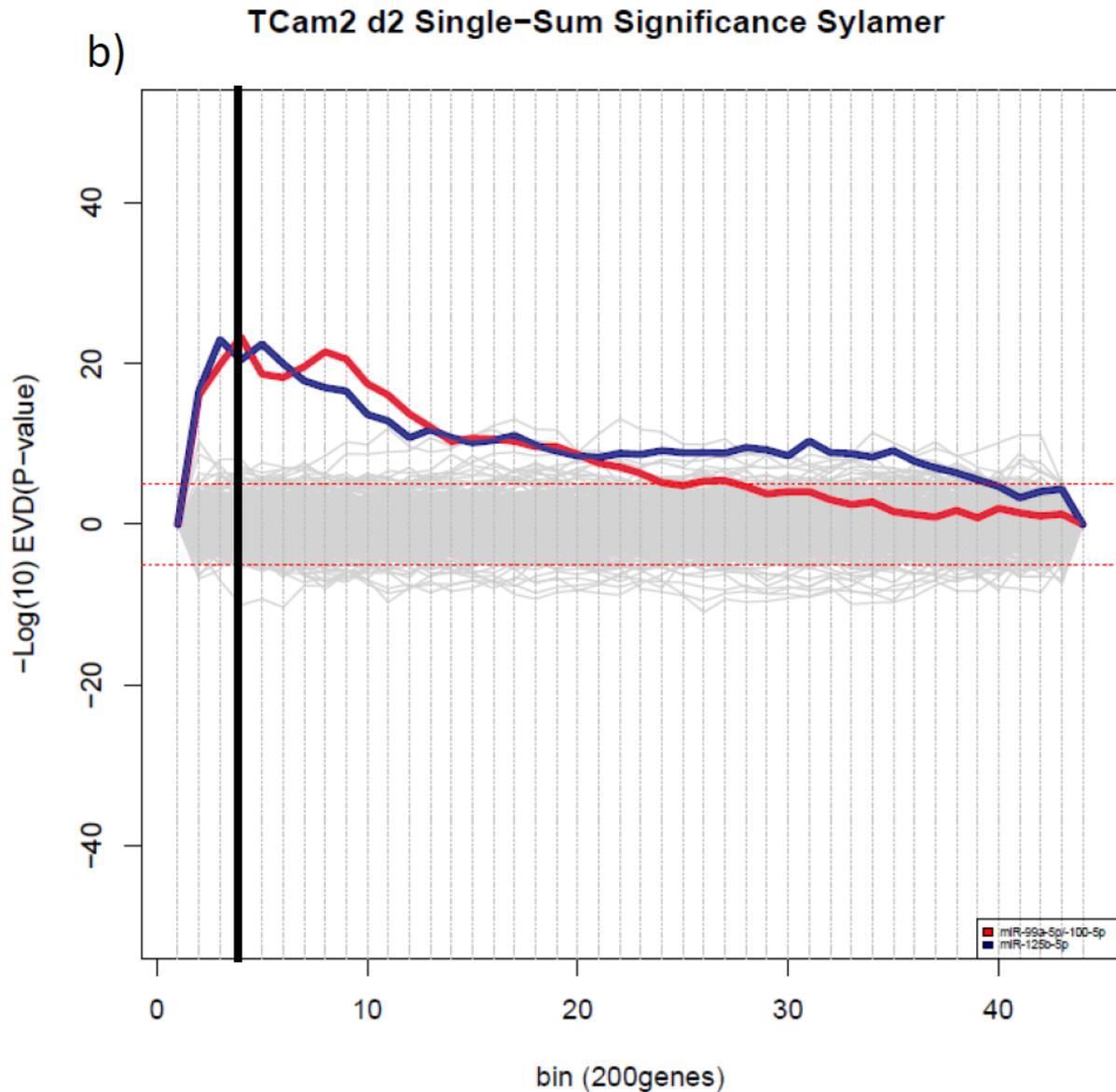
Figure 5.16: **SSS plots derived from analysis of differential expression data**. The graphs show in (a) 1411H, and (b) TCam2 cell lines treated with MNC vs Combo. Highlighted are the significance curves for miR-99a-5p/-100-5p (red line) and miR-125b-5p (blue line). The differentially expressed genes are plotted on the x-axis left to right, from the most downregulated to the most upregulated, each bin corresponds to 200 genes. The black vertical line corresponds to the assigned cut-off and the circle on its left underlines the genes that were used for further pathway analysis.

Significant Single-Sum curves were seen in both 1411H and TCam2 cells. Interestingly, in both 1411H and TCam2 cells downregulation of genes containing miRNA seed complementary regions for either miRNA appeared to have the same level of significance. This is in contrast to the data for the same analysis performed at day 7 post replenishment that showed significance only for genes containing the miR-125b-5p seed complementary region. The data suggests that at earlier timepoints both miRNAs mediate the downregulation of their target genes at similar levels, while at later timepoints miR-125b-5b is able to downregulate its targets more efficiently than miR-100-5p. For each cell line the peak of the SSS curve for miR-125b-5p and miR-100-5p was set as threshold for the identification of significant genes to take into consideration for further analysis.

5.2.8 Pathway analysis at an early timepoint after miRNA replenishment

From the Significant Single-Sum Sylamer plots, an appropriate cut off for each cell line was determined, and the most differentially expressed genes identified as containing the miR-125b-5p and miR-100-5p seed complementary regions above the threshold were utilised to determine the pathways regulated by the replenishment of miR-100-5p and miR-125b-5p. 400 genes were identified in 1411H cells (second bin) and 600 genes (third bin) were identified in TCam2 cells.

Pathway analysis using KEGG (Kyoto Encyclopedia of Genes and Genomes), gene onthology (GO) and reactome were carried out. The analysis was executed as follows. Firstly, each database was used individually and only pathways enriched for three or more genes were taken into account. Secondly, these resultant pathways found were compared among the three different bioinformatical systems. Finally, only the pathways found to be commonly enriched among the three approaches were further investigated. Significance of a particular pathway was then utilised to represent graphically the results obtained for 1411H (Figure 5.17 panel a) and TCam2 (Figure 5.17 panel b) cells.
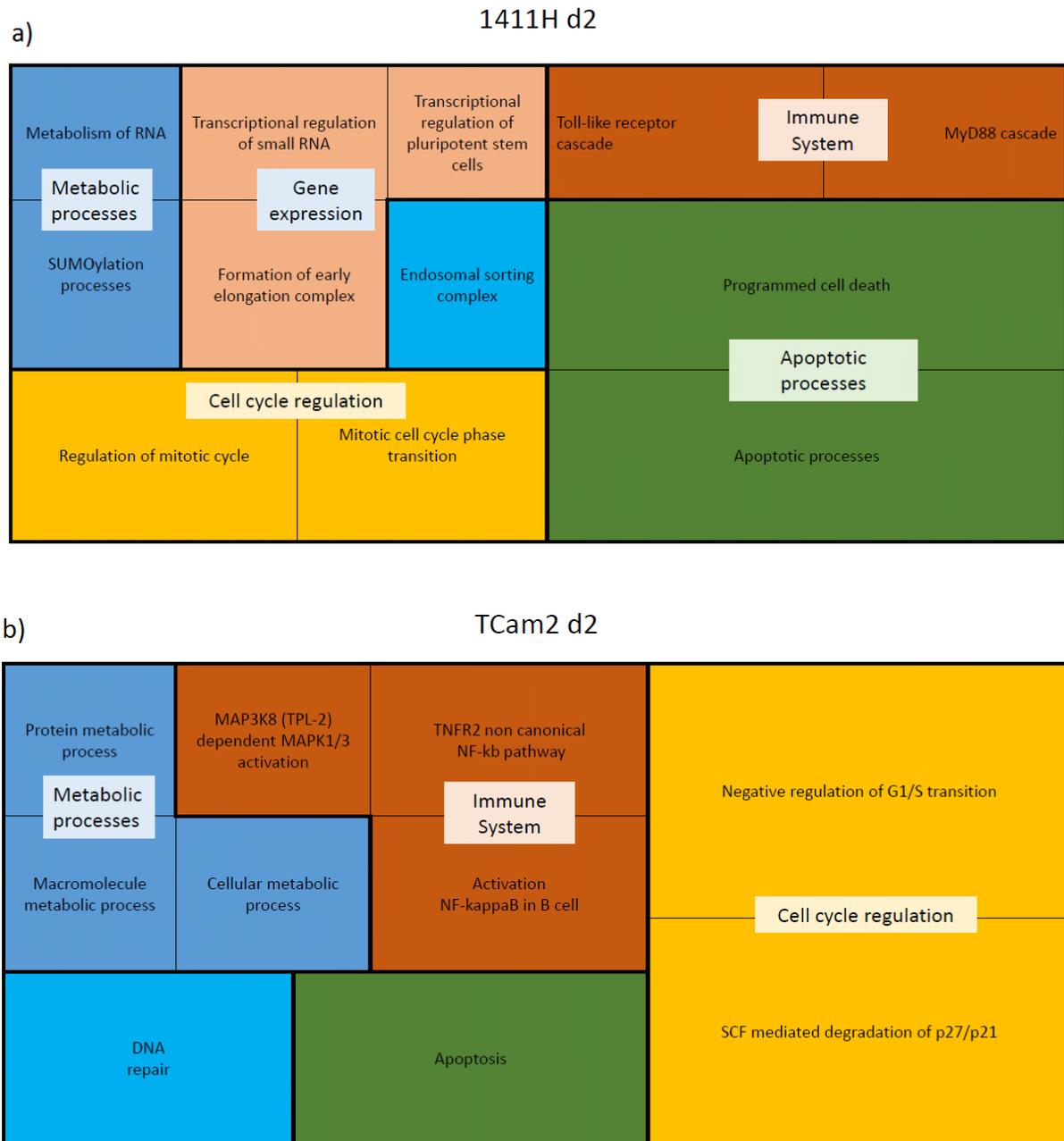
Figure 5.17: **Pathway analysis**. The most significant pathways affected by the replenishment of miR-100-5p and miR-125b-5p and identified in 1411H (panel a) and TCam2 (panel b) cells are shown. The size of each quadrant is relative to the significance of the pathway analysed.

The analysis of the pathways affected at day 2 post miRNA replenishment showed a much more consistent picture between cell lines than the one seen at day 7. For both cell lines the

majority of the genes differentially expressed at day 2 following replenishment of the miRNA code for proteins involved in similar processes, such as the regulation of cell cycle, apoptosis and immune system related processes. The regulation of cell cycle and apoptosis was also identified at day 7 and is consistent with the observed phenotypic changes on cell growth seen. Some differentially expressed genes appeared to be involved in cell line specific pathways, but this was only a minor proportion of the genes affected by the miRNAs replenishment.

Genes that were deregulated in both cell lines were then investigateded at in more details. 62 genes were found to be differentially regulated in both cell lines after replenishment with mimics of miR-100-5p and miR-125b-5p. These genes are of particular interest as they highlight common pathways of miRNA regulation in tumours that are otherwise highly heterogeneous, characterised by histologically different subtypes and occurring at varied anatomical site. Moreover, it puts forward the tantalising suggestion that miR-100-5p and miR-125b-5p, which are dysregulated in all mGCT, could be used as an all-encompassing potential therapeutic target.

The 62 differentially expressed genes which are affected by miR-100-5p and miR-125b-5p replenishment in both 1411H and TCam2 were then utilised for pathway analysis to investigate potential common mechanisms of miRNAs regulation (Figure 5.18).
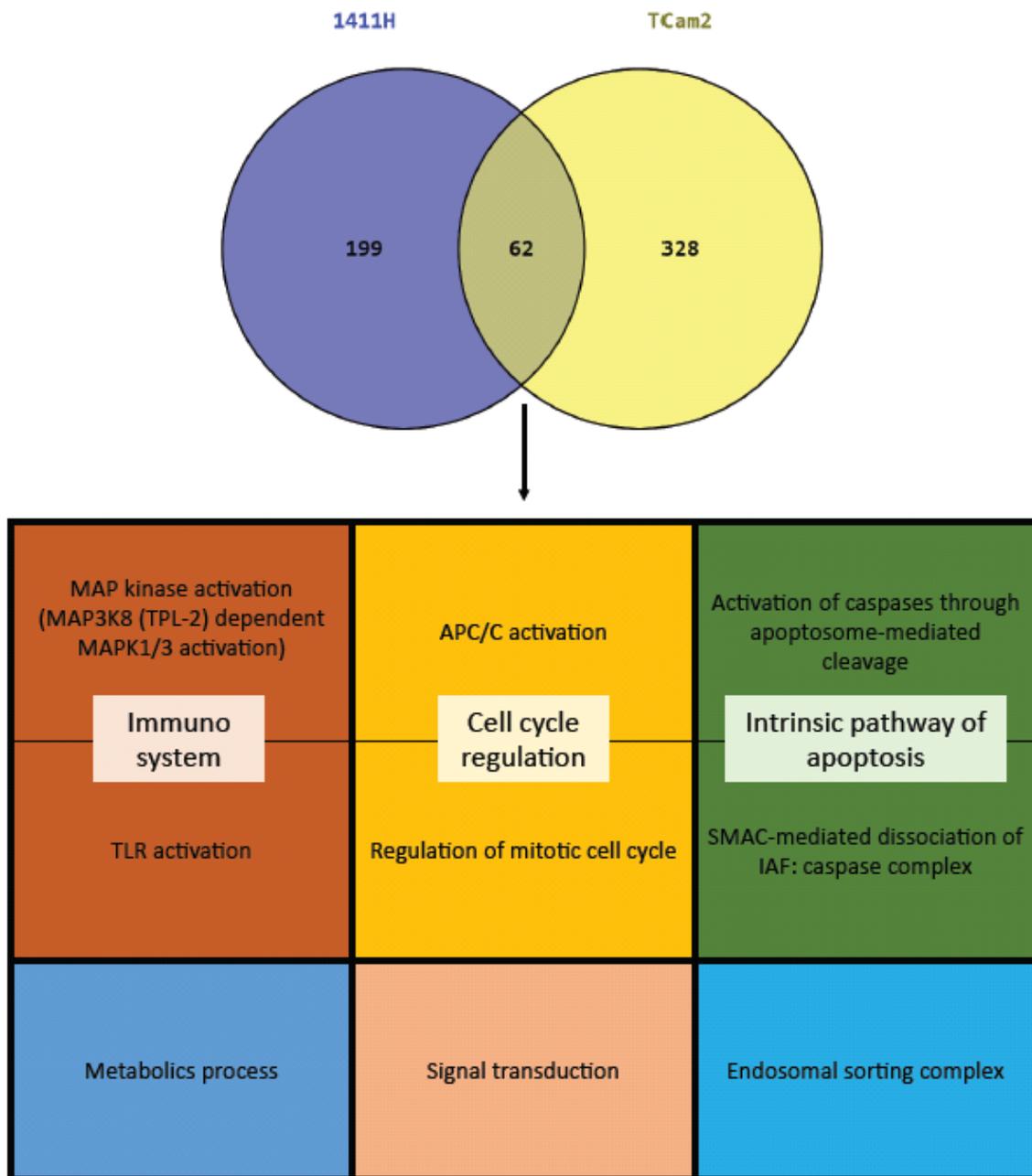
Figure 5.18: **Pathway analysis**. 62 differentially expressed genes were found to be affected by miRNA replenishment in both 1411H and TCam2 cell lines. The communal pathways identified are represented in the table. The size of each quadrant is relative to the significance of the pathway analysed. Pathway analyses were performed using Reactome and Gene Onthology (GO).

Pathways analysis with the 62 communal genes identified all the pathways that were initially identified in each cell line separately, and again immune system related processes, cell cycle regulation and apoptosis were the main area of gene involvement, in agreement with the phenotypic changes observed in the replenishment experiments. Further work was therefore carried out to validate these findings by RT-qPCR and western blotting.

5.2.9 Validation of genes involved in cell cycle regulation

Having correlated a role in the regulation of the cell cycle to some of the genes identified as containing seed complementary sequences in their 3' UTR to miR-99a-5p/-100-5p and/or miR-125b-5p and being significantly under-expressed in both 1411H and TCam2 cells, a panel of eight genes fulfilling these criteria was selected for validation. Of these, five, were targeted by both miR-99a-5p/-100-5p and miR-125b-5p (ARID 3A, E2F7, TRIM71, FGFR3, and CTDSPL); one was targeted by miR-99a-5p/-100-5p only (mTOR) and two genes were targeted by miR-125b-5p only (ARID 3B and LIN28).

5.2.9.1 Expression of selected genes in tissue

To assess if the genes identified were deregulated in mGCT tissues, data from The Cancer Genome Atlas (TCGA) combined with Genotype-Tissue Expression (GTEx) data via GEPIA, an online tool, was analysed. 137 samples derived from TGCT patients were compared with 165 normal samples. The results are shown in a Log2 scale and the different colours represent the degree of expression (dark green being over-expressed and light green being under-expressed). Interestingly, some of the genes of interest were upregulated in TGCT tissue compared with normal tissue, suggesting that such genes may contribute to malignancy (Figure 5.19).

Figure 5.19: **Gene expression in tumour tissue (T) vs normal tissue (N).** The data were acquired from GEPIA, an online platform that integrates data from TCGA and GTEx. Overexpression is shown in dark green, underexpression is shown in light green.

5.2.9.2 **Validation of miRNA gene targets via RT-qPCR and Western blot**

To validate the results obtained by microarray analysis and also to study the potential contribution of the replenishment of the individual miRNAs to the observed changes of genes involved in the regulation of the cell cycle, cells were treated with either the single miRNA (miR-100-5p or miR-125b-5p) at two different concentration: 8.3 nM, and 16.7 nM or with the combination (Combo) of the two miRNAs (miR-100-5p + miR-125b-5p) at 16.7 nM. As reference, cells treated using mimic non-targeting control (MNC) at the appropriate matched concentration were used (Figure 5.20, Figure 5.21, Figure 5.22, Figure 5.23).

Figure 5.20: **Effects of miRNA replenishment on TRIM71 in 1411H and TCam2 cells.** a) gene expression as determined by RT-qPCR: TRIM71 expression levels were examined at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the key. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPL13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged. b) Western blot analysis of protein expression at day 2 and day 3 post replenishment with 16.7 nM mimics. Shown are representative blots, the experiment was performed twice with similar results.

RT-qPCR analysis of TRIM71 expression confirmed the microarray data, as this gene was downregulated 2 days post Combo transfection in both cell lines. Interestingly, expression of TRIM71 was modulated by the combination of mimics with different kinetics in the two cell lines. In 1411H cells, the downregulation of TRIM71 after replenishment of both miRNAs peaked at day 1, and then decreased with time, so that it was lost by day 3, whereas in TCam2

cells the maximum downregulation was seen at day 2, with both day 1 and day 3 showing similar levels of downregulation (Figure 5.20, panel a). When comparing the effects of the mimics on their own to the Combo, the results obtained again differ between the two cell lines. In 1411H cells the combination appeared to have an effect equivalent to the sum of the effects seen for each mimic (at 8.3 nM), while the relationship between the effects seen with each mimic individually and the Combo was more variable in TCam2 cell. For example, at day 1 post transfection there did not appear to be any difference in the ability of the mimics either individually or in combination to downregulate the expression of TRIM71, whereas at day 2 the combination was significantly more efficient than the mimics alone (p-value: 0.002). Furthermore, the data suggested that doubling the dose of either individual mimic, did not result in the same effects on transcription shown by their combination at least at days 1 and 2 post transfection. The data at day 3 was more complex, as in both cell lines the effects of miR-100-5p replenishment on TRIM71 seemed to disappear more rapidly than those induced by miR-125b-5p replenishment. This could potentially explain the rapid reduction in TRIM71 downregulation seen also with the Combo. Analysis of protein levels following miRNA replenishment by Western blotting showed a downregulation of TRIM71 protein induced by either each mimic alone or in combination in both cell lines, both at day 2 and 3 after transfection. However, all miRNAs replenishment was carried out using 16.7 nM mimics either alone or in combination, and, in agreement with the RT-qPCR data, the combination of both mimics together always appeared more effective than a single mimic at the equivalent dose. Moreover, the downregulation appeared more prominent at day 3, particularly in TCam2 cells.
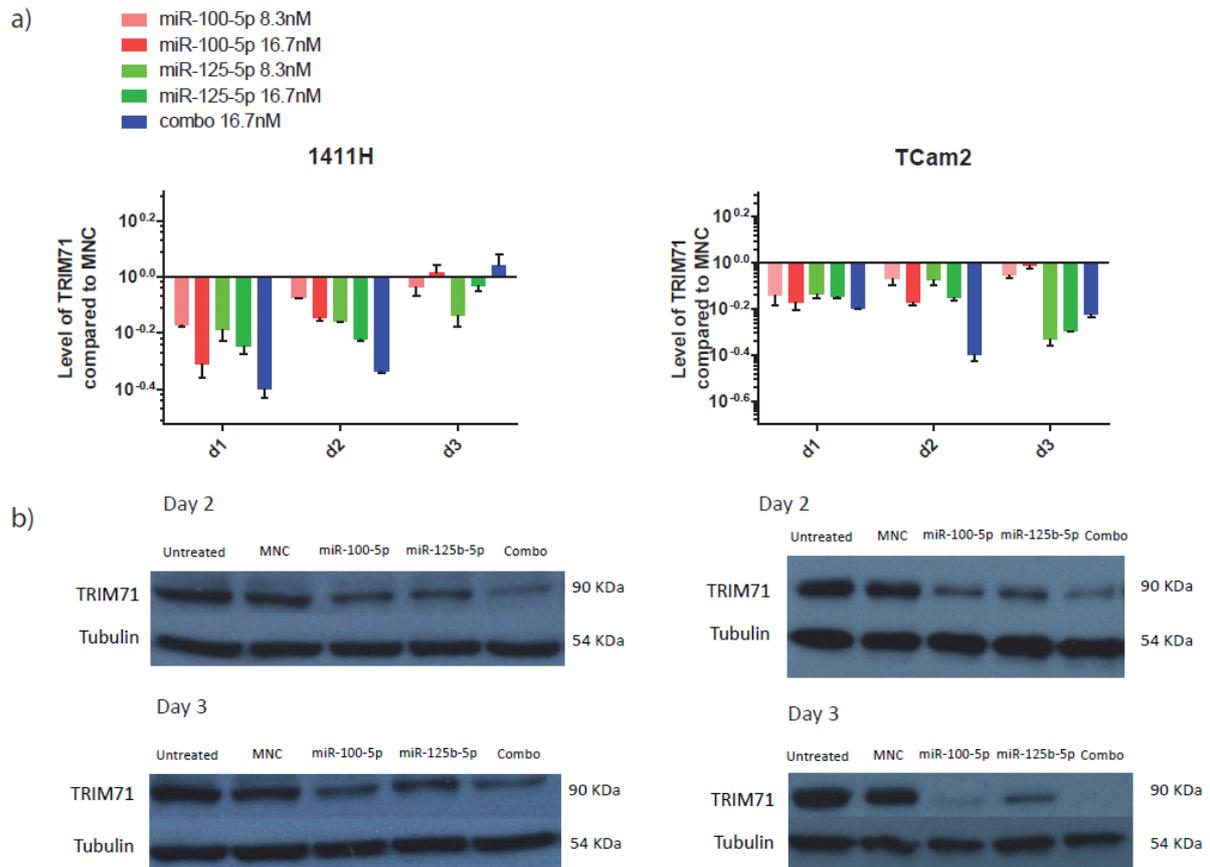
Figure 5.21: **Effects of miRNA replenishment on ARID 3A in 1411H and TCam2 cells.** a) gene expression as determined by RT-qPCR: ARID 3A expression levels were examined at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the key. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPL13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged. b) Western blot analysis of protein expression at day 2 and day 3 post replenishment with 16.7 nM mimics. Shown are representative blots, the experiment was performed twice with similar results.

RT-qPCR analysis of ARID 3A expression also confirmed the microarray data, as this gene was downregulated 2 days post Combo transfection in both cell lines. Expression of ARID 3A was modulated by the combination of mimics with similar kinetics in the two cell lines and the maximum downregulation was seen at day 2, with both day 1 and day 3 showing similar levels

of downregulation. When comparing the effects of the mimics on their own to the Combo, the results obtained were consistent between the two cell lines. In both cell lines replenishment of miR-100-5p alone did not downregulate the transcription of ARID 3A at any of the doses tested: upregulation of these transcripts was in fact seen at day 2 and day 3 post transfection both with 8.3 nM and 16.7 nM mimics. On the other hand, replenishment of miR-125b-5ps alone did resulted in downregulation of ARID 3A transcription, however, the relationship between the effects of this mimic alone and the effects seen with the Combo were sometimes conflicting. Although in some cases the effects seen by the Combo were similar to those seen with miR-125b-5p alone (for example in 1411H cells at day 2 and in TCam2 cells at day 1) (Figure 5.21, panel a), the combination at other times appeared to have an effect statistically greater than seen with the mimic alone, for instance in 1411H cells at day 3 (p-value: 0.02) and in TCam2 cells at day 2 (p-value=0.006). Furthermore, the data suggested that doubling the dose of the miR-125b-5p mimic on its own did not result in the same effects on transcription shown by their combination: differently from TRIM71, doubling the dose of miR-125b-5p had a much greater effect than the Combo of mimics, suggesting that in this instance the two mimics acted in an antagonistic manner.

Analysis of protein levels following miRNA replenishment by Western blotting showed a downregulation of ARID 3A protein induced by miR-125b-5p alone or in combination in both cell lines, both at day 2 and 3 after transfection when replenishment was carried out using 16.7 nM mimics either alone or in combination. Consistently with the RT-qPCR data, no effect was seen on protein levels using the mimic for miR-100-5p alone.
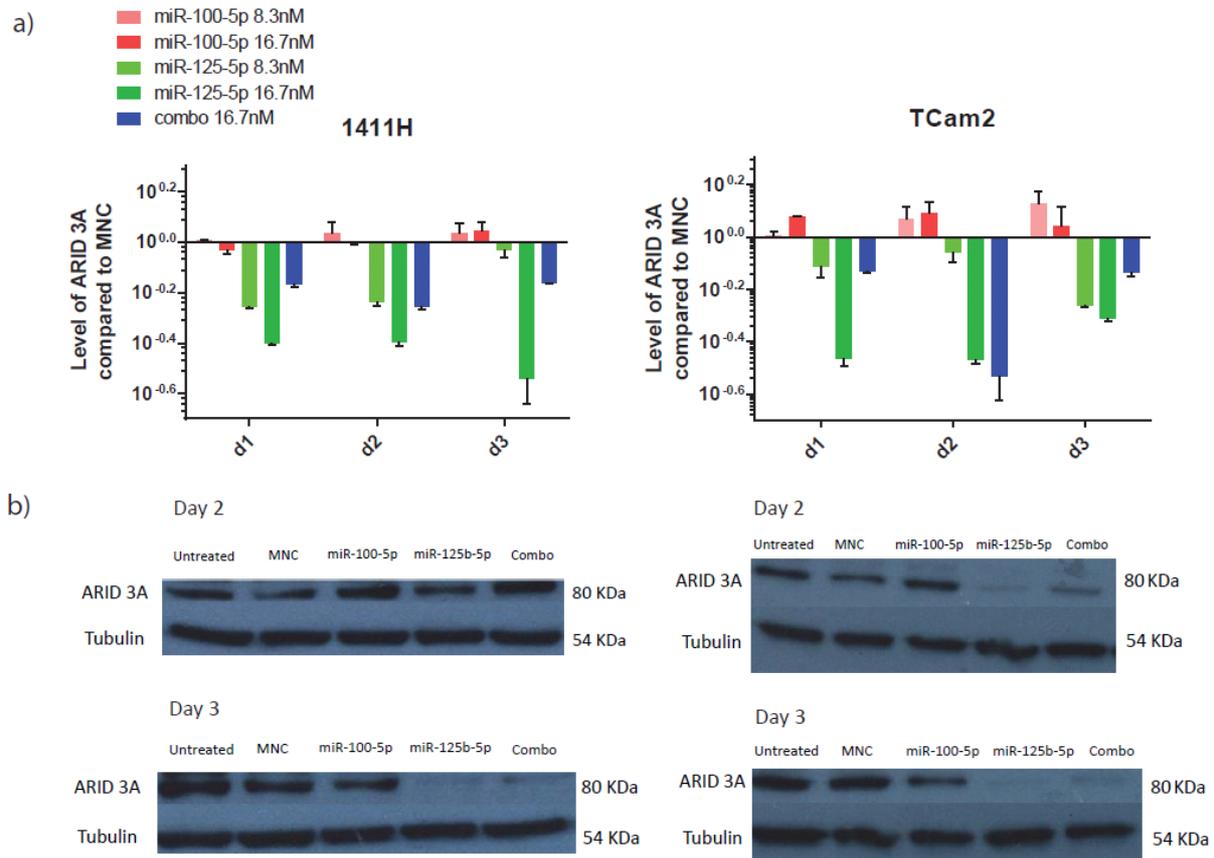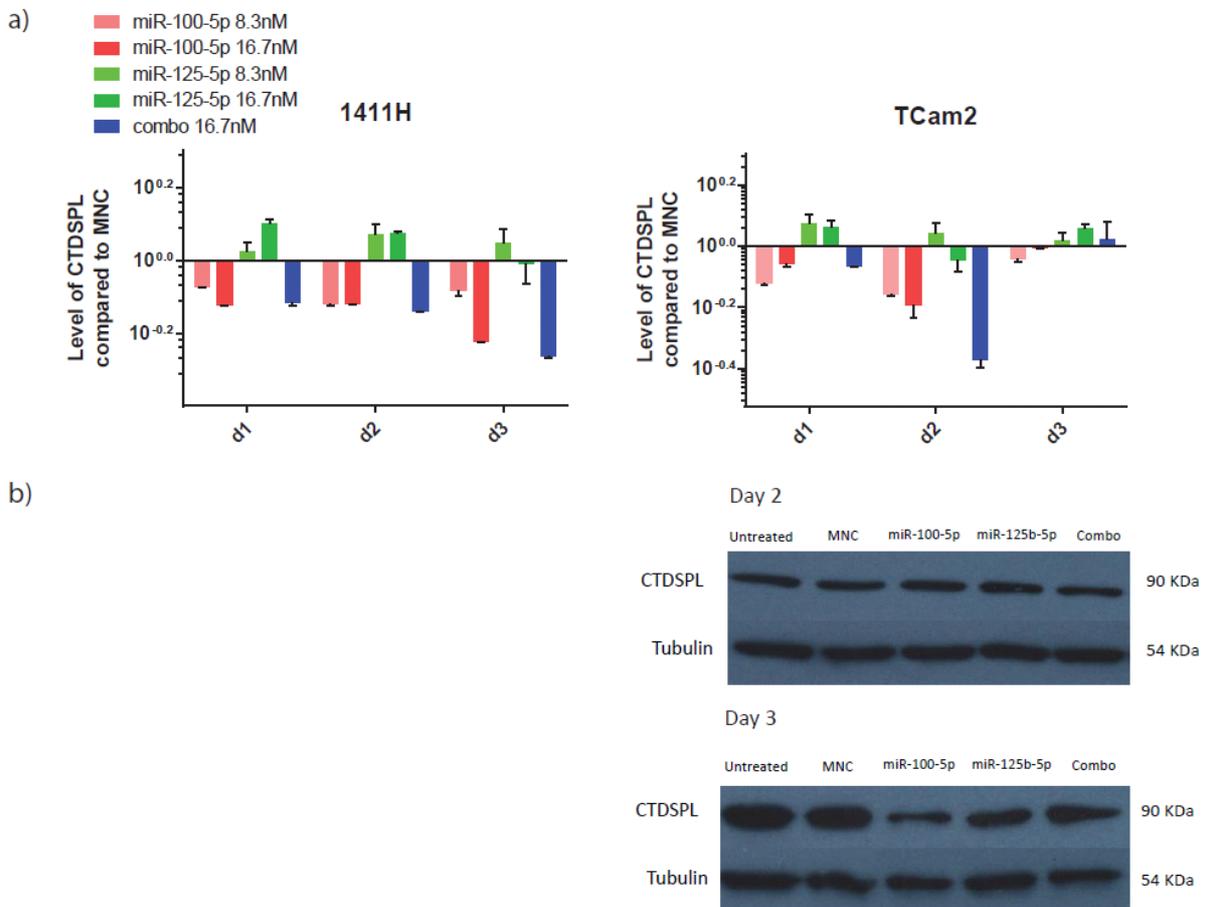
Figure 5.22: **Effects of miRNA replenishment on CTDSPL in 1411H and TCam2 cells.** a) gene expression as determined by RT-qPCR: CTDSPL expression levels were examined at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the key. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPL13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged. b) Western blot analysis of protein expression at day 2 and day 3 post replenishment with 16.7 nM mimics. Shown are representative blots, the experiment was performed twice with similar results.

RT-qPCR analysis of CTDSPL expression confirmed the microarray data, as this gene was downregulated 2 days post Combo transfection in both cell lines. Expression of CTDSPL was modulated by the combination of mimics with different kinetics in the two cell lines. In 1411H cells, the downregulation of CTDSPL after replenishment of both miRNAs in combination,

increased with time and peaked at day 3, whereas in TCam2 cells the maximum downregulation was seen at day 2, with both day 1 and day 3 showing similar levels of downregulation. When comparing the effects of the mimics on their own to the Combo, the results obtained were consistent between the two cell lines. In both cell lines replenishment of miR-125b-5p alone did not downregulate the transcription of CTDSPL at any of the doses tested: upregulation of these transcripts was in fact seen at day 1, day 2 and day 3 post transfection both with 8.3 nM and 16.7 nM mimics. On the other hand, replenishment of miR-100-5ps alone did result in downregulation of CTDSPL transcription, however, in a manner parallel to ARID 3A, the relationship between the effects of this mimic alone and the effects seen with the Combo were sometimes conflicting. Although in some cases the effects seen by the Combo were similar to those seen with miR-100-5p alone (for example in 1411H cells at day 2), the combination at other times appeared to have an effect statistically greater than seen with the mimic alone, for instance in 1411H cells at day 3 (p-value: 0.0002) and in TCam2 cells at day 2 (p-value: 0.001). Furthermore, the data suggested that doubling the dose of miR-100-5p, had similar effects on transcription to the combination: with the exception of day 2 in TCam2 cells where doubling the dose of miR-100-5p had a much smaller effect than the Combo of mimics on CTDSPL transcription.

Analysis protein levels following miRNA replenishment by Western blotting, showed a downregulation of CTDSPL protein induced by miR-100-5p alone or in combination in TCam2 cells, at day 3 after transfection when replenishment was carried out using 16.7 nM mimics. In 1411H no changes were detected by Western blot (data not shown).

5.2.9.3 Seed Complementary Region length influences regulation of gene expression

ARID 3A and CTDSPL were identified as targets of both miR-100-5p and miR-125b-5p by *Sylamer* analysis. However, the data presented above suggested that these genes are regulated differently from each other. ARID 3A transcription was not affected by the mimic for miR-100-5p alone while it was affected by the mimic for miR-125b-5p or the combination of the

two mimics. On the other hand, CTDSPL transcription was not affected by the mimic for miR-125b-5p alone, while it was affected by the mimic for miR-100-5p or the Combo.

To identify the reasons for this differential regulation the number of potential binding sites for each individual miRNA in the 3'UTR of the target gene was assessed. A correlation between the numbers of 6-mers, 7-mers and 8-mers SCR for either miR-100-5p or miR-125b-5p on the 3'UTR of the target gene and the level of its downregulation was found (Table 5.2).

Gene name	6-mers		7-mers		8-mers	
	miR-100-5p	miR-125b-5p	miR-100-5p	miR-125b-5p	miR-100-5p	miR-125b-5p
ARID 3A	1	6	1	6	0	1
CTDSPL	2	4	4	2	2	0

Table 5.2: **SCR sites specific for miR-100-5p and miR-125b-5p in the 3'UTR of ARID 3A and CTDSPL transcripts.**

Six SCR binding sites for miR-125b-5p were found in the 3'UTR of ARID 3A, one of which was complementary to the miRNA for a stretch of 8 nucleotides. In the same area only one SCR for miR-100-5p was identified, with the maximal length of complementarity of 7 nt. Conversely, four SCR binding sites for miR-100-5p were found in the 3'UTR of CTDSPL, two of which were complementary to miR-100-5p for a stretch of 8 nucleotides, while only two SCR for miR-125b-5p were identified, with the maximal length of complementarity of 7 nt. It is therefore possible therefore that the specific regulation of the transcription of these genes by either of these miRNAs is a result of the differential number of SCR sites and the length of the SCR sites themselves.
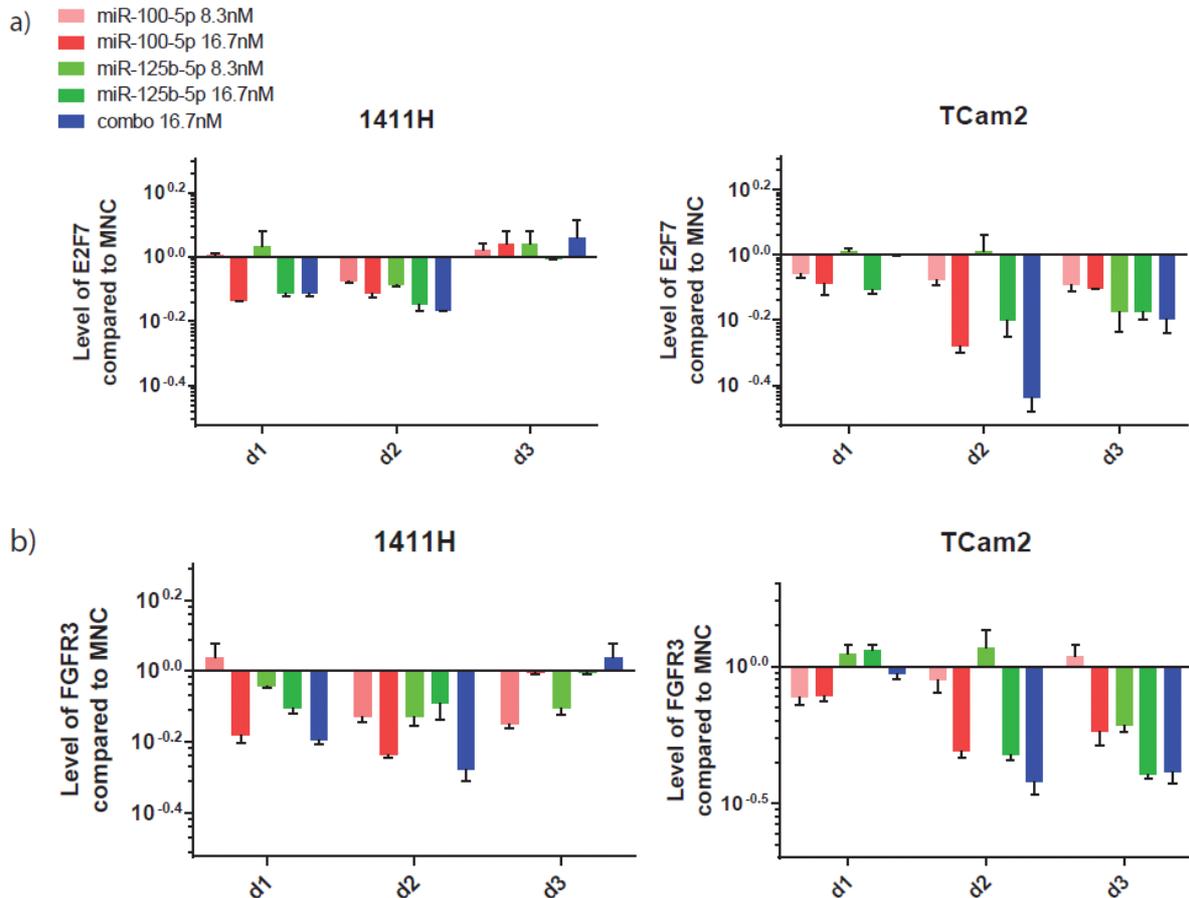
Figure 5.23: **Effects of miRNA replenishment on a) E2F7 and b) FGFR3 in 1411H and TCam2 cells.** Gene expression was determined by RT-qPCR at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the legend. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPL13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged.

Two more genes identified to be regulated by miR-100-5p and miR-125b-5p were validated by RT-qPCR. Both E2F7 and FGFR3 were shown to be downregulated at day 2 post transfection of the Combo mix of mimics in both 1411H and TCam2 cells by RT-qPCR, confirming the microarray data.

Expression of E2F7 was modulated by the combination of mimics with similar kinetics in the two cell lines, with the maximum downregulation seen at day 2. This consistency between cell lines was also seen when comparing the effects of the mimics on their own to the Combo. In both cell lines replenishment of miR-100-5p or miR-125b-5p alone downregulated the transcription of E2F7 both with 8.3 nM and 16.7 nM mimics. Although in some cases the effects seen by the Combo were similar to those seen with a mimic alone (for example in TCam2 cells at day 3), the combination at all other times appeared to have an effect greater than seen with the mimic alone. Furthermore, the data suggested that doubling the dose of miR-100-5p, or miR-125b-5p had less effects on transcription than the combination. These observations were also true for FGFR3, and overall the two genes behaved very similarly.

E2F7 is fundamental to the control of several different steps of the cell cycle and it is hence differentially regulated at various stages of it. For this reason, it was difficult to have a consistent evaluation of its protein levels in culture, particularly as the cells were not synchronised.

As mentioned earlier, some of the genes discovered using *Sylamer* and chosen for validation are targeted by either miR-100-5p or miR-125b-5p alone (Figure 5.24). The effects of the replenishment of the miRNAs either on their own or in combination (Combo) on the transcription of these genes were tested by RT-qPCR.
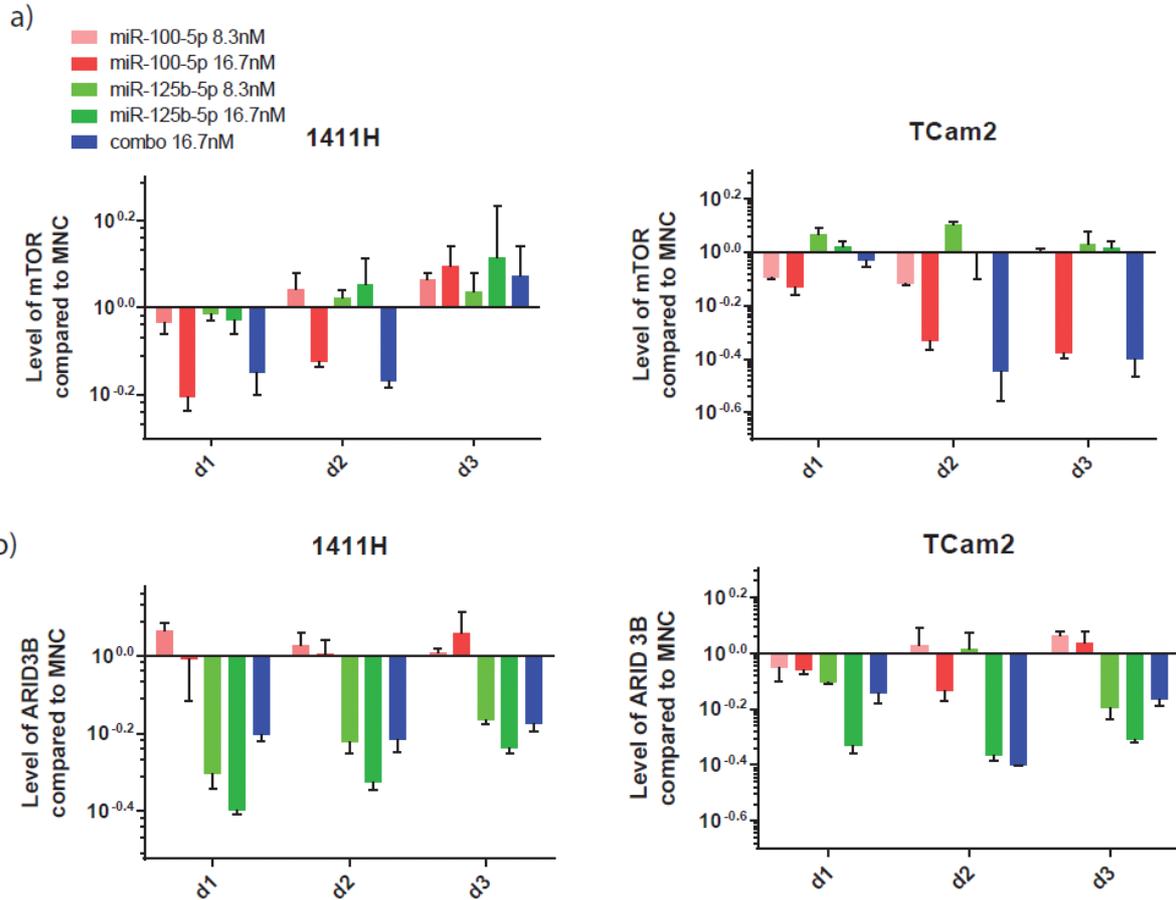
Figure 5.24: **Effects of miRNA replenishment on a) mTOR and b) ARID 3B in 1411H and TCam2 cells.** Gene expression was determined by RT-qPCR at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the legend. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPL13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged.

mTOR is targeted only by miR-100-5p. RT-qPCR analysis of mTOR expression confirmed the microarray data, as this gene was downregulated 2 days post Combo transfection in both cell lines (Figure 5.24, panel a). Expression of mTOR was modulated by the combination of mimics

with similar kinetics in the two cell lines and the maximum downregulation was seen at day 2, with both day 1 and day 3 showing lower levels of downregulation. When comparing the effects of the mimics on their own to the Combo, the results obtained were consistent between the two cell lines. As predicted, in both cell lines replenishment of miR-125b-5p alone did not downregulate the transcription of mTOR at any of the doses tested, and some minor upregulation of these transcripts was in fact seen. On the other hand, replenishment of miR-100-5ps alone did result in downregulation of mTOR transcription, but generally at lower level than the corresponding Combo treatment. Furthermore, as seen for other genes, the data suggested that doubling the dose of the miR-100-5p mimic on its own, did not result in the same effects on transcription shown by their combination.

ARID 3B is targeted only from miR-125b-5p. RT-qPCR analysis of ARID 3B expression confirmed the microarray data, as this gene was also downregulated 2 days post Combo transfection in both cell lines (Figure 5.24, panel b). As for mTOR, expression of ARID 3B was modulated by the combination of mimics with similar kinetics in the two cell lines and the maximum downregulation was seen at day 2, with both day 1 and day 3 showing lower levels of downregulation. Moreover, when comparing the effects of the mimics on their own to the Combo, the results obtained were consistent between the two cell lines. As predicted, in both cell lines replenishment of miR-100-5p alone did not downregulate the transcription of ARID 3B at any of the doses tested, and some minor upregulation of these transcripts was in fact seen. On the other hand, replenishment of miR-125b-5ps alone did result in downregulation of ARID 3B transcription, but generally at lower level than the corresponding Combo treatment. Furthermore, the data suggested that doubling the dose of the miR-125b-5p mimic on its own, results in the more prominent effects on transcription than shown by the combination of the two mimics.

All together these data confirm the selective nature of the downregulation of the expression of some genes, as identified by *Sylamer* analysis.

5.2.9.4 **LIN28A is specifically targeted by miR-125b-5p**

LIN28A is one of the gene identified using *Sylamer* and chosen for validation. The effects of the replenishment of the miRNAs either on their own or in combination (Combo) on the

transcription of this gene was therefore tested by RT-qPCR. This gene is of particular interest because previous work carried out in our lab had shown that LIN28 downregulation by siRNA resulted in an increase of the levels of let-7 (Figure 5.25). Interestingly, a let-7 sequence is present at both chromosomal sites where miR-125b-5p and mir-100-5p are located but is regulated differently from the other miRNAs in the same cluster. In particular, let-7 has been shown to be downregulated by posttranscriptional mechanisms involving LIN28, while LIN28 levels are regulated by let-7 via direct downregulation of its transcription.

This mechanism of regulation has been shown in several contexts and let-7 is thought to have a role in a variety of cancers, including mGCTs.
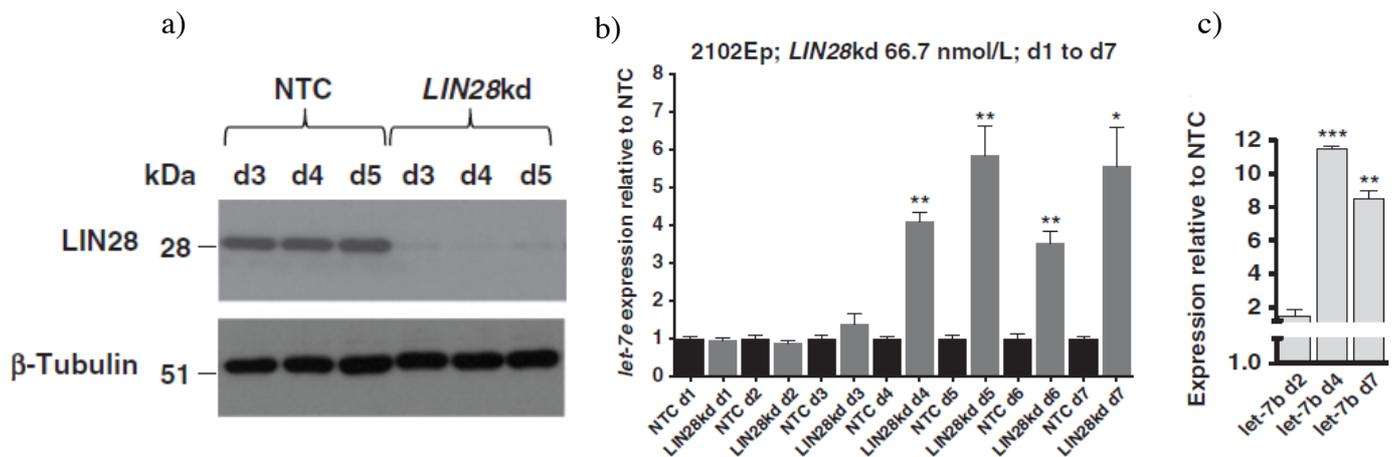
Figure 5.25: **LIN28A let-7 feedback**. a) Depletion of LIN28 protein levels after treatment with siRNAs for LIN28. Shown are protein levels at day 3 to 5. b) Levels of let-7e and c) levels of let-7d, both established by RT-qPCR. Figure adapted from Murray et al, CR, 2013.

As shown in Figure 5.25, the levels of *let*-7 family members start to rise from day 4 following LIN28 knock-down, implying a double-negative feedback loop mechanism regulating the expression of these two proteins. Importantly, LIN28 was identified as a target of miR-125b-5p by the *Sylamer* analysis (Figure 5.26).
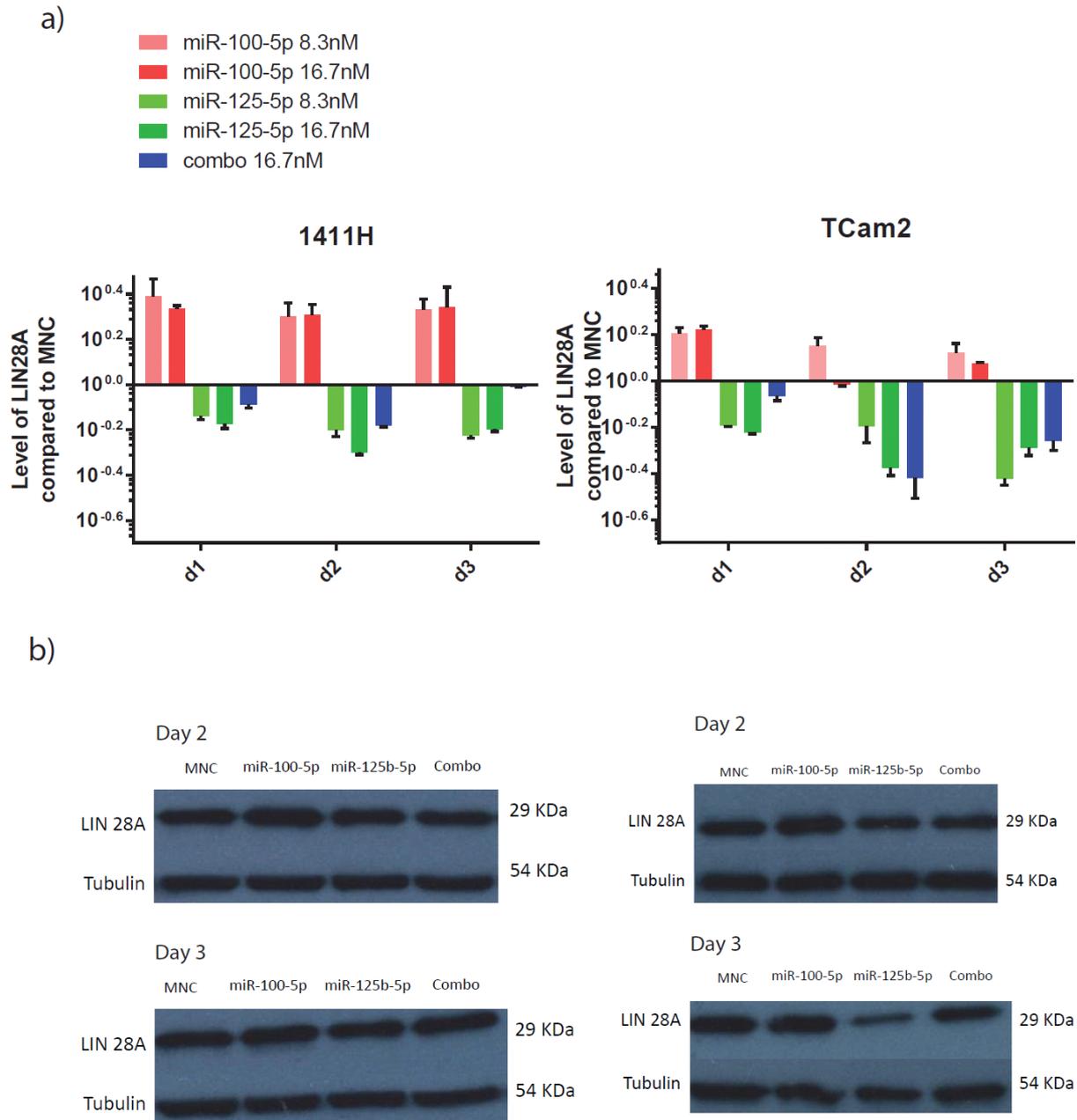
Figure 5.26: **Effects of miRNA replenishment on LIN28A in 1411H and TCam2 cells.** a) gene expression as determined by RT-qPCR: LIN28A expression levels were examined at day 1, day 2 and day 3 post transfection. Different shades of the same colour illustrate different concentrations of mimics in the same condition (single or combination of mimics), as described in the legend. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Data presented as mean ± SEM of n=3 biological replicates. Relative gene expression was normalised using three housekeeping genes (RPLI13A, YWAHZ and ACTβ) and independently calculated expression ratios were then averaged. b) Western blot analysis of protein expression at day

2 and day 3 post replenishment with 16.7 nM mimics. Shown are representative blots, the experiment was performed twice with similar results.

RT-qPCR analysis of LIN28 expression confirmed the microarray data; this gene was downregulated 2 days post Combo transfection in both cell lines. Expression of LIN28 was modulated by the combination of mimics with similar kinetics in the two cell lines and the maximum downregulation was seen at day 2, with both day 1 and day 3 showing similar levels of downregulation. In both cell lines replenishment of miR-100-5p alone did not downregulate the transcription of LIN28 at any of the doses tested: upregulation of these transcripts was in fact consistently seen at all analysed times post transfection both with 8.3 nM and 16.7 nM mimics, probably due to secondary effects. On the other hand, replenishment of miR-125b-5ps alone did result in downregulation of LIN28 transcription. In most cases the effects seen after the transfection of the Combo were similar to those seen with miR-125b-5p alone.

Analysis of protein levels following miRNA replenishment by Western blotting showed a downregulation of LIN28 protein induced by miR-125b-5p alone or in combination in TCam2 cells, mostly at day 3 after transfection when replenishment was carried out using 16.7 nM mimics either alone or in combination. Consistently with the RT-qPCR data, no effect was seen on protein levels using the mimic for miR-100-5p alone in TCam2 cells, and minimal effect was seen on 1411h cells at both time points.

5.2.10 Regulation of the expression of LIN28A, let-7 and miR-125b-5p

To further dissect out the role of miR-125b-5p on the regulation of LIN8 and *let*-7 family members further studies were carried out.

Firstly, the effect of replenishing miR-125b-5p was studied over a longer time frame, particularly in view of the lack of change in protein levels seen at d3 post transfection with the Combo treatment. This suggested that at day 3 post transfection, the upregulation of LIN28 induced by the presence of miR-100-5p in the Combo has an antagonistic effect on the downregulation of LIN28 induced by miR-125b-5p. Therefore, I tested whether effects

mediated by miR-125b-5p were more significant than the effects mediated by miR-100-5p at later timepoints (Figure 5.27).
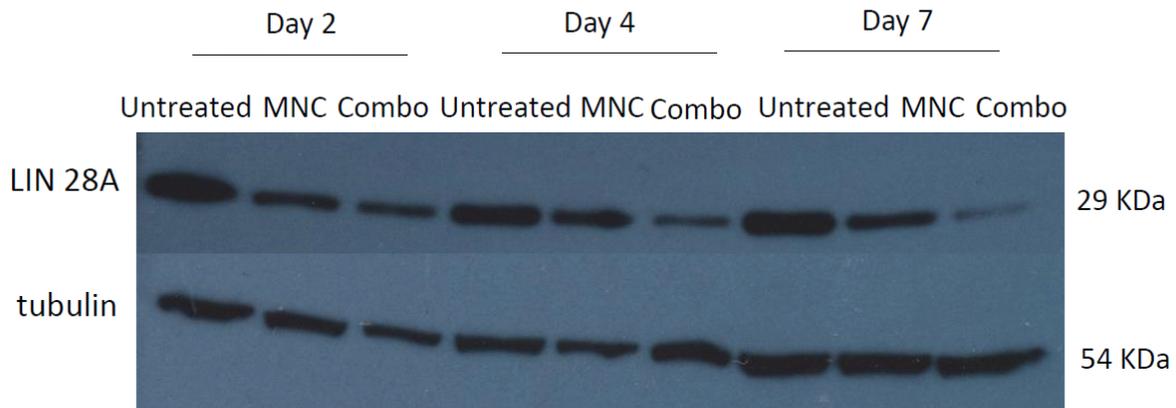
Figure 5.27: **Western blot of LIN 28A**. Levels of LIN 28A in TCam2 cell lines were tested at day 2, 4 and 7 post transfections with Combo or MNC.

As shown in Figure 5.27, the effects of the transfection of the mimic combination on LIN28 is greater at day 7 than at earlier timepoints. The effects of the mimics alone or in combination on transcription of LIN28 was therefore further investigated.

5.2.10.1   **Level of let-7 after miR-125b-5p transfection**

To further study the potential contribution of miR-125b-5p to the let-7 / LIN28 feedback loop, cells were treated with mimics either alone or in combination and levels of transcription of let-7 family members assessed over the course of six days (Figure 5.28).
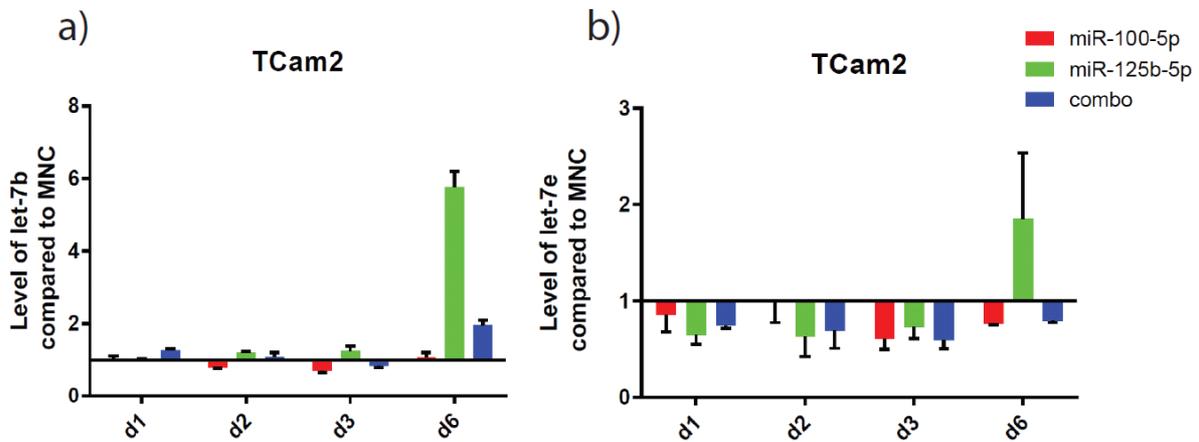
Figure 5.28: **Effects of miRNA replenishment on of let-7 family members in TCam2 and 1411H cells.** Gene expression was determined by RT-qPCR. a) let-7b and b) let-7e expression levels were examined at day 1, day 2 day 3 and day 6 post transfection. Replenishment of miR-100-5p is shown in red, miR-125b-5p in green and the Combo (combination of miR-100-5p+miR-125b-5p) in blue. Relative gene expression was normalised using RNU24 and independently calculated expression ratios were then averaged. Data presented as mean ± SEM of n=3 biological replicates.

The levels of two *let*-7 family members (*let*-7b and *let7*-e, Figure 5.28 a and b, respectively) were shown to be upregulated at day 6 following replenishment of miR-125b-5p alone. This suggests that the decrease in the amount of available LIN28 at day 7 post replenishment of miRNAs in GCT cell lines may be mediated by upregulation of *let*-7 transcription. All together the data presented here suggest a role for miR-125b-5p in the regulation of the levels of let-7 that may in turn influence the levels of LIN28. More broadly, the data presented has also shown that the transcription of a variety of genes involved in the regulation of the cell cycle are indeed affected by the replenishment of the miRNAs of interest, suggesting a potential role for these

genes in the phenotypic change in cell growth observed after transfection of the Combo in mGCT cell lines.

5.2.11 **FACS analysis of cell cycle arrest**

As a number of genes shown to be deregulated by the replenishment of miR-100-5p and miR-125b-5p are involved in the cell cycle, further experiments were carried out to determine whether cell cycle arrest could be the potential cause of reduced cell growth following transfection of the Combo in mGCT cell lines. In particular, I aimed to determine at which stage of the cell cycle a potential arrest was occurring. Incorporation of EdU (5-ethynyl-2'-deoxyuridine) into newly synthesised DNA was utilised to obtain information about the cellular DNA content that was then used to decipher the particular stage of the cell cycle (G1, S, and G2/M) cell were at, with the hypothesis that a block at a particular stage would have increased the percentage of population found at that stage itself.

EdU is a thymidine analogue that is incorporated into DNA during the synthesis phase (S phase) and has a high intracellular penetration capability due to its small size. It is added directly to the media of cells in culture, but conditions need to be optimised for different cell lines. A time course was utilised to establish the optimal time of exposure and concentration of this reagent for mGCT cell lines. In optimised experiments, incorporation of EdU occurred at a concentration of 10 μM, over a 2 hrs period, which is the time required for the EdU to be incorporated into the DNA during active DNA synthesis (Figure 5.29).
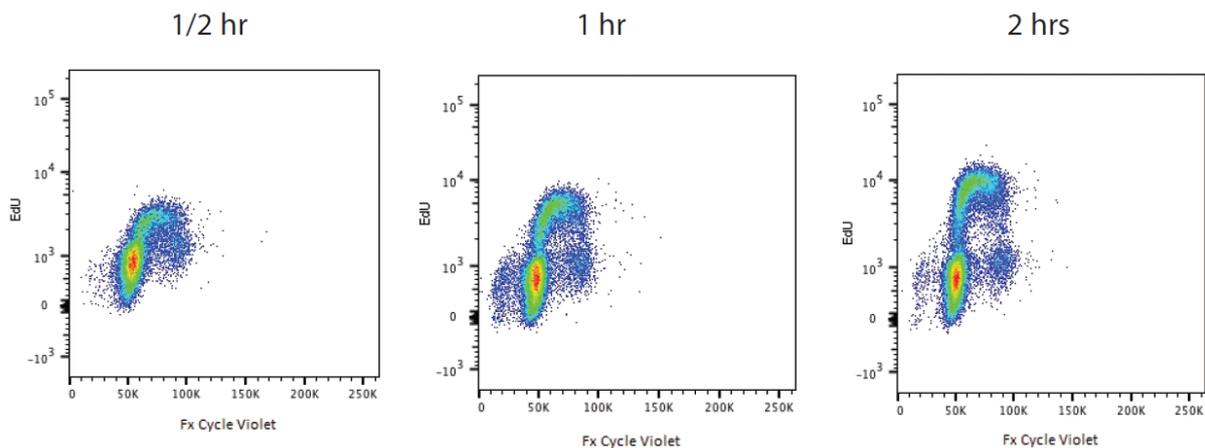
Figure 5.29: **EdU incorporation.** Example of EdU incorporation at three different time points in TCam2 cell line, showing the different populations that can be identified by this method. The x-axis shows the Fx-Cycle Violet used to detect DNA content, while the y-axis shows the EdU incorporation.

Further optimisation was carried out to determine the best way to select (i.e. gate) the cells of interest. In particular cell were selected as 'alive' and in a single cell suspension. This is particularly important to avoid misjudging the amount of intracellular DNA (Figure 5.30).
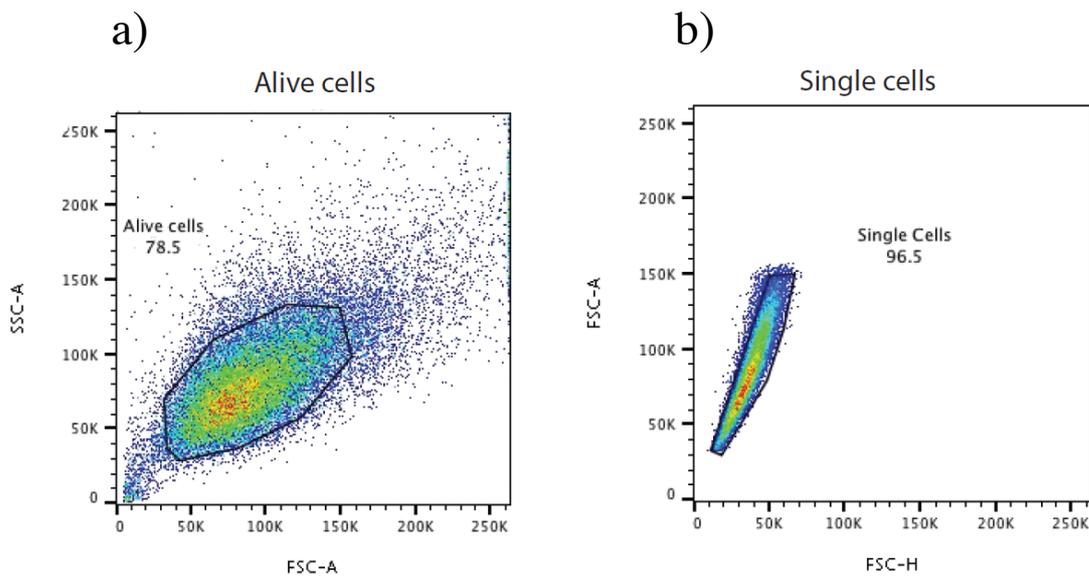
Figure 5.30: **Examples of cell selection (i.e. gating) needed before cell cycle analysis.** a) Along the X-axis is the FSC (Forward SCatter) parameter. This parameter is a measurement of the amount of the laser beam that passes around the cell. This gives a relative size for the cell. Along the Y-axis is the SSC (Side SCatter) parameter. This parameter is a measurement of the amount of the laser beam that bounces off the particulates inside of the cell. This gives information about the granularity. Cellular debris are usually low in the FSC parameters and dead cells have high SSC parameters. The graphs show a gated 'alive' population that fulfils the requirements for further analysis b) Doublets discrimination graph. Doublets can be distinguished from single cells by plotting FSC height vs FCS area. Doublets have increased area whilst similar height to single cells. The gate shows a typical single cells population that fulfils the requirement for analysis.

Once cells were appropriately gated, further analysis was carried out by FACS. TCam2 cells were collected for analysis at day 1 (Figure 5.31), day 2 (Figure 5.32), day 3 (Figure 5.33) and day 4 (Figure 5.34) after miRNA replenishment. For these experiments two concentrations of mimics were utilised, 16.7 nM and 33 nM. This was done to test whether a greater dose of mimic may result in a larger proportion of cells showing phenotypic changes.

Figure 5.31: **FACS analysis of TCam2 cells one day after transfection with miRNAs mimics**. Cells were either untreated or treated with mimic non-targeting control (MNC) or the combination of miR-100-5p and miR-125b-5p (Combo) at two different concentration (16.7 nM and 33.3 nM). a) schematic view of the different stages of the cell cycle: G1/G0 phase is shown in red, S phase in blue and G2/M phase in green. b) FACS acquisitions defining different stages of the cell cycle. Fx-Cycle Violet emission used to detect DNA content is shown on the x-axis while EdU incorporation is shown on the y-axis c) histogram showing DNA content. The experiments were repeated 3 times, a representative example is shown. The numbers relate to 'alive population' (parent).

At day 1 after miRNA replenishment, no differences were seen between MNC and Combo at both the concentration. A small decrease in the percentage of cells in S phase was seen in cells treated with MNC or Combo, compared with untreated control cells, probably due to toxicity associated to the transfection procedure which utilises Lipofectamine RNAiMAX reagent. Cells usually recover one day after transfection (data not shown).

Figure 5.32: **FACS analysis of TCam2 cells two days after transfection with miRNAs mimics**. Cells were either untreated or treated with mimic non-targeting control (MNC) or the combination of miR-100-5p and miR-125b-5p (Combo) at two different concentration (16.7 nM and 33.3 nM). a) schematic view of the different stages of the cell cycle: G1/G0 phase is shown in red, S phase in blue and G2/M phase in green. b) FACS acquisitions defining different stages of the cell cycle. Fx-Cycle Violet emission used to detect DNA content is shown on the x-axis while EdU incorporation is shown on the y-axis c) histogram showing DNA content. The experiments were repeated 3 times, a representative example is shown. The numbers relate to 'alive population' (parent).

At day 2 after miRNA replenishment, no significant differences were seen between untreated cells and cells treated with MNC at either 16.7 nM or 33.3 nM, suggesting that cells had recovered from the slight toxicity associated with the transfection procedure. An 8.1% increase in G0/G1 and a corresponding 7.9% decrease in S phase populations was seen in cells treated with Combo 16.7 nM compared to MNC 16.7 nM. A similar 7% increase in G0/G1 and a 7.8% decrease in S phase was obtained in cells treated with Combo 33.3 nM compared with MNC 33.3 nM. No significant difference was seen in G2/M phase fractions between MNC and Combo.
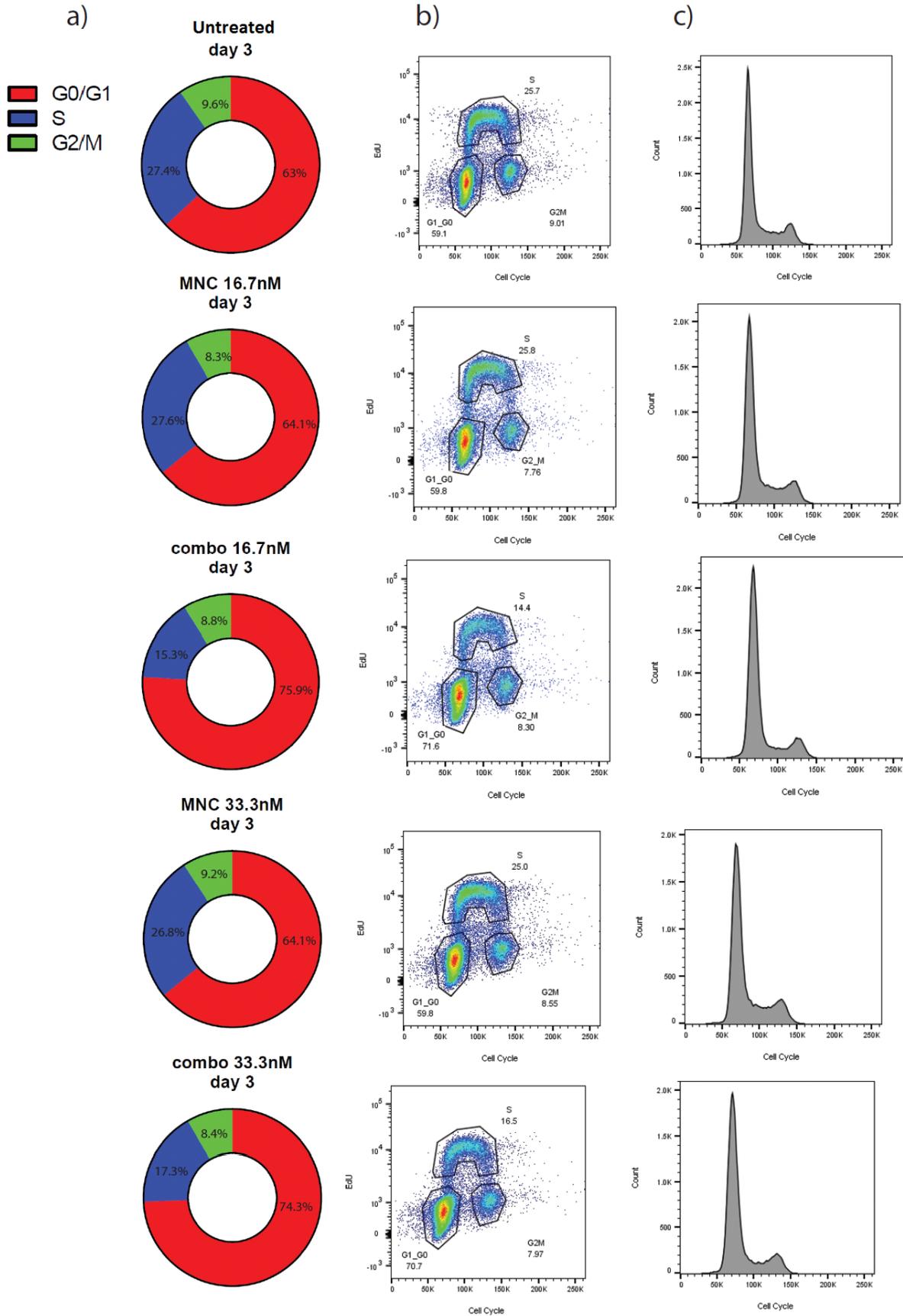
Figure 5.33: **FACS analysis of TCam2 cells three days after transfection with miRNAs mimics.** Cells were either untreated or treated with mimic non-targeting control (MNC) or the combination of miR-100-5p and miR-125b-5p (Combo) at two different concentration (16.7 nM and 33.3 nM). a) schematic view of the different stages of the cell cycle: G1/G0 phase is shown in red, S phase in blue and G2/M phase in green. b) FACS acquisitions defining different stages of the cell cycle. Fx-Cycle Violet emission used to detect DNA content is shown on the x-axis while EdU incorporation is shown on the y-axis c) histogram showing DNA content. The experiments were repeated 3 times, a representative example is shown. The numbers relate to 'alive population' (parent).

At day 3 after miRNA replenishment, no significant differences were seen between untreated cells and cells treated with MNC at either 16.7 nM or 33.3 nM. An 11.8% increase in G0/G1 cellular fraction and 12.3% decrease in S phase population was oserved in cells treated with Combo 16.7 nM compared with MNC 16.7 nM. A similar 10.2% increase in G0/G1 fraction and a 9.5% decrease in S phase population was seen in cells treated with Combo 33.3 nM compared with MNC 33.3 nM. No significant differences were seen in the G2/M phase fraction when comparing MNC and Combo.
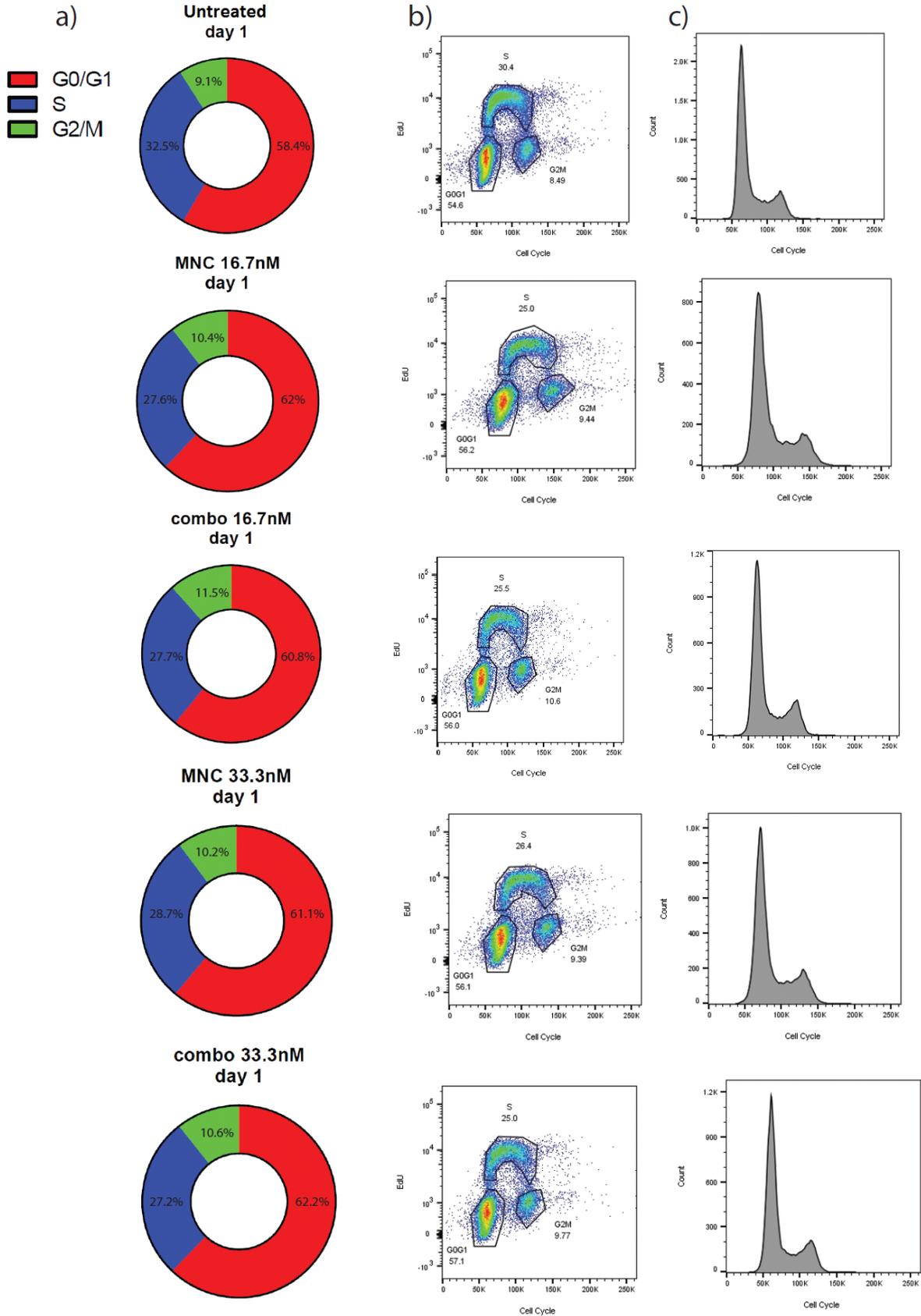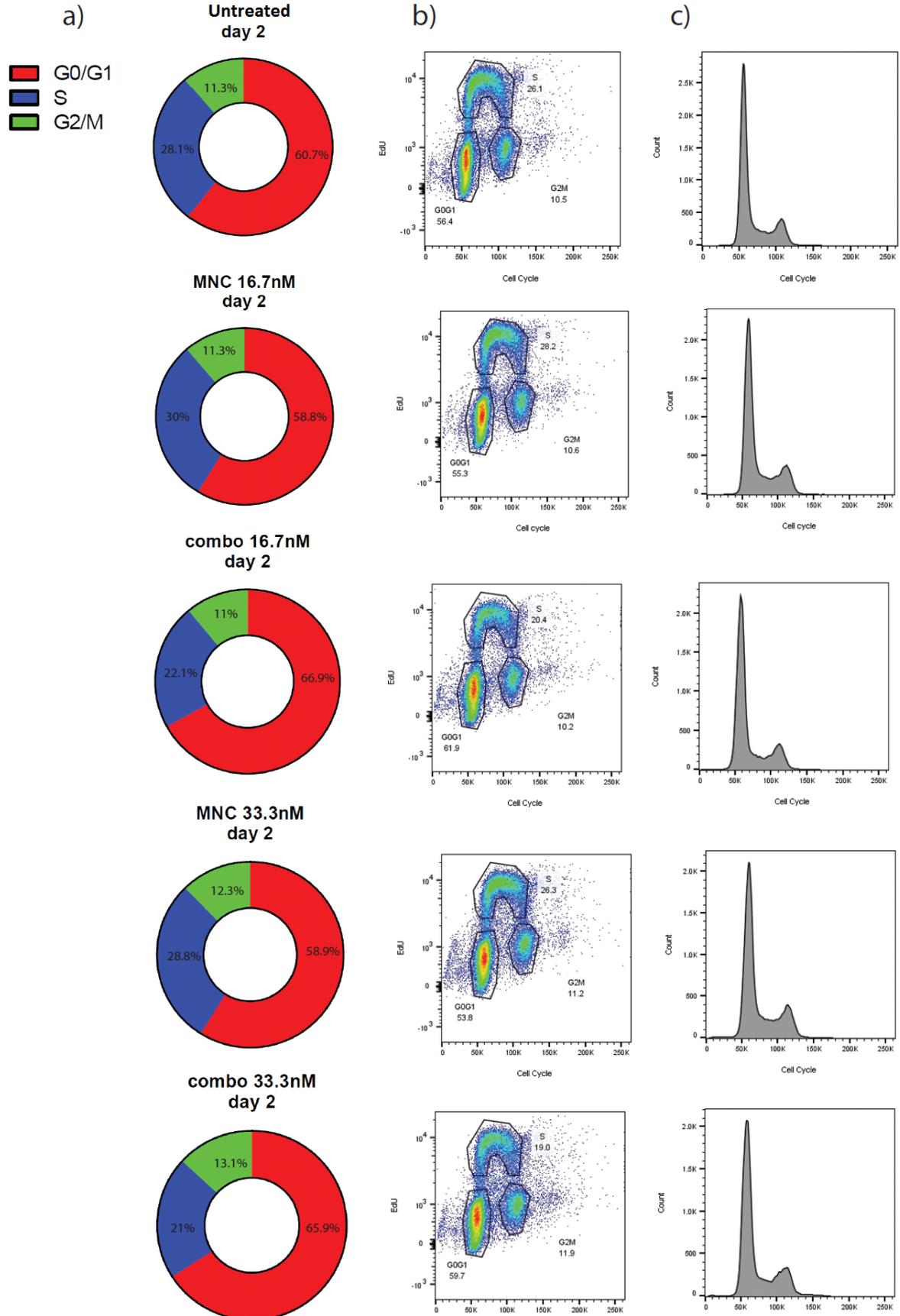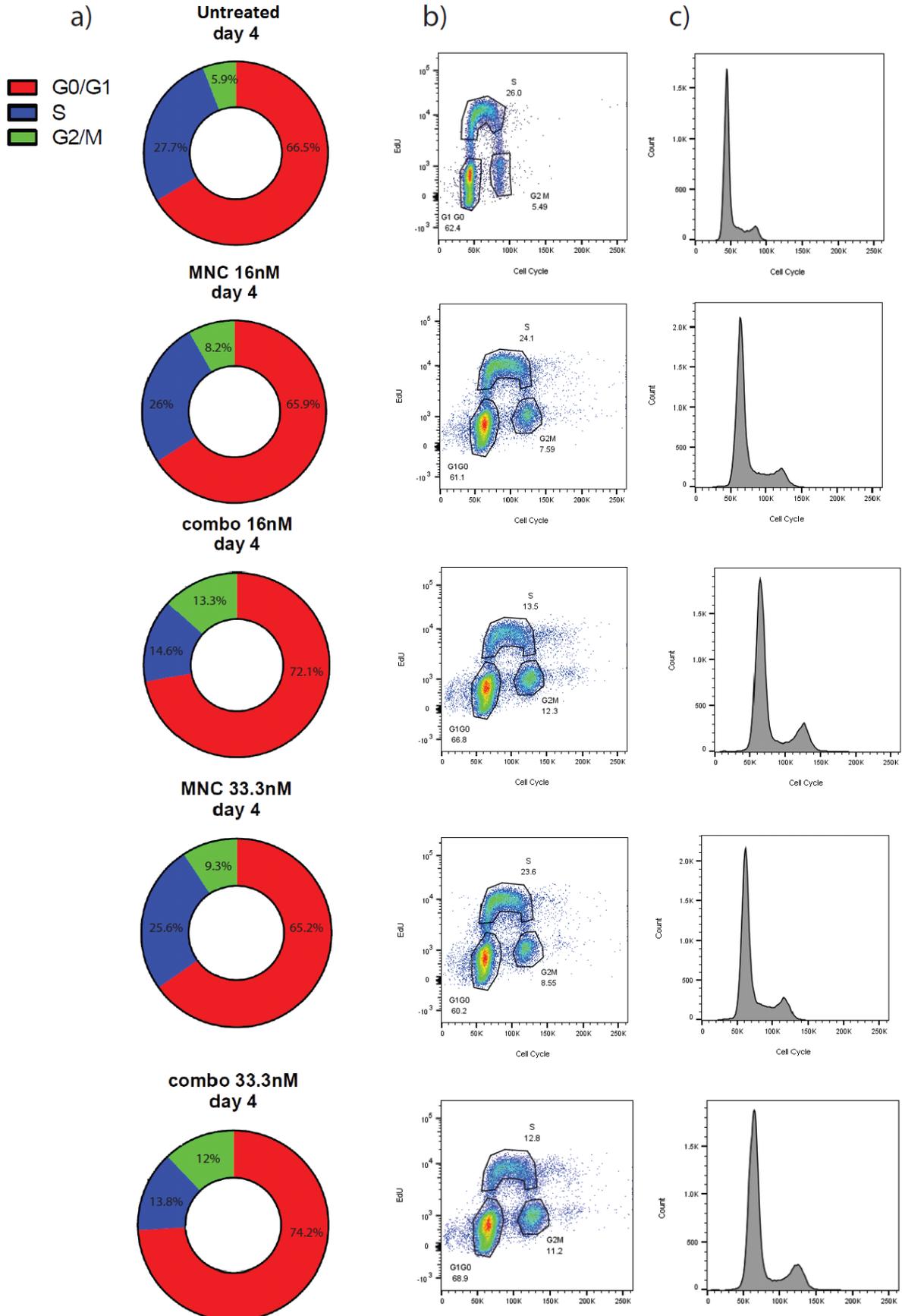
Figure 5.34: **FACS analysis of TCam2 cells four days after transfection with miRNAs mimics**. Cells were either untreated or treated with mimic non-targeting control (MNC) or the combination of miR-100-5p and miR-125b-5p (Combo) at two different concentration (16.7 nM and 33.3 nM). a) schematic view of the different stages of the cell cycle: G1/G0 phase is shown in red, S phase in blue and G2/M phase in green. b) FACS acquisitions defining different stages of the cell cycle. Fx-Cycle Violet emission used to detect DNA content is shown on the x-axis while EdU incorporation is shown on the y-axis c) histogram showing DNA content. The experiments were repeated 3 times, a representative example is shown. The numbers relate to 'alive population' (parent).

At day 4 after miRNA replenishment, no significant differences were seen between untreated cells and cells treated with MNC at either 16.7 nM or 33.3 nM. A 6.2% increase in G0/G1 population, an 11.4% decrease in S phase fraction and a 5.1% increase in G2/M populations was observed in cells treated with Combo 16.7 nM compared with MNC 16.7 nM. A similar 9% increase in G0/G1 fraction, an 11.8% decrease in S phase population and a 2.7% increase in G2/M were seen in cells treated with Combo 33.3 nM compared with MNC 33.3 nM.

All together the data presented here suggest that the phenotypic effect seen on cell growth in TCam2 cells after replenishment of miRNAs is indeed associated with a small but significant decrease in the S phase at day 3 (*p*-value: 0.05). This primarily results in the accumulation of cells in G0/G1 phase, however, a small proportion of cells also rests in G2/M phase. It is unclear from this data if cells are temporarily arrested in these phases and then resume cycling, or whether these blocks result in the induction of apoptosis or other processes that ultimately lead to cell death.

5.3   Discussion

In the present study, we aimed to investigate the contribution of miR-99a-5p/-100-5p and miR-125b-5p in the progression of mGCTs. Replenishment of miR-99a-5p/-100-5p and miR-125b-5p alone or in combination, have been previously shown to reduce tumour proliferation in breast cancer, hepatocellular carcinoma, oral squamous cell carcinoma and colorectal cancer (Henson *et al.*, 2009; Akhavantabasi *et al.*, 2012; Jia *et al.*, 2012; Wang *et al.*, 2018). However, their role in mGCTs is still unknown.

Data presented in Chapter 3, points strongly to the fact that both families of miRNAs are coregulated in these tumours. Furthermore, overall expression of miR-99a-5p/-100-5p and miR-125b-5p was similar in three representative cancer cell lines, suggesting equal contribution of these miRNAs in the overall downregulation. From these studies, we hypothesised that miR-99a-5p and miR-100-5p work together and in tandem with miR-125b-5p and the reintroduction of both miRNA mimics at the same time (named Combo) was achieved.

These miRNAs were then replenished in several mGCTs cell lines and the effect of such replenishment analysed. Results showed that restoration of miRNAs levels using both miRNAs mimics at the same time (Combo) resulted in a more pronounced phenotypic effect compared with transfection of single miRNAs. In particular, no significant phenotypic change on cell growth was observed following transfection with mimics for miR-100-5p alone or miR-125b-5p alone at 16.7 nM each, whereas, when both mimics were transfected individually at lower doses (8.3 nM + 8.3 nM each) to achieve a final concentration of 16.7 nM a significant phenotypic change on cell growth was observed. This observation has implications for the potential use of mimic combination to replenish downregulated miRNAs as a therapeutic approach, as lower doses are generally beneficial to reduce off-target toxicity. However, in view of the redundancy between these miRNAs, it cannot be excluded that the combination of the miRNA mimics may have additional off-target effects, in a manner similar to their enhanced activity. Further studies assessing potential additive or synergistic effects of these mimics' combinations in a detailed manner would have to be carried out to draw firmer

conclusions. Interestingly, when higher miRNA concentrations were used (up to a final concentration of 66.6 nM), additional phenotypic effects were not achieved. I then decided to continue my experiments using 16.7 nM of total concentration as adding high concentration of mimics can resulted in off-target effect and presence of high molecular weight RNA (Jin *et al.*, 2015).

In TCam2 and 1411H cell lines, which were derived from a testicular seminoma and yolk sac tumour respectively, replenishment of the miRNA resulted in a phenotypic effect on cell growth at day 7 (Bian *et al.*, 2012), while in 2102Ep cells, which were derived from an embryonal carcinoma, no changes in growth were seen, even at higher transfection concentration.

Embryonal carcinomas are known to be the most aggressive and the most undifferentiated of the non-seminomatous mGCTs. Seminomas, which are only of one histological type and are considered to be relatively undifferentiated, are associated with a better clinical prognosis and are highly sensitive to chemotherapy. In contrast, non-seminomatous GCTs have a worse prognosis and in general are thought to be more resistant to systemic chemotherapy and are therefore treated more aggressively.

In order to understand the reason why 2102Ep cells were resistant to the mimic treatment, different experiments were performed. First of all, labelled mimic was utilised to visually examine mimic uptake by the cells. In the first few hours following transfection, 2102Ep cells incorporated the labelled mimic with similar efficiency to TCam2 or 1411H cells. Interestingly, after 12 hrs, it appeared that almost all labelled mimic disappeared from 2102Ep cells (data not shown). This was not observed in the other two cell lines. These initial findings indicated that the mimic may be actively transported out of 2102Ep cells. Recently, a number of studies have highlighted the importance of exosomes in drug-resistance (Zhang *et al.*, 2015; Samuel *et al.*, 2018), so, accordingly, I investigated the potential involvement of exosomes in the GCT system. The first step was to measure the levels of the miRNAs mimics packages into exosomes (Appendix 1). I noticed that when 2102Ep cells were transfected, only a small proportion of the mimic remained inside the cells; almost all of the remaining mimic was found inside the cells' exosomes from day 1 post transfection. By comparing TCam2 and 1411H with 2102Ep

cells I could see a significant difference between the level of mimics packaged into 2102Ep exosomes compared with TCam2 and 1411H exosomes. This preliminary result gave me more confidence to study exosome involvement in drug resistant more extensively. Unsuccessful attempts to knock-out genes involved in the production of exosomes (nSMase and RAB27A), using CRISPR/Cas9 and small interfering RNA (siRNA), were performed on 2102Ep cells (Alexander *et al*., 2017; Menck *et al*., 2017) (Appendix 1). 2102Ep cell lines are highly undifferentiated; to maintain this state of undifferentiation cells have to be grown in their optimal condition. These cells have little cytoplasm and large nuclei and grow in tight clusters. Their undifferentiated state could only be maintained when cells were seeded at a density of at least 5×10^6 per 75 cm^2 flask. At lower densities the larger, flatter cells began to appear and the cells start to lose their undifferentiated state (Andrews *et al*., 1982). For this reason, it was hard, if not impossible, to select single clones necessary to perform CRISPR/Cas9. A heterogeneous population was initially selected due to the loss of genomic nSMase gene, but after a few passages the gene reappeared and the protein levels of nSMase increased. Due to these challenges, future studies should focus on drugs that inhibit exosome production.

To elucidate the mechanisms that led to the phenotypic effect on cell growth seen in the mGCT cell lines studied, gene expression was quantified using mRNA isolated at day 7 post miRNAs replenishment on Illumina mRNA microarrays. Three cell lines - TCam2, 1411H and 2102Ep – untreated or treated with a mimic non-targeting control (MNC) or the combination of miR-100-5p and miR-125b-5p mimics (Combo) were studied.

As expected, no gene expression changes were seen after replenishment with the Combo compared with MNC in 2102Ep cells. On the other hand, TCam2 and 1411H treated with the Combo showed differential gene expression when compared with MNC. Fold change analysis showed significant differentially expressed genes for both TCam2 and 1411H. Fewer differentially expressed genes were found when untreated cells were compared with MNC, but these changes were smaller with much lower significance (*p*-value). These analyses informed on the magnitude of the changes in expression resulting from the replenishment of miRNAs.

Further analysis carried out using *Sylamer* aimed to identify genes with Seed Complementary Regions in their 3'UTR corresponding to miR-99a-5p/-100-5p and miR-125b-5p that were differentially expressed. When examined independently, no significant curve was seen for the hexamers, heptamers and the octamer representing the SCRs of interests, for either 1411H or TCam2. However, if all SCR permutations complementary to a specific microRNA were tested together, a stronger significance was seen, particularly for genes containing SCRs to miR-125b-5p. Genes containing SCR to miR-100-5p did not appear to be amongst the most downregulated genes. This is interesting as the quantification of the intracellular mimics suggest that both miRNAs were replenished to a similar level, with miR-125b-5p being slightly less abundant at day 7 in 1411H cells. On the other hand, at day 7 post transfection, the effects of miR-100-5p and miR-125b-5p on target genes is lower than at earlier time points, as the amount of intracellular mimic decreased with time and the majority of the deregulated genes are likely to be downstream of the direct miRNAs' targets. It is likely that this differential significance is an artefact of this late timepoint and bears no biological significance.

An arbitrary cut-off corresponding approximately to the highest peak of downregulation of genes targeted by miR-125b-5p (and miR-100-5p, although the peak for these genes did not reach significance) was selected. The genes on the left of the peak were the most significantly downregulated and consequently were considered the most significant genes in terms of biological effects. As mentioned above, at day 7 the genes deregulated would probably be downstream the genes directly targeted by miR-125b-5p and miR-100-5p, therefore pathway analysis was carried out using the gene list derived from *Sylamer* analysis as well as the ranked gene list derived from the differential expression analysis of the microarray data using as baseline 0.5 fold change with a *p*-value of 0.05 for TCam2 cells and 0.3 fold change with a *p*-value of 0.05 in 1411H cells. Different baselines (or cut-offs) were applied to the two cell lines as the number of differentially expressed genes was lower in 1411H than TCam2 cells and this allowed the analysis of a sufficient number of genes to be able to deduce pathways of importance. Similar pathways were identified in both cell lines, such as cell cycle regulation, and apoptotic process. These pathways are consistent with the phenotypic changes seen in the cell lines after replenishment of the miRNAs. Others, such as transport of small molecules were found only TCam2 cells and may be more relevant to this particular mGCT subtype.

Unfortunately, the level of deregulation was very low, and it was impossible to validate these findings experimentally by RT-qPCR and western blotting.

Having elucidated some of the pathways influenced by the replenishment of miR-125b-5p and miR-100-5p seven days post transfection, it then was important to narrow down the genes that were the initial source of the observed changes. These were partially identified by the day 7 arrays, however, identification of direct targets of these miRNAs was challenging as their targets may in turn have instigated downstream signalling. Moreover, a reduction of the amount of mimic present intracellularly at this timepoint would have reduced interactions observed between these miRNAs and their targets. Therefore, a second set of arrays were utilised to study differential regulation of genes by the replenished miRNAs at a much earlier timepoint after transfection (day 2). The 2102Ep cell line was not included in this second study, as no phenotypic effect or significant gene expression changes were identified in this cell line at day 7. Moreover, as untreated control and cells treated with the mimic non-targeting control showed some minor phenotypic differences 7 days post transfection but no significant transcriptional changes and did not shown any phenotypic differences at this early timepoint, the untreated control was also excluded from the experiment.

The resulting analysis identified numerous differentially expressed genes at day 2 post transfection in both TCam2 and 1411H cells treated with the mimics' combination (Combo). In both 1411H and TCam2 cells, downregulation of genes containing miRNA seed complementary regions of variable length (hexamers, heptamers or octamer) for either miRNA was significant. When the effect of all SCR elements was considered simultaneously, a higher significance was demonstrated. The peak of significance for both miRNAs mimics was utilised as cut-off, and the genes on the left were used for pathway analysis using the online browsers: KEGG, gene ontology (GO) and reactome. More target genes were significantly enriched in the SCR of miR-99a-5p/-100-5p and miR-125b-5p at day 2 post treatment than at day 7. Pathways involved in cell cycle regulation, apoptotic processes, metabolic processes and immune system were identified as significantly affected by the miRNAs replenishment. Multiple genes involved in cell cycle regulation were identified as downregulated in both 1411H and TCam2 gene lists. To further validate the importance of these genes in clinical

setting, data from GEPIA (Gene Expression Profiling Interactive Analysis) was utilised. GEPIA is a web-based tool that combines TCGA and GTEx data (Tang *et al*., 2017) and can be used to study gene expression in mGCTs clinical samples and compared with normal tissue. The genes identified as regulators of the cell cycle which were downregulated in both TCam2 and 1411H cells were therefore investigated in this database. Interestingly, the vast majority of the genes identified by *Sylamer* analysis were found to be upregulated in tumour samples compared with normal controls. This suggests that the mechanisms observed *in vitro* may also be relevant in clinical settings, irrespective of tumour subtype. The discovery of genes that are commonly deregulated in different mGCT cell lines and in tissue samples from a variety of mGCTs suggest that a common mechanism of deregulation may exist in these tumours and opens the possibility of a single therapeutic approach for these highly heterogeneous cancers.

To validate the pathways and genes identified by microarray gene expression analysis, genes deregulated in both cell lines after transfection with miR-100-5p and miR-125b-5p and identified as having a role in the regulation of the cell cycle were further investigated. In particular *ARID 3A*, *E2F7*, *TRIM71*, *FGFR3* and *CTDSPL* were identified by *Sylamer* analysis as being regulated by both miRNAs in TCam2 and 1411H cells and were also shown to be overexpressed in mGCTs compared with normal tissue. These five genes have been described as important regulator of the G1/G0 to S transition in the cell cycle.

ARID 3A has already been identified as a gene targeted by miR-125b-5p in B-cell acute lymphoblastic leukemia (B-ALL). It has also been shown to have a role in cancer progression and apoptosis (Puissegur *et al*., 2012). Similarly, CTDSPL has been identified as a target of miR-99b-5p, although this gene is generally thought to be a tumour suppressor gene (Zheng *et al*., 2012). However, overexpression of CTDSPL was shown to lead to changes in expression of ribosomal genes and genes involved in cellular migration and metabolism (Winans *et al*., 2017). Importantly, a truncated transcript of CTDSPL found in virus induced B-cell lymphomas promoted immortalization when overexpressed in primary cell culture (Winans *et al*., 2017).

FGFR3 is another known target of miR-100. It has been shown to be upregulated in bladder cancer and in oral squamous cell carcinoma (OSCC) as a consequence of the downregulation

of miR-100-5p (Henson *et al.*, 2009; Song *et al.*, 2010). FGFR3 is one of four members of the FGFR family of receptor tyrosine kinases that serve as cell surface receptors for the FGF ligands. Signalling by the FGFRs leads to the activation of multiple intracellular signalling pathways including the extracellular signalling-regulated kinase/mitogen-activated protein kinase (ERK/MAPK) cascade and PI3K signalling, which in turn play important roles in various biological processes including embryogenesis, differentiation, angiogenesis, cell proliferation, migration, and apoptosis (Bi, Jing and Cao, 2015).

TRIM71 has been shown to associate with 3′UTRs of many mRNAs and this association represses protein synthesis, accompanied by the decrease in mRNA abundance due to accelerated mRNA decay (Loedige *et al.*, 2013). In some instances, TRIM71 acts together with miRNAs to achieve this mRNA decay. For example, TRIM71 functions with the miR-290 and miR-302 families to repress Cdkn1a expression in embryonic stem cells (ESCs) (Chang *et al.*, 2012). Cdkn1a is the key inhibitor of the cyclin E-Cdk2 complex whose constitutive activity allows progression through the G1–S phase in ESCs. The repression of Cdkn1a expression in ESCs is thought to be key in the maintenance of their rapid proliferation. Interestingly, the miR-302 family of miRNAs which partners TRIM71 in this regulation of Cdkn1a is highly overexpressed in mGCTs.

The family of E2F transcription factors are known to be important regulator of the cell cycle. E2F7 is thought to be important in the deactivation of cell cycle-regulated genes. E2F7 is highly expressed during mid to late S-phase, occupies promoters of G1/S-regulated genes and represses their transcription. In particular, together with E2F8, E2F7 directly repress directly E2F1 with consequent anti-proliferative properties (Carvajal *et al.*, 2012). However, its ability to inhibit E2F1 also mediates the block of E2F1-induced apoptosis in response to DNA damage. These pro-survival actions of E2F7 suggest that, under certain circumstances, E2F7 may be oncogenic. Consistent with this idea, E2F7 has been reported to be selectively expressed in proliferation-competent keratinocytes and to be overexpressed in cutaneous squamous cell carcinoma (CSCC) (Endo-Munoz *et al.*, 2009).

Interestingly, while these 8 genes were all identified by *Sylamer* analysis to be targets of both miRNAs, not all of these genes were downregulated equally by miR-100-5p and miR125b-5p. TRIM71, E2F7 and FGFR3 were affected by the replenishment of both miRNAs, but ARID 3A was affected only by the replenishment of miR-125b-5p, while CTDSPL was affected only by the replenishment of miR-100-5p. Analysis of the 3'UTR sequences of ARID 3A and CTDSPL identified a different number of seed complementary regions to each individual miRNA in these regions. Interestingly, the miRNA with more SCR was indeed the one that the genes appeared to be regulated by. Moreover, alongside the higher number of SCRs for a particular miRNA, one or more of these SCRs would be characterised by a longer stretch of homology. It therefore appeared that as abundance and length of the seed complementary region for a specific miRNA (i.e. the mRNA binding site) increased, so did the extent of downregulation of the gene by that miRNA. A number of studies have looked at the relationship between miRNA binding sites in 3'UTRs and the effects on their target genes, as this is fundamental to understanding how miRNAs carry out their function. Unfortunately, some of this process is still poorly understood. Many tools are available to predict potential seed complementary regions, and most of them – such as such as PicTar, EIMMO or PACMIT - search primarily for seeds seven or eight nucleotides long, giving 8-mer SCRs more statistical weight (Ellwanger *et al*., 2011). On the other hand, a recent study on colorectal cancer has shown that the length of the SCR does not influence the level of downregulation mediated by the miRNA, as no additional downregulation was achieved in this system when a longer SCR was introduced (Mullany *et al*., 2016). The observation that despite the presence of binding sites for both miRNAs, greater transcript downregulation was achieved using the miRNA with a longer SCR sequence suggests that this may be a relevant factor. Manipulation of the SCRs to change miRNA affinity would be needed to strengthen this observation. The use of *Sylamer* to predict specific miRNAs target genes is more flexible than other available platforms as it allows the interrogation of SCRs of variable length (6-mers, 7-mers and 8-mers). Moreover, it allows analysis using both miRNA and mRNA expression data from matched samples, something that other platforms do not provide.

Reassuringly, the genes that were thought to be targeted by only one miRNA only responded to miRNA replenishment in the expected manner, supporting the validity of the *Sylamer*

analysis. mTOR, which is a well-known target of miR-100-5p in a variety of cancers such as ovarian cancer, responded only to replenishment with miR-100-5p mimics while ARID 3B and LIN28A which are targeted by miR-125b-5p (Nagaraja *et al*., 2010; Bobbs *et al*., 2015; Balzeau *et al*., 2017) responded only to replenishment with mimics of the latter.

Of interest was the investigation of the effects of the replenishment of miR-125b-5p (or the use of the Combo), on the expression of LIN28. The reason for this is that the upregulation of the expression of LIN28 (an RNA binding protein and known negative regulator of *let-7* biogenesis) in mGCT tissue samples was originally identified in the lab by microarray. LIN28 acts by blocking the process of maturation of precursor *let-7* molecules (Viswanathan, Daley and Gregory, 2008). Once LIN28 binds *pri-let-7* and *pre-let-7*, it prevents their processing by DROSHA and DICER1 (Viswanathan and Daley, 2010b). This interaction occurs due to a stem loop motif that includes the nucleotide sequence 'GGAG' and that in turn leads to the recruitment of a terminal-uridyl-transferase (TUTase), so that *pre-let-7* undergoes uridylation and degradation (Heo *et al*., 2009). In GCTs cell lines, *LIN28* knockdown (KD) by siRNA results in a specific upregulation of *let-7* family members, suggesting that LIN28 and *let-7* regulate each other via a negative feedback loop. Furthermore, replenishment of *let-7* levels in mGCT cell lines using a mimic *let-7* molecule (namely *let-7e*) resulted in reduced cell growth (as established by total cell number counts) associated with reductions in the levels of *let-7* targets such as *MYCN*. In addition, a reduction in LIN28 expression itself were also observed, demonstrating again a feedback loop, mediated by the presence of a *let-7* binding site in the *LIN28* 3'UTR (Murray *et al*., 2013). The observation that miR-125b-5p alone can partially restore *let-7* levels has potential clinical implications. Due to the regulatory loops described, the replenishment of a downregulated miRNA (miR-125b-5p), also has the additional benefit of restoring the level of another downregulated miRNA (*let-7*). The low LIN28A protein level found at day 7 can be explained by this additional effect of *let-*7 to that of miR-125b-5p.

To further confirm that the phenotypic effects on cell growth seen after replenishment of miR-100-5p and miR-125b-5p were mediated by a modulation of the cell cycle, FACS analysis was performed. These experiments demonstrated that the replenishment of miR-100-5p and miR-125b-5p indeed resulted in changes in cell cycle, the most significant of which was seen at day

3 post mimic transfection. At this timepoint, an increase of approximately 12% in the number of cells in G0/G1 phase and a 12% decrease in the number of cells in S phase was seen. This increase of the G0/G1 population is likely to be due to failure to pass the G1 checkpoint that monitors the DNA integrity and cell machinery. Cells that fail the G1 checkpoint due to DNA damage can initiate a process of DNA repair or initiate apoptosis. Cells can also fail the G1 checkpoint due to limited nutrients or extracellular signalling, in which case they can enter the G0 phase of the cell cycle, which describes a state of quiescence. A similar checkpoint is found in G2/M phase.

Interestingly, no further accumulation in G1/S was seen by doubling the number of mimics replenished. This was unexpected, however further work would be needed to investigate this lack of additional effect. It is in fact unclear if the additional mimic was indeed delivered intracellularly as no quantification of the intracellular mimics was performed.

Unfortunately, it was technically impossible to obtain reliable data on the cell cycle at timepoints later than day 5. This was due to the fact that GCT cell lines cannot be seeded sparsely and therefore by day 5 cells became too overconfluent (this was particularly true of untreated cells and those treated with MNC). Consequently, these cells started to show abnormal cell cycle distribution due to the fact that they were responding to inhibition by contact. Therefore, almost no cells in these control samples were seen in G2/M phase, whereas a large fraction of the population was blocked in the G1/G0 phase. Attempts were made to synchronise cells to see if this would improve the ability to determine where cells were blocked, however this was never successfully achieved, despite trying different methods including: serum starvation, double thymidine block and nocodazole. Taken together the cell cycle data shows that following replenishment with miR-100-5p and miR-125b-5p, TCam2 could no longer overcome the block in G1/G0. This is in agreement with the suggestion that the genes regulated by these miRNAs have an important role in the control of the cell cycle and that this is circumvented in these cancer cells by the removal of the miRNA mediated control of these genes.

Ideally the data on cell cycle regulation shown for TCam2 cells would have been confirmed using another cell line. Attempts were made to use the 1411H cells, which were extensively

utilised for other parts of this work. Unfortunately, while various attempts were made to optimise the protocol with these cells, none produced conclusive data. For reference data is shown in the Appendix 2. In particular EdU was not incorporated in the cellular DNA after treatment of 1411H cells. Different concentrations of EdU and different length of treatment were tested, but no sufficient EdU incorporation was achieved. It is possible that these cells do not express a pyrimidine salvage pathway that is key for the incorporation of EdU to occur. A salvage pathway is the pathway in which nucleotides are synthesized from intermediates in the degradative pathway for nucleotides. Without this pathway, the EdU is not phosphorylated intracellularly, a step essential for incorporation into replicating DNA to happen. Alternatively, it is possible that the permeabilisation step was insufficient to allow staining of the EdU. The kit that was utilised for this experiment (Click-iT EdU) only requires a mild detergent permeabilisation because of the size of the reagent, but it is possible, although unlikely, that a harsher permeabilisation was required.

Future work will be undertaken to look into the role of apoptosis and the immune system in mGCTs cell lines after miRNAs replenishment, as suggested from the pathway analysis data.

Chapter 6   Concluding discussion and future work

All malignant GCTs are thought to arise from the same progenitors (PGCs), however, until recently, no common biological abnormality present in these tumours independently of patient age (paediatric or adult), histological subtype (YST, germinoma/seminoma or EC) or anatomical site (gonadal or extragonadal) had been identified. Members of the Coleman lab, have identified that mGCTs could be segregated from non-malignant controls by their miRNA expression profile, thus identifying for the first time a biological abnormality shared by all mGCTs (Palmer *et al*., 2010).

This doctoral thesis originates from these findings and aimed to investigate the mechanism and functional significance of the downregulation of miRNA expression in mGCTs. This presented work covers cell biological studies to understand the upstream regulation of miRNA expression, and the identification of mechanisms underlying the malignant effects of miRNA downregulation in mGCTs by testing the consequence of miRNA replenishment *in vitro*. Ultimately, a better understanding of miRNA dysregulation in mGCTs may provide insight to therapeutic approaches targeting miRNAs in the future.

The material described and discussed in Chapter 3 can be divided into two sections.

In the first section, data from the microarray previously performed in the lab (Palmer *et al*., 2010) was utilised to analyse the individual expression data for each of the five most downregulated miRNA (miR-125b-5p, miR-100-5p, miR-99a-5p, miR-204-5p and miR-152-3p) in mGCT samples and representative cell lines, and compare it to expression in non-malignant controls (gonadal controls and benign teratomas) (Palmer *et al*., 2010). This analysis showed that miR-99a-5p/-100-5p and miR-125b-5p were the only miRNAs likely to have a significant functional effect in mGCTs as their expression level was consistently downregulated across all mGCT samples and representative cell lines. Moreover, the 2-7 nt

seed sequences of miR-99a-5p/-100-5p and miR-125b-5p was also substantially decreased in both clinical malignant GCT samples and representative GCT cell lines, compared with non-malignant control samples. Once the downregulation of miR-99a-5p/-100-5p and miR-125b-5p in mGCTs was validated using the array data and seed abundance analysis, the chapter moved to a second section that aimed to understand the mechanism by which this downregulation occurs. The microarray data suggested that miR-99a-5p/-100-5p and miR-125b-5p are co-regulated as implied by their similar pattern of downregulation. Additionally, miR-99a-5p/-100-5p and miR-125b-5p are physically linked at genomic level. MiR-125b-5p is the product of two genomic loci, each of which is clustered with either miR-99a-5p or miR-100-5p. The cluster on chromosome 11q24.1 contains miR-100 and miR-125b-1, while the cluster on chromosome 21q21.1 contains miR-99a and miR-125b-2. The similar pattern of regulation and location of these miRNAs suggested a co-regulation in their expression, which was further explored in Chapter 3.

This second section of the Chapter 3 reported three main findings. Firstly, the lower level of primary transcripts (pri-miR-99a/-miR-100 and pri-miR-125b-2/-miR-125b-1) and host genes (MIR99AHG and MIR100HG) found in mGCTs, compared with non-malignant control, indicated that miRNA downregulation occurs at the transcriptional level. Secondly, both chromosomes were shown to have a role in miRNAs downregulation. These studies used differences in the primary (pri-miRNAs) and the -3p strands (miRNA-3p) transcripts which made them specific to a chromosomal site to define the relative contribution of each chromosome to the downregulation of the common mature miRNA-5p. To further confirm this point, levels of the host genes, MIR99AHG and MIR100HG, were also assessed. The assessment of the transcription of *BLID*, a gene embedded in chromosome 11 between miR-100 and miR-125b-1, brought another interesting finding. This gene has its own promoter and its transcription is separate from the miRNAs', however its expression was found to be also low in mGCTs, suggesting that the mechanism underlying the downregulation of the miRNAs was unlikely to be miRNA specific. These two points together brought us to the third and last point discussed in Chapter 3. The strong correlation that was observed within chromosomes and between the two chromosomal loci by assessing primary miRNA, mature miRNA-5p, mature miRNAs-3p and lincRNA transcripts suggests that both miR-99a, miR-125b-2 on

chromosome 21, and miR-100, miR-125b-1 on chromosome 11, are regulated at the same time by a common process which results in broader changes in transcription.

The work presented in Chapter 4 of this thesis sought to identify the common mechanism through which miR-99a, miR-125b-2 on chromosome 21, and miR-100, miR-125b-1 on chromosome 11, are regulated. Chapter 4 focused on DNA methylation as a possible regulator of miRNAs expression in mGCTs. Initial experiments showed that miR-99a/-100-5p and miR-125b-5p expression increased in 2102Ep and 1411H cell lines upon treatment with a de-methylating agent - 5-azacytidine. This suggested that methylation had a suppressive role on the transcription of the miRNAs. Specific regulatory regions differentially methylated in mGCTs compared with non-malignant control were then identified using bisulfite conversion. Four mGCT cell lines, six normal cell lines and one normal tissue sample derived from the testis on a healthy patient were interrogated. Using a semi-quantitative analysis, five CpG islands (4 on chromosome 11 and 1 on chromosome 21) were shown to be statistically differentially methylated in mGCTs compared with non-malignant controls. Pyrosequencing was then used to quantify methylation at each individual CpG dinucleotide. No single CpG was identified as selectively methylated in mGCT samples however an overall increase in CpG island methylation was observed in mGCTs compared with non-malignant controls, suggesting an involvement of methylation in miRNAs downregulation. This overall increased methylation as illustrated by cumulative methylation graphs was seen on both chromosomes 11 and 21 and could cause the downregulation of the miRNAs transcription by potentially affecting the local chromatin structure, making it more condensed and hardly accessible. It is becoming apparent that DNA methylation and histone modification are dependent on one another and that this crosstalk can lead to changes in chromatin conformation (Cedar and Bergman, 2009). Other studies have shown that DNA methylation works together with histone methylation during replication to ensure that a correct methylation pattern is passed to the daughter cells (Sarraf and Stancheva, 2004). Moreover, the establishment of DNA methylation has been shown to be dependent on the H3K9 histone-methyl-transferase *dim-5* in a silencing pathway utilized by Neurospora crassa, (Tamaru and Selker, 2001). More recent studies are showing a direct link between DNA methylation and the histone lysine methylation system (Rose and Klose, 2014). Some of the proteins involved in CpG methylation belong to the MBD family. These proteins

are also associated with different chromatin-modifying enzymes including histone deacetylases and histone lysine methyl-transferases (Fuks *et al*., 2003). Another important class of proteins essential for DNA methylation of CpG sites is the family of DNA methyl-transferases (DNMTs). In mammals, there are three DNMTs: Dnmt1, a maintenance methyl-transferase and Dnmt3a/b, which are involved in *de novo* methylation. Several studies have linked H3K9 and H3K36 with *de novo* DNA methylation either through direct domain binding, such as the binding of H3K36 to the PWWP domain on the DNMT3a/b or indirect binding such is the case of H3K9 (Li *et al*., 2006; Ooi *et al*., 2007; Dhayalan *et al*., 2010; Sasidharan Nair *et al*., 2018). H3K9 and H3K36 are histone modification that regulate the structure of chromatin. H3K9 tri-methylation is used as a marker of epigenetic silencing and constitutive heterocromatin (Schotta *et al*., 2004; Stewart, Li and Wong, 2005). H3K36 comprises three states of methylation and was shown to be required for the formation of the highly condensed chromatin structure (Matsuda *et al*., 2015; Suzuki, Murakami and Takahata, 2017).

As part of future work, it will be important to investigate the relationship between the observed methylation and chromatine structure in mGCT cell lines. Chromatin immunoprecipitation (ChIP) is a well-established technique for *in vitro* detection of histone modifications (Milne, Zhao and Hess, 2009). It aims to determine which specific histone modifications are associated with specific genomic regions, informing on the transcriptional landscape of the genomic area. Performing ChIP analysis to characterise the area surrounding the region where the host genes and miRNAs are located will help elucidate the nature of the aforementioned repression. Antibodies specific for a variety of markers should be used to evaluate the relative presence of markers of active transcription (such as acetylated histone H3 and trimethylated lysine 4 of histone H3) versus those of transcriptional repression (dimethylated lysine 9 of histone H3 and dimethylated lysine 27 of histone H3).

Two main issues still need to be addressed with regards to the studies presented in Chapter 4. Firstly, concerns have been raised over using cell lines for methylation studies because cancer cell lines can develop aberrant methylation profiles as a consequence of the tissue culture process (Smiraglia *et al*., 2001). This concern is valid, and the findings here reported should be validated on clinical material, reassuring that the data presented, is in line with observations made on tissue samples. Secondly, the unbalanced methylation density between the two

chromosomes raises reservations about the hypothesised common mechanism of regulation. The area surrounding miR-99a/-125b-2 is more than 50Kb long, and not all CpG islands have been tested, it is possible that methylation is occurring in areas that have not been looked at. This hypothesis seems likely, as the area upstream the starting site of miR-125b-2 was heavily methylated. More importantly, downregulation of the pri-transcripts was very similar in both chromosomal loci, which again suggest that both loci are regulated in similar manner. ChIP analysis as previously mentioned will help to further elucidate this mechanism on chromosome 21. It cannot be excluded that other additional mechanism are responsible for the downregulation of the transcription of miR-99a-5p/-100-5p and miR-125b-5p. It is also possible that there may be chromosome specific additional regulatory pathways. The two loci are structurally similar and are likely to have arisen by gene duplication. During species evolution, different strategies for the regulation of these two clusters might have evolved. As an extra control to dictate their expression, not only epigenetic mechanisms, but also transcription factors, might be important. However, more needs to be done to locate the promoter of the miRNAs before studies looking at differential regulation by transcription factors can be carried out.

The work presented in Chapter 5 of this thesis explored and discussed the effect of miR-99a-5p/-100-5p and miR-125b-5p replenishment in mGCTs cell lines. MiR-99a-5p/-100-5p and miR-125b-5p replenishment caused the repression of multiple tumour suppressor genes such as *ARID 3A, ARID 3B, E2F7, TRIM71, FGFR3, CTDSPL, mTOR* and *LIN28*, resulting in cell cycle arrest in G0/G1. Most importantly, these deregulated genes have been shown to be upregulated in mGCTs patients, as reported from TCGA data combined with GTEx data, highlighting the importance of miRNA regulation as potential therapeutic agent for mGCTs patients.

Some of the genes listed and other identified by pathway analysis suggest that replenishment of miR-99a-5p/-100-5p and miR-125b-5p in mGCTs may also have pro-apoptotic effects. Unfortunately, though many experimental attempts were undertaken to demonstrate the possible activation of apoptotic pathways none of the data obtained was informative. As part of future work, different assays could be used. For example, Click-iT TUNEL Alexa Fluor, is an apoptosis assay based on the identification of fragmented DNA. It uses the same modified

nucleotides as the Click-iT reaction used for cell cycle detection. Since the cell cycle assay could be used with our cells, an apoptosis assay using the same chemical procedure is likely to work well. Alternatively, mitochondrial permeability based-assays could be used. These assays use of a dye that accumulates in healthy mitochondria and aggregates there as a multimer. Upon disruption of the mitochondrial membrane, the dye is released, changing colour due to monomerization as it spreads through the cytoplasm. These assays will be investigated in our system to further explore the potential role of apoptosis in mGCT cells where downregulated miRNAs have been replenished.

Another interesting effect obtained by the replenishment of miR-99a-5p/-100-5p and miR-125b-5p, was to change the level of expression of some members of the *let-7* family, another class of downregulated miRNAs in mGCTs. Briefly, replenishment of miR-125b-5p, by targeting LIN28 and suppressing its expression, resulted in increased levels of *let-7*, which in turns is downregulated by LIN28 (Murray *et al.*, 2013). The replenishment of miR-125b-5p, has an indirect effect on the level of expression of another deregulated tumour suppressor miRNA (*let-7*), helping the cell reverting its tumorigenic behaviour. This model is presented in the Figure 6.1. This result opens a new prospective for clinical practice, where one can envisage a more successful treatment using a carefully chosen miRNA targeting multiple downregulated pathways, some directly and some via another downregulated miRNA.
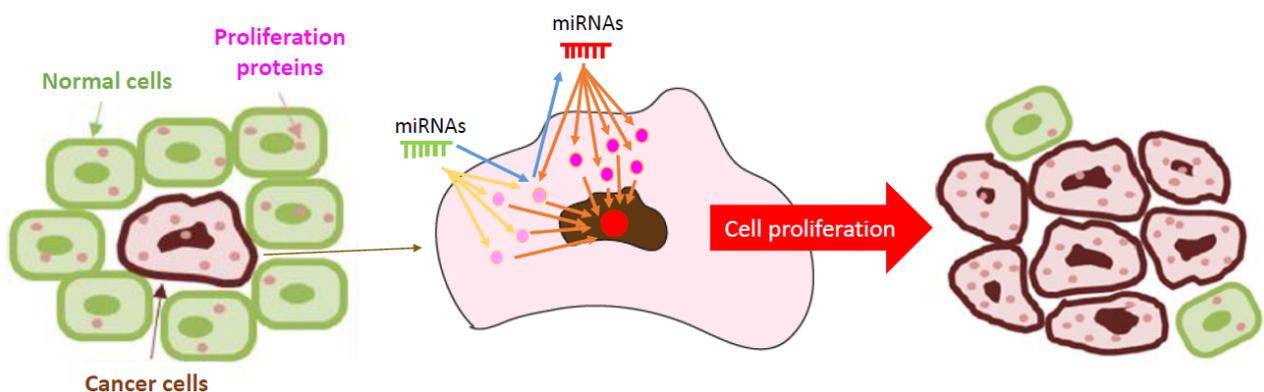

Figure 6.1: **Model of miRNAs regulation.** MiRNAs post-trascriptionally regulate the expression of protein-coding genes by directly bind to the 3'UTR region of mRNAs. A single miRNA can suppress the production of a multitude of mRNA. The targeted mRNA can then regulate other miRNAs, as seen from the miR-125b-5p-LIN28A-let-7 system described in Chapter 5. In mGCTs the downregulation of

miR-99a-5p/-100-5p and miR-125b-5p leads to cell cycle deregulation and uncontrolled cell proliferation, due to the up-regulation of their pro-proliferative target mRNA.

As part of the future work, which has already commenced in the lab, an *in vivo* system will now be utilized. Injection of TCam2 cell lines in CD1 nude mice in three different locations (in the flank, testicles and brain) has been previously shown to generate tumour (Nettersheim *et al.*, 2012). Depending on the location of injection, tumour were generated with a seminoma or embryonal carcinoma-like phenotype (Nettersheim *et al.*, 2012). Experiments to establish subcutaneous and lung models using TCam2 and 2102Ep cell lines in CD1-*Foxn1^{nu}* nude mouse (Charles River) are currently undergoing in the lab. Once the tumour model is being undertaken, a delivery system for the miRNAs needs to be selected. As a proof of principle, the first step could be to transfect TCam2 and 2102Ep cells *in vitro*, using the combination of mimics (Combo) and the mimic non-targeting control (MNC), inject the transfected cells into the flank of the mouse and follow their growth. However, this method is unlikely to give significant results as the effect of the mimics is short lived compared with the length of time needed to establish tumours. Delivery of mimics in vivo is problematic, and requires the use of high doses, which could result in toxicity. To overcome this problem transfection of TCam2 and 2102Ep using an inducible lentiviral system would be preferable and will be used in our studies. In detail, a Tet-inducible system (Tet-pLKO-neo) lentiviral plasmid will be used (Wiederschain *et al.*, 2009). Primary sequence of miR-100 and miR-125b-1 will be cloned into the plasmid using restriction enzymes. A non-targeting control plasmid will also be created by cloning a non-human pri-miRNA sequence. A lentiviral vector will be generated in 293T cells using the cloned plasmid together with a packaging and envelope vector (named psPAX2 and pMD2.G, respectively). The formed lentiviral vector will then be used to transduce TCam2 and 2102Ep cell lines. The expression of the miRNA will be induced by antibiotic treatment of the animals after tumour establishment, mimicking the activity of small molecule inhibitor in the clinical setting. To activate the plasmid, administration of doxycycline will be performed either by oral gavage or in drinking water.

The findings presented in this doctoral thesis further our understanding of the mechanisms and functions of downregulated miRNAs in mGCTs. This study contributed to identify a possible regulatory mechanism of miR-99a-5p/-100-5p and miR-125b-5p downregulation, which opened a new avenue in mGCTs studies. From the clinical perspective, the presented findings created the basis for an *in vivo* study, which hopefully will be brought forward in the future in clinical trials as targeted therapy. Taken together, these results might set the basis for the development of a targeted therapy which could reduce the toxicity effect due to chemotherapy treatment.

Chapter 7　　Appendix 1

a)

2102Ep cells
50,000 cells seeded @d-1; 8 hours transfection

1411H
100,000 cells seeded @d-1; 6 hrs transfection

b)

2102Ep

1411H

c)

2102Ep d1

1411H d1

level of miR-100-5p in MNC
level of miR-100-5p in combo
level of miR-125b-5p in MNC
level of miR-125b-5p in combo

d)
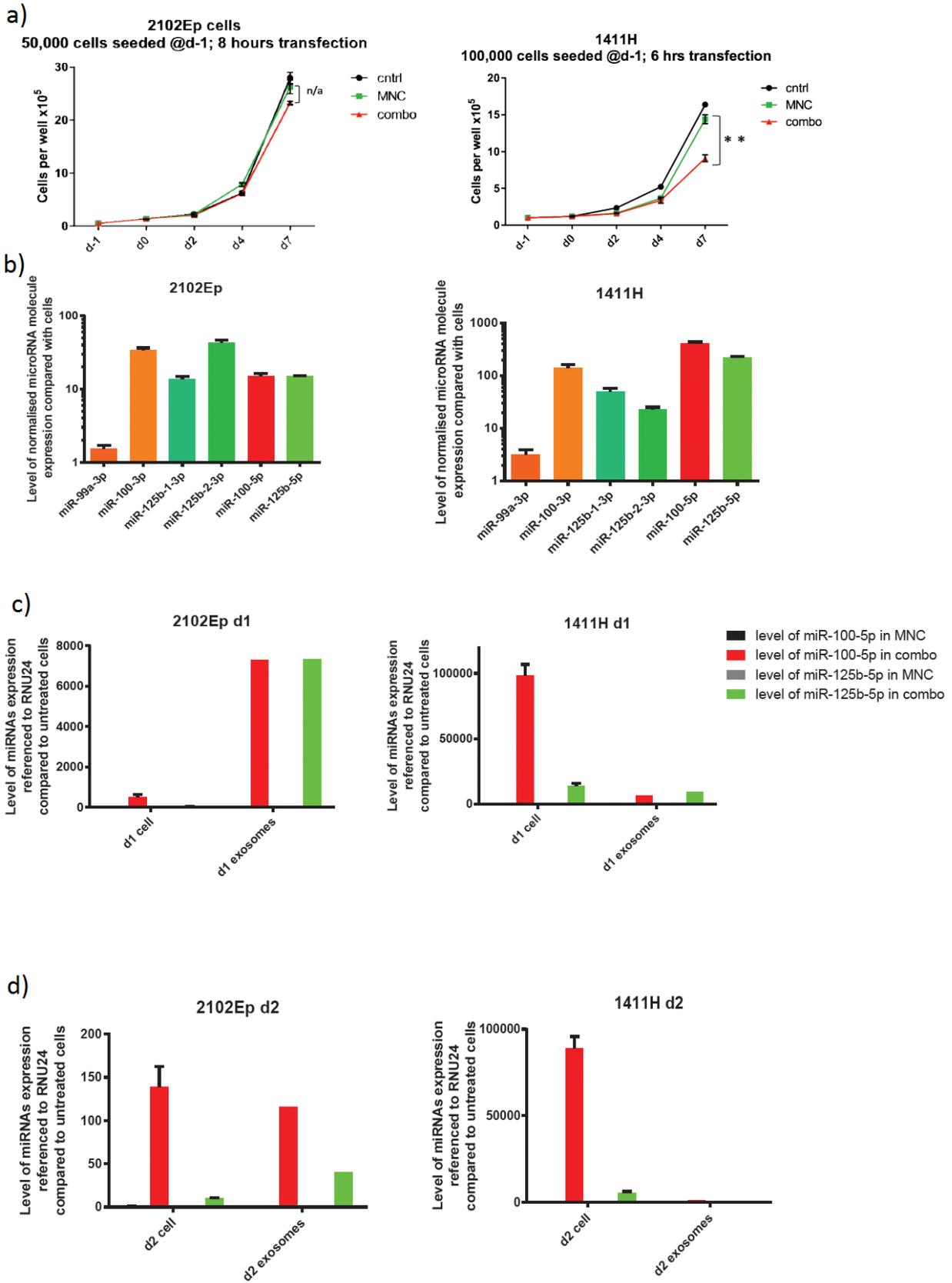
2102Ep d2

1411H d2

Figure 7.1: **Possible involvement of exosome in 2102Ep transfection failure.** The left panel shows data from 2102Ep cell lines, while the right panel shows 1411H data. a) Growth curve of untransfected cells, transfected Combo (2102Ep with 66.6 nM and 1411H with 16.7 nM) and non-targeting control (MNC) (2102Ep with 66.6 nM and 1411H with 16.7 nM). b) Level of miRNAs expression in the exosomes compared with cells by normalising the data to RNU24. c) Levels of miR-100-5p and miR-125b-5p in the cells and in the exosomes at day 1 after Combo transfection and d) at day 2 after Combo transfection.
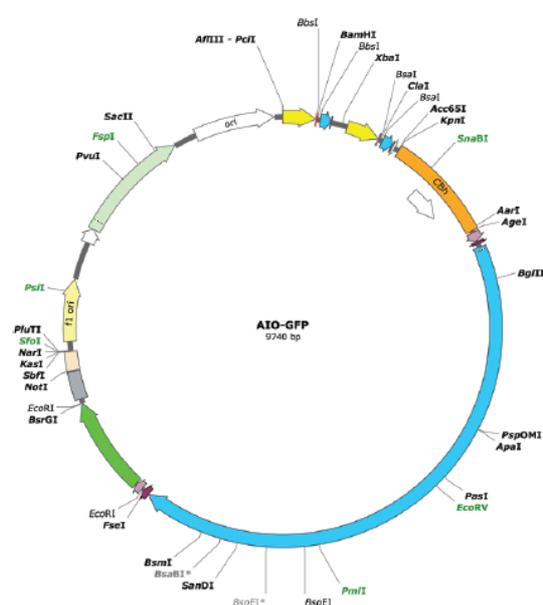
Figure 7.2: **Creation of CRISP-Cas9 plasmid for nSMase2 knock-out.** The Cas9-D10A vector was kindly provided by Professor Stephen Jackson (Chiang et al 2016). Two gRNAs were inserted into the plasmid in two different restriction enzyme points: BsaI site and BbsI site.
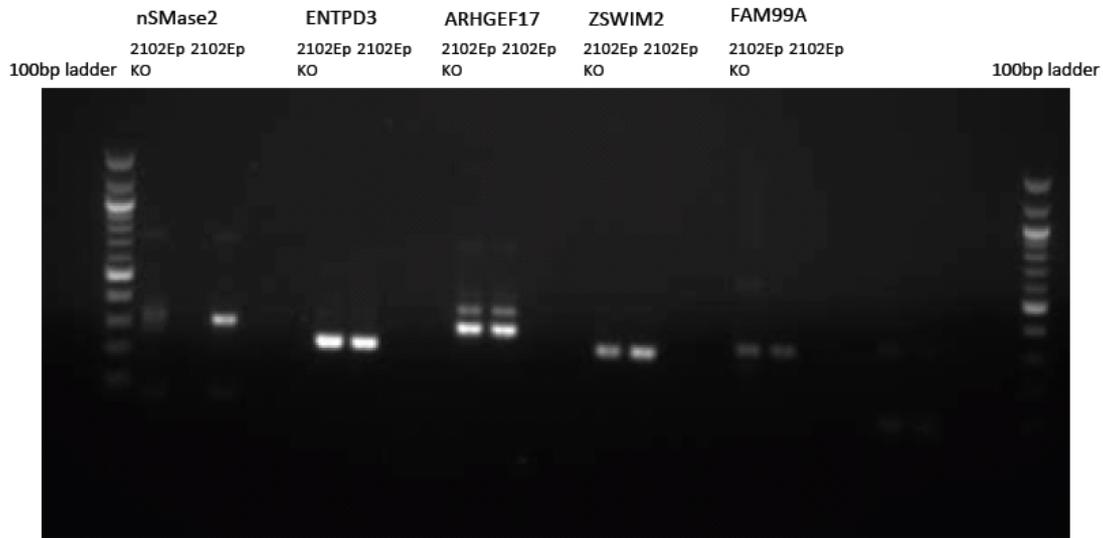
Figure 7.3: **QPCR gel to detect genomic level of nSMase2 after CRISP transfection.** From left to right: *nSMase2* (2102Ep KO and 2102Ep), *ENTPD3* (2102Ep KO and 2102Ep), *ARHGEF17* (2102Ep KO and 2102Ep), *ZSWIM2* (2102Ep KO and 2102Ep) and *FAM99A* (2102Ep KO and 2102Ep). The other 4 genes have been assessed as they had high risk of being targeted by the CRISP-Cas9 plasmid due to similarity of their sequence with the gRNA used for knocking-out nSMase2.
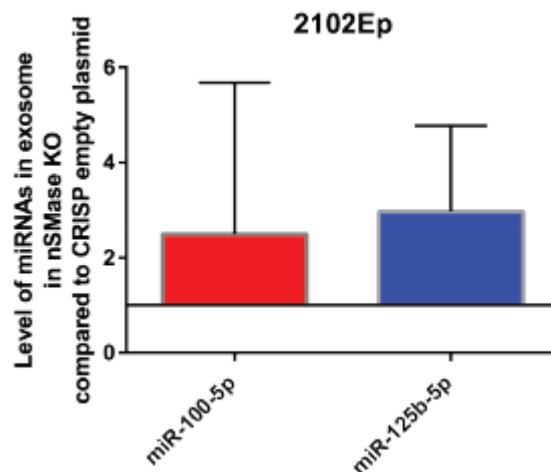
Figure 7.4: **Level of miR-100-5p and miR-125b-5p after combo transfection in 2102Ep cell line.** nSMase2 KO cells and 2102Ep CRISP empty plasmid were transfected using Combo and level of miR-100-5p and miR-125b-5p were tested. nSMase2 KO cells were compared with 2102Ep CRISP empty plasmid and normalised to *RNU24*. No significant differences were noticed. Data presented as mean ± SEM n=3.

Chapter 8　Appendix 2

1411H
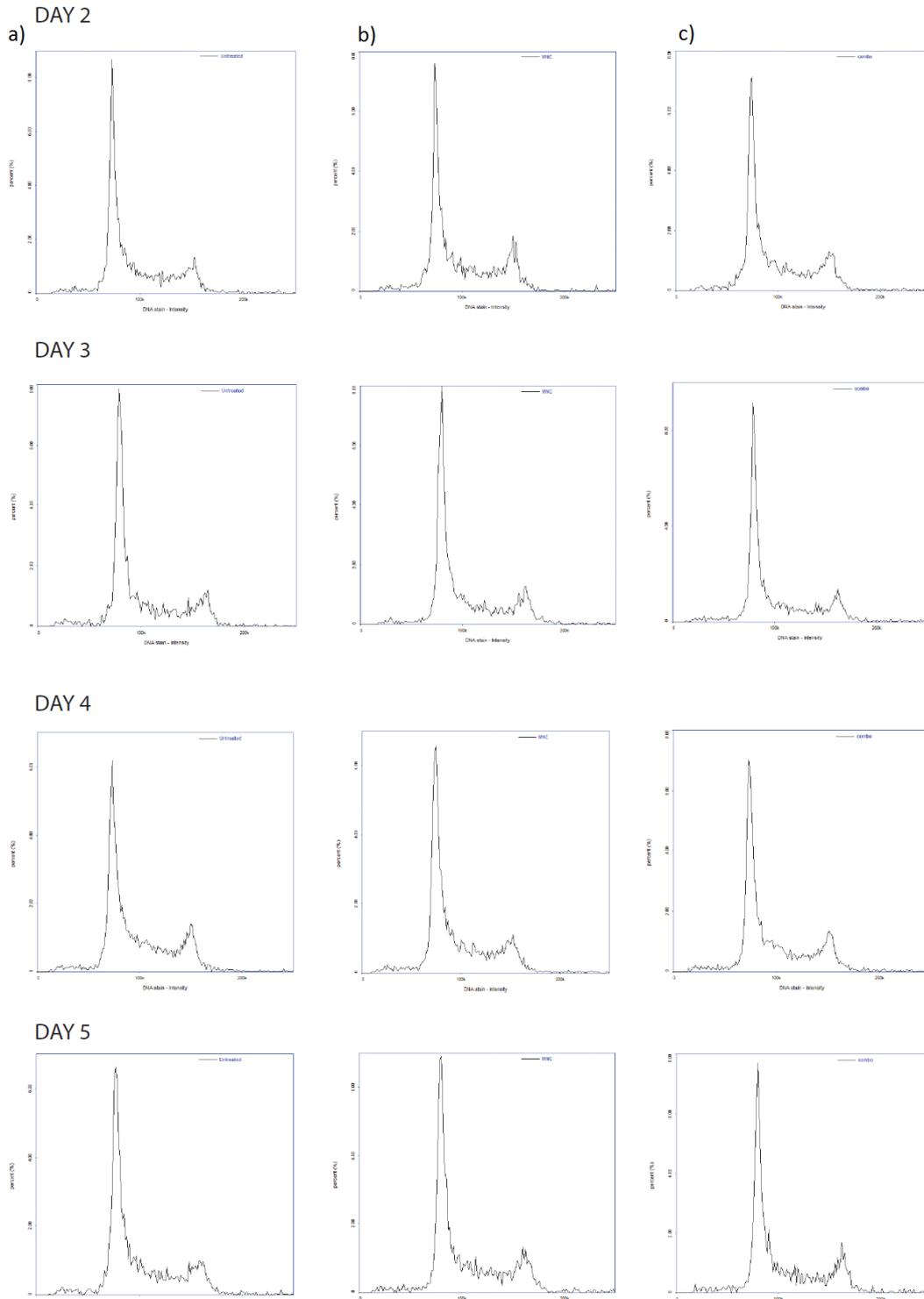
Figure 8.1: **Cell cycle analysis of 1411H cells using NC3000.** 1411H cells were analysed at day 1, 2, 3,4 and 5 after treatment. Row a) untreated cells, row b) non-targeting control (MNC) and row c) Combo treated cell.

References

Afonso, M. B., Rodrigues, P. M., Simão, A. L., Gaspar, M. M., Carvalho, T., Borralho, P., … Rodrigues, C. M. P. (2018). miRNA-21 ablation protects against liver injury and necroptosis in cholestasis. *Cell Death & Differentiation*, *25*(5), 857–872. https://doi.org/10.1038/s41418-017-0019-x

Aigner, A., & Fischer, D. (2016). Nanoparticle-mediated delivery of small RNA molecules in tumor therapy. *Die Pharmazie*, *71*(1), 27–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26867350

Akhavantabasi, S., Sapmaz, A., Tuna, S., & Erson-Bensan, A. E. (2012). miR-125b targets ARID3B in breast cancer cells. *Cell Structure and Function*, *37*(1), 27–38. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22307404

Alagaratnam, S., Lind, G. E., Kraggerud, S. M., Lothe, R. A., & Skotheim, R. I. (2011). The testicular germ cell tumour transcriptome. *International Journal of Andrology*, *34*(4 Pt 2), e133-50; discussion e150-1. https://doi.org/10.1111/j.1365-2605.2011.01169.x

Albany, C., Hever-Jardine, M. P., von Herrmann, K. M., Yim, C. Y., Tam, J., Warzecha, J. M., … Spinella, M. J. (2017). Refractory testicular germ cell tumors are highly sensitive to the second generation DNA methylation inhibitor guadecitabine. *Oncotarget*, *8*(2), 2949–2959. https://doi.org/10.18632/oncotarget.13811

Alexander, M., Ramstead, A. G., Bauer, K. M., Lee, S.-H., Runtsch, M. C., Wallace, J., … O'Connell, R. M. (2017). Rab27-Dependent Exosome Production Inhibits Chronic Inflammation and Enables Acute Responses to Inflammatory Stimuli. *The Journal of Immunology*, *199*(10), 3559–3570. https://doi.org/10.4049/jimmunol.1700904

Amatruda, J. F., Ross, J. A., Christensen, B., Fustino, N. J., Chen, K. S., Hooten, A. J., … Poynter, J. N. (2013a). DNA methylation analysis reveals distinct methylation signatures in pediatric germ cell tumors. *BMC Cancer*, *13*(1), 313. https://doi.org/10.1186/1471-2407-13-313

Amatruda, J. F., Ross, J. A., Christensen, B., Fustino, N. J., Chen, K. S., Hooten, A. J., … Poynter, J. N. (2013b). DNA methylation analysis reveals distinct methylation signatures in pediatric germ cell tumors. *BMC Cancer*, *13*(1), 313. https://doi.org/10.1186/1471-2407-13-313

Ambros, V. (2000). Control of developmental timing in Caenorhabditis elegans. *Current*

Opinion in Genetics & Development, *10*(4), 428–433. https://doi.org/10.1016/S0959-437X(00)00108-8

Anding, A. L., & Baehrecke, E. H. (2015). Autophagy in Cell Life and Cell Death. In *Current topics in developmental biology* (Vol. 114, pp. 67–91). https://doi.org/10.1016/bs.ctdb.2015.07.012

Andrews, P. W., Goodfellow, P. N., Shevinsky, L. H., Bronson, D. L., & Knowles, B. B. (1982). Cell-surface antigens of a clonal human embryonal carcinoma cell line: morphological and antigenic differentiation in culture. *International Journal of Cancer*, *29*(5), 523–531. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7095898

Anglesio, M., Wang, Y., Yang, W., Senz, J., Wan, A., Heravi-Moussavi, A., … Morin, G. (2013). Cancer-associated somatic *DICER1* hotspot mutations cause defective miRNA processing and reverse-strand expression bias to predominantly mature 3p strands through loss of 5p strand cleavage. *The Journal of Pathology*, *229*(3), 400–409. https://doi.org/10.1002/path.4135

Arora, R. S., Alston, R. D., Eden, T. O. B., Geraci, M., & Birch, J. M. (2012). Comparative incidence patterns and trends of gonadal and extragonadal germ cell tumors in England, 1979 to 2003. *Cancer*, *118*(17), 4290–4297. https://doi.org/10.1002/cncr.27403

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. https://doi.org/10.1038/75556

Aslan, D., Garde, C., Nygaard, M. K., Helbo, A. S., Dimopoulos, K., Hansen, J. W., … Kristensen, L. S. (2016). Tumor suppressor microRNAs are downregulated in myelodysplastic syndrome with spliceosome mutations. *Oncotarget*, *7*(9), 9951–9963. https://doi.org/10.18632/oncotarget.7127

Atkin, N. B., & Baker, M. C. (1982, December). Specific chromosome change, i(12p), in testicular tumours? *Lancet*. ENGLAND.

Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P., & Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature*, *455*(7209), 64–71. https://doi.org/10.1038/nature07242

Balzeau, J., Menezes, M. R., Cao, S., & Hagan, J. P. (2017). The LIN28/let-7 Pathway in Cancer. *Frontiers in Genetics*, *8*, 31. https://doi.org/10.3389/fgene.2017.00031

Banzhaf-Strathmann, J., & Edbauer, D. (2014). Good guy or bad guy: the opposing roles of microRNA 125b in cancer. *Cell Communication and Signaling*, *12*(1), 30. https://doi.org/10.1186/1478-811X-12-30

Bartel, D. P. (2004a). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*,

116(2), 281–297.

Bartel, D. P. (2004b). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281–297. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14744438

Baskerville, S., & Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*. https://doi.org/10.1261/rna.7240905

BASKERVILLE, S., & Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, *11*(3), 241–247. https://doi.org/10.1261/rna.7240905

Bastien, S., Smallwood, A., & Kelsey, G. (2011). De novo DNA methylation: a germ cell perspective. *Trends in Genetics*, *28*, 33–42. https://doi.org/10.1016/j.tig.2011.09.004

Batool, A., Wang, Y.-Q., Hao, X.-X., Chen, S.-R., & Liu, Y.-X. (2018). A miR-125b/CSF1-CX3CL1/tumor-associated macrophage recruitment axis controls testicular germ cell tumor growth. *Cell Death & Disease*, *9*(10), 962. https://doi.org/10.1038/s41419-018-1021-z

Bergeron, J. M., Crews, D., & McLachlan, J. A. (1994). PCBs as environmental estrogens: turtle sex determination as a biomarker of environmental contamination. *Environmental Health Perspectives*, *102*(9), 780–781. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9657710

Bergström, R., Adami, H. O., Möhner, M., Zatonski, W., Storm, H., Ekbom, A., … Hakulinen, T. (1996). Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. *Journal of the National Cancer Institute*, *88*(11), 727–733. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8637026

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, *409*(6818), 363–366. https://doi.org/10.1038/35053110

Bhutra, S., Lenkala, D., LaCroix, B., Ye, M., & Huang, R. S. (2014). Identifying and Validating a Combined mRNA and MicroRNA Signature in Response to Imatinib Treatment in a Chronic Myeloid Leukemia Cell Line. *PLoS ONE*, *9*(12), e115003. https://doi.org/10.1371/journal.pone.0115003

Bi, Y., Jing, Y., & Cao, Y. (2015). Overexpression of miR-100 inhibits growth of osteosarcoma through FGFR3. *Tumor Biology*. https://doi.org/10.1007/s13277-015-3581-1

Bian, K., Fan, J., Zhang, X., Yang, X.-W., Zhu, H.-Y., Wang, L., … Zhang, R. (2012). MicroRNA-203 leads to G1 phase cell cycle arrest in laryngeal carcinoma cells by directly targeting survivin. *FEBS Letters*, *586*(6), 804–809.

https://doi.org/10.1016/J.FEBSLET.2012.01.050

Bobbs, A., Gellerman, K., Hallas, W. M., Joseph, S., Yang, C., Kurkewich, J., & Cowden Dahl, K. D. (2015). ARID3B Directly Regulates Ovarian Cancer Promoting Genes. *PLOS ONE*, *10*(6), e0131961. https://doi.org/10.1371/journal.pone.0131961

Bohnsack, M. T., Czaplinski, K., & Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, N.Y.)*, *10*(2), 185–191. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14730017

Brueckner, B., Stresemann, C., Kuner, R., Mund, C., Musch, T., Meister, M., … Lyko, F. (2007). The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. *Cancer Research*, *67*(4), 1419–1423. https://doi.org/10.1158/0008-5472.CAN-06-4074

Bueno, M. J., & Malumbres, M. (2011). MicroRNAs and the cell cycle. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1812*(5), 592–601. https://doi.org/10.1016/J.BBADIS.2011.02.002

Bussey, K. J., Lawce, H. J., Himoe, E., Shu, X. O., Suijkerbuijk, R. F., Olson, S. B., & Magenis, R. E. (2001). Chromosomes 1 and 12 abnormalities in pediatric germ cell tumors by interphase fluorescence in situ hybridization. *Cancer Genetics and Cytogenetics*, *125*(2), 112–118. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11369053

Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, N.Y.)*, *10*(12), 1957–1966. https://doi.org/10.1261/rna.7135204

Caldwell, B. T., Wilcox, D. T., & Cost, N. G. (2017). Current Management for Pediatric Urologic Oncology. *Advances in Pediatrics*, *64*(1), 191–223. https://doi.org/10.1016/j.yapd.2017.04.001

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., … Croce, C. M. (2002). Nonlinear partial differential equations and applications: Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, *99*(24), 15524–15529. https://doi.org/10.1073/pnas.242606799

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., … Croce, C. M. (2004a). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences*, *101*(9), 2999–3004. https://doi.org/10.1073/pnas.0307323101

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., … Croce, C. M. (2004b). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences*, *101*(9),

2999–3004. https://doi.org/10.1073/pnas.0307323101

Campaner, S., & Amati, B. (2012). Two sides of the Myc-induced DNA damage response: from tumor suppression to tumor maintenance. *Cell Division*, *7*(1), 6. https://doi.org/10.1186/1747-1028-7-6

Cao, Y., Song, J., Ge, J., Song, Z., Chen, J., & Wu, C. (2018). MicroRNA-100 suppresses human gastric cancer cell proliferation by targeting CXCR7. *Oncology Letters*, *15*(1), 453–458. https://doi.org/10.3892/ol.2017.7305

Carvajal, L. A., Hamard, P.-J., Tonnessen, C., & Manfredi, J. J. (2012). E2F7, a novel target, is up-regulated by p53 and mediates DNA damage-dependent transcriptional repression. *Genes & Development*, *26*(14), 1533–1545. https://doi.org/10.1101/gad.184911.111

Casey, M.-C., Kerin, M. J., Brown, J. A., & Sweeney, K. J. (2015). Evolution of a research field-a micro (RNA) example. *PeerJ*, *3*, e829. https://doi.org/10.7717/peerj.829

Cecchetto, G. (2014). Gonadal germ cell tumors in children and adolescents. *Journal of Indian Association of Pediatric Surgeons*, *19*(4), 189–194. https://doi.org/10.4103/0971-9261.141995

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: Patterns and paradigms. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2540

Chaganti, R. S., & Houldsworth, J. (2000). Genetics and biology of adult human male germ cell tumors. *Cancer Research*, *60*(6), 1475–1482. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10749107

Chang, H.-M., Martinez, N. J., Thornton, J. E., Hagan, J. P., Nguyen, K. D., & Gregory, R. I. (2012). Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. *Nature Communications*, *3*, 923. https://doi.org/10.1038/ncomms1909

Chang, T.-C., Yu, D., Lee, Y.-S., Wentzel, E. A., Arking, D. E., West, K. M., … Mendell, J. T. (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. *Nature Genetics*, *40*(1), 43–50. https://doi.org/10.1038/ng.2007.30

Chen, D., Sun, Y., Yuan, Y., Han, Z., Zhang, P., Zhang, J., … Ma, L. (2014). miR-100 Induces Epithelial-Mesenchymal Transition but Suppresses Tumorigenesis, Migration and Invasion. *PLoS Genetics*, *10*(2), e1004177. https://doi.org/10.1371/journal.pgen.1004177

Chen, H., & Xu, Z. (2015). Hypermethylation-Associated Silencing of miR-125a and miR-125b: A Potential Marker in Colorectal Cancer. *Disease Markers*, *2015*, 1–7. https://doi.org/10.1155/2015/345080

Chen, Y.-F., Liu, H., Luo, X.-J., Zhao, Z., Zou, Z.-Y., Li, J., … Liang, Y. (2017). The roles of

reactive oxygen species (ROS) and autophagy in the survival and death of leukemia cells. *Critical Reviews in Oncology/Hematology*, *112*, 21–30. https://doi.org/10.1016/j.critrevonc.2017.02.004

Cheung, H.-H., Lee, T.-L., Rennert, O. M., & Chan, W.-Y. (2009). DNA methylation of cancer genome. *Birth Defects Research Part C: Embryo Today: Reviews*, *87*(4), 335–350. https://doi.org/10.1002/bdrc.20163

Chi, S. W., Hannon, G. J., & Darnell, R. B. (2012). An alternative mode of microRNA target recognition. *Nature Structural & Molecular Biology*, *19*(3), 321–327. https://doi.org/10.1038/nsmb.2230

Chorn, G., Klein-McDowell, M., Zhao, L., Saunders, M. A., Flanagan, W. M., Willingham, A. T., & Lim, L. P. (2012). Single-stranded microRNA mimics. *RNA*, *18*(10), 1796–1804. https://doi.org/10.1261/rna.031278.111

Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., … Croce, C. M. (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences*, *102*(39), 13944–13949. https://doi.org/10.1073/pnas.0506654102

Cloonan, N. (2015). Re-thinking miRNA-mRNA interactions: intertwining issues confound target discovery. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *37*(4), 379–388. https://doi.org/10.1002/bies.201400191

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., … Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes & Development*, *26*(15), 1714–1728. https://doi.org/10.1101/gad.194209.112

Collinson, K., Murray, M. J., Orsi, N. M., Cummings, M., Shipley, J., Joffe, J. K., … Stark, D. (2014). Age-related biological features of germ cell tumors. *Genes, Chromosomes and Cancer*, *53*(3), 215–227. https://doi.org/10.1002/gcc.22131

Cui, L., Zhou, H., Zhao, H., Zhou, Y., Xu, R., Xu, X., … He, X. (2012). MicroRNA-99a induces G1-phase cell cycle arrest and suppresses tumorigenicity in renal cell carcinoma. *BMC Cancer*, *12*(1), 546. https://doi.org/10.1186/1471-2407-12-546

Curtis, H. J., Seow, Y., Wood, M. J. A., & Varela, M. A. (2017). Knockdown and replacement therapy mediated by artificial mirtrons in spinocerebellar ataxia 7. *Nucleic Acids Research*, *45*(13), 7870–7885. https://doi.org/10.1093/nar/gkx483

Dai, E., Yu, X., Zhang, Y., Meng, F., Wang, S., Liu, X., … Jiang, W. (2014). EpimiR: a database of curated mutual regulation between miRNAs and epigenetic modifications. *Database*, *2014*. https://doi.org/10.1093/database/bau023

Danial, N. N., & Korsmeyer, S. J. (2004). Cell death: critical control points. *Cell*, *116*(2), 205–219. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14744432

de Jong, J., Stoop, H., Gillis, A., van Gurp, R., van de Geijn, G.-J., Boer, M. de, … Looijenga, L. (2008). Differential expression of SOX17 and SOX2 in germ cells and stem cells has biological and clinical implications. *The Journal of Pathology*, *215*(1), 21–30. https://doi.org/10.1002/path.2332

Dehner, L. P. (1983). Gonadal and extragonadal germ cell neoplasia of childhood. *Human Pathology*, *14*(6), 493–511. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6343221

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S., & Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M109.089433

Dobi, K. C., & Winston, F. (2007). Analysis of Transcriptional Activation at a Distance in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *27*(15), 5575–5586. https://doi.org/10.1128/MCB.00459-07

Drosten, M., Sum, E. Y. M., Lechuga, C. G., Simón-Carrasco, L., Jacob, H. K. C., García-Medina, R., … Barbacid, M. (2014). Loss of p53 induces cell proliferation via Ras-independent activation of the Raf/Mek/Erk signaling pathway. *Proceedings of the National Academy of Sciences*, *111*(42), 15155–15160. https://doi.org/10.1073/pnas.1417549111

Du, P., Kibbe, W. A., & Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics (Oxford, England)*, *24*(13), 1547–1548. https://doi.org/10.1093/bioinformatics/btn224

Eichhorn, S. W., Guo, H., McGeary, S. E., Rodriguez-Mias, R. A., Shin, C., Baek, D., … Bartel, D. P. (2014). mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAs by the Time Substantial Repression Ensues. *Molecular Cell*, *56*(1), 104–115. https://doi.org/10.1016/J.MOLCEL.2014.08.028

Einhorn, L. H., & Williams, S. D. (1979). Chemotherapy of disseminated testicular cancer. *The Western Journal of Medicine*, *131*(1), 1–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/90432

Ellwanger, D. C., Büttner, F. A., Mewes, H.-W., & Stümpflen, V. (2011). The sufficient minimal set of miRNA seed types. *Bioinformatics*, *27*(10), 1346–1350. https://doi.org/10.1093/bioinformatics/btr149

Emmrich, S., Rasche, M., Schöning, J., Reimer, C., Keihani, S., Maroz, A., … Klusmann, J.-H. (2014). miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell

homeostasis by shifting the balance between TGFβ and Wnt signaling. *Genes & Development*, *28*(8), 858–874. https://doi.org/10.1101/gad.233791.113

Emmrich, S., Streltsov, A., Schmidt, F., Thangapandi, V., Reinhardt, D., & Klusmann, J.-H. (2014). LincRNAs MONC and MIR100HG act as oncogenes in acute megakaryoblastic leukemia. *Molecular Cancer*, *13*(1), 171. https://doi.org/10.1186/1476-4598-13-171

Endo-Munoz, L., Dahler, A., Teakle, N., Rickwood, D., Hazar-Rethinam, M., Abdul-Jabbar, I., … Saunders, N. (2009). E2F7 can regulate proliferation, differentiation, and apoptotic responses in human keratinocytes: implications for cutaneous squamous cell carcinoma formation. *Cancer Research*, *69*(5), 1800–1808. https://doi.org/10.1158/0008-5472.CAN-08-2725

Esquela-Kerscher, A., & Slack, F. J. (2006). Oncomirs — microRNAs with a role in cancer. *Nature Reviews Cancer*, *6*(4), 259–269. https://doi.org/10.1038/nrc1840

Exelby, P. R. (1980). Testicular Cancer in Children. *Cancer*, *45*, 1803–1809. https://doi.org/10.1002/cncr.1980.45.s7.1803

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., … D'Eustachio, P. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkx1132

Farazi, T. A., Spitzer, J. I., Morozov, P., & Tuschl, T. (2011). miRNAs in human cancer. *The Journal of Pathology*, *223*(2), 102–115. https://doi.org/10.1002/path.2806

Fazi, F., Racanicchi, S., Zardo, G., Starnes, L. M., Mancini, M., Travaglini, L., … Nervi, C. (2007). Epigenetic Silencing of the Myelopoiesis Regulator microRNA-223 by the AML1/ETO Oncoprotein. *Cancer Cell*, *12*(5), 457–466. https://doi.org/10.1016/j.ccr.2007.09.020

Fazi, F., Rosa, A., Fatica, A., Gelmetti, V., De Marchis, M. L., Nervi, C., & Bozzoni, I. (2005). A Minicircuitry Comprised of MicroRNA-223 and Transcription Factors NFI-A and C/EBPα Regulates Human Granulopoiesis. *Cell*, *123*(5), 819–831. https://doi.org/10.1016/j.cell.2005.09.023

Fontana, L., Pelosi, E., Greco, P., Racanicchi, S., Testa, U., Liuzzi, F., … Peschle, C. (2007). MicroRNAs 17-5p–20a–106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. *Nature Cell Biology*, *9*(7), 775–787. https://doi.org/10.1038/ncb1613

Fujimoto, T., Miyayama, Y., & Fuyuta, M. (1977). The origin, migration and fine morphology of human primordial germ cells. *The Anatomical Record*, *188*(3), 315–329. https://doi.org/10.1002/ar.1091880305

Fujino, Y., Takeishi, S., Nishida, K., Okamoto, K., Muguruma, N., Kimura, T., … Takayama,

T. (2017). Downregulation of microRNA-100/microRNA-125b is associated with lymph node metastasis in early colorectal cancer with submucosal invasion. *Cancer Science*, *108*(3), 390–397. https://doi.org/10.1111/cas.13152

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., & Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M210256200

Galluzzi, L., & Kroemer, G. (2008). Necroptosis: A Specialized Pathway of Programmed Necrosis. *Cell*, *135*(7), 1161–1163. https://doi.org/10.1016/J.CELL.2008.12.004

Genovesi, L. A., Anderson, D., Carter, K. W., Giles, K. M., & Dallas, P. B. (2012). Identification of suitable endogenous control genes for microRNA expression profiling of childhood medulloblastoma and human neural stem cells. *BMC Research Notes*, *5*, 507. https://doi.org/10.1186/1756-0500-5-507

Ghazarian, A. A., Trabert, B., Devesa, S. S., & McGlynn, K. A. (2015a). Recent trends in the incidence of testicular germ cell tumors in the United States. *Andrology*, *3*(1), 13–18. https://doi.org/10.1111/andr.288

Ghazarian, A. A., Trabert, B., Devesa, S. S., & McGlynn, K. A. (2015b). Recent trends in the incidence of testicular germ cell tumors in the United States. *Andrology*, *3*(1), 13–18. https://doi.org/10.1111/andr.288

Gilam, A., Conde, J., Weissglas-Volkov, D., Oliva, N., Friedman, E., Artzi, N., & Shomron, N. (2016). Local microRNA delivery targets Palladin and prevents metastatic breast cancer. *Nature Communications*, *7*, 12868. https://doi.org/10.1038/ncomms12868

Gozuacik, D., Akkoc, Y., Ozturk, D. G., & Kocak, M. (2017). Autophagy-Regulating microRNAs and Cancer. *Frontiers in Oncology*, *7*, 65. https://doi.org/10.3389/fonc.2017.00065

Griffiths-Jones, S., Saini, H. K., Van Dongen, S., & Enright, A. J. (2008). miRBase: Tools for microRNA genomics. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkm952

Gronbaek, Kirsten Hother, Christoffer Jones, P. A. (2007). Epigenetic changes in cancer. *APMIS*, *115*(10), 1039–1059. https://doi.org/10.1111/j.1600-0463.2007.apm_636.xml.x

Gu, L., Li, H., Chen, L., Ma, X., Gao, Y., Li, X., … Zhang, X. (2015). MicroRNAs as prognostic molecular signatures in renal cell carcinoma: a systematic review and meta-analysis. *Oncotarget*, *6*(32), 32545–32560. https://doi.org/10.18632/oncotarget.5324

Guillette, L. J., Gross, T. S., Masson, G. R., Matter, J. M., Percival, H. F., Woodward, A. R., & Woodward, A. R. (1994). Developmental abnormalities of the gonad and abnormal sex hormone concentrations in juvenile alligators from contaminated and control lakes in Florida. *Environmental Health Perspectives*, *102*(8), 680–688. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/7895709

Gulyaeva, L. F., & Kushlinskiy, N. E. (2016). Regulatory mechanisms of microRNA expression. *Journal of Translational Medicine*, *14*(1), 143. https://doi.org/10.1186/s12967-016-0893-x

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., … Qiao, J. (2015). The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell*, *161*(6), 1437–1452. https://doi.org/10.1016/j.cell.2015.05.015

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010a). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, *466*(7308), 835–840. https://doi.org/10.1038/nature09267

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010b). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, *466*(7308), 835–840. https://doi.org/10.1038/nature09267

Guo, Q., Zhang, J., Li, J., Zou, L., Zhang, J., Xie, Z., … Wu, Y. (2013). Forced miR-146a expression causes autoimmune lymphoproliferative syndrome in mice via downregulation of Fas in germinal center B cells. *Blood*, *121*(24), 4875–4883. https://doi.org/10.1182/blood-2012-08-452425

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology*, *15*(8), 509–524. https://doi.org/10.1038/nrm3838

Han, H., Sun, D., Li, W., Shen, H., Zhu, Y., Li, C., … Li, Y. (2013). A c-Myc-MicroRNA functional feedback loop affects hepatocarcinogenesis. *Hepatology*, *57*(6), 2378–2389. https://doi.org/10.1002/hep.26302

Hardell, L., van Bavel, B., Lindström, G., Carlberg, M., Dreifaldt, A. C., Wijkström, H., … Kolmert, T. (2003). Increased concentrations of polychlorinated biphenyls, hexachlorobenzene, and chlordanes in mothers of men with testicular cancer. *Environmental Health Perspectives*, *111*(7), 930–934. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12782494

Harrington, C. T., Lin, E. I., Olson, M. T., & Eshleman, J. R. (2013). Fundamentals of Pyrosequencing. *Archives of Pathology & Laboratory Medicine*, *137*(9), 1296–1303. https://doi.org/10.5858/arpa.2012-0463-RA

Hata, A., & Kashima, R. (2016). Dysregulation of microRNA biogenesis machinery in cancer. *Critical Reviews in Biochemistry and Molecular Biology*, *51*(3), 121–134. https://doi.org/10.3109/10409238.2015.1117054

Haugnes, H. S., Bosl, G. J., Boer, H., Gietema, J. A., Brydøy, M., Oldenburg, J., … Fosså, S. D. (2012). Long-term and late effects of germ cell testicular cancer treatment and

implications for follow-up. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *30*(30), 3752–3763. https://doi.org/10.1200/JCO.2012.43.4431

Haugnes, H. S., Wethal, T., Aass, N., Dahl, O., Klepp, O., Langberg, C. W., … Fosså, S. D. (2010). Cardiovascular risk factors and morbidity in long-term survivors of testicular cancer: a 20-year follow-up study. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *28*(30), 4649–4657. https://doi.org/10.1200/JCO.2010.29.9362

Hausser, J., Syed, A. P., Bilen, B., & Zavolan, M. (2013). Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Research*, *23*(4), 604–615. https://doi.org/10.1101/gr.139758.112

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., … Takahashi, T. (2005). A Polycistronic MicroRNA Cluster, *miR-17-92* , Is Overexpressed in Human Lung Cancers and Enhances Cell Proliferation. *Cancer Research*, *65*(21), 9628–9632. https://doi.org/10.1158/0008-5472.CAN-05-2352

He, X.-X., Chang, Y., Meng, F.-Y., Wang, M.-Y., Xie, Q.-H., Tang, F., … Lin, J.-S. (2012). MicroRNA-375 targets AEG-1 in hepatocellular carcinoma and suppresses liver cancer cell growth in vitro and in vivo. *Oncogene*, *31*(28), 3357–3369. https://doi.org/10.1038/onc.2011.500

Heifetz, S. A., Cushing, B., Giller, R., Shuster, J. J., Stolar, C. J., Vinocur, C. D., & Hawkins, E. P. (1998). Immature teratomas in children: pathologic considerations: a report from the combined Pediatric Oncology Group/Children's Cancer Group. *The American Journal of Surgical Pathology*, *22*(9), 1115–1124. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9737245

Henson, B. J., Bhattacharjee, S., O'Dee, D. M., Feingold, E., & Gollin, S. M. (2009). Decreased expression of miR-125b and miR-100 in oral cancer cells contributes to malignancy. *Genes, Chromosomes and Cancer*, *48*(7), 569–582. https://doi.org/10.1002/gcc.20666

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., … Kim, V. N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell*, *138*(4), 696–708. https://doi.org/10.1016/j.cell.2009.08.002

Hermeking, H. (2010). The miR-34 family in cancer and apoptosis. *Cell Death & Differentiation*, *17*(2), 193–199. https://doi.org/10.1038/cdd.2009.56

Herrington, C. S. (1999). Tumors of the Ovary, Maldeveloped Gonads, Fallopian Tube and Broad Ligament. Atlas of Tumor Pathology. Third Series, Fascicle 23 Robert E. Scully, Robert H. Young and Philip B. Clement. Armed Forces Institute of Pathology, Washington, DC, 1998. No. of pages: 527. Price: $95.00. ISBN: 1 881041 43 3. *The*

Journal of Pathology, *189*(1), 145–145. https://doi.org/10.1002/(SICI)1096-9896(199909)189:1<145::AID-PATH420>3.0.CO;2-F

Höbel, S., & Aigner, A. (2013). Polyethylenimines for siRNA and miRNA delivery *in vivo*. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, *5*(5), 484–501. https://doi.org/10.1002/wnan.1228

Hochstetter, A. R., & Hedinger, C. E. (1982). The differential diagnosis of testicular germ cell tumors in theory and practice. *Virchows Archiv A Pathological Anatomy and Histology*, *396*(3), 247–277. https://doi.org/10.1007/BF00431386

Hoffman, B., & Liebermann, D. A. (2008). Apoptotic signaling by c-MYC. *Oncogene*, *27*(50), 6462–6472. https://doi.org/10.1038/onc.2008.312

Holler, N., Zaru, R., Micheau, O., Thome, M., Attinger, A., Valitutti, S., … Tschopp, J. (2000). Fas triggers an alternative, caspase-8–independent cell death pathway using the kinase RIP as effector molecule. *Nature Immunology*, *1*(6), 489–495. https://doi.org/10.1038/82732

Honecker, F., Aparicio, J., Berney, D., Beyer, J., Bokemeyer, C., Cathomas, R., … Horwich, A. (2018). ESMO Consensus Conference on testicular germ cell cancer: diagnosis, treatment and follow-up. *Annals of Oncology*, *29*(8), 1658–1686. https://doi.org/10.1093/annonc/mdy217

Honecker, F., Stoop, H., Mayer, F., Bokemeyer, C., Castrillon, D. H., Lau, Y.-F. C., … Oosterhuis, J. W. (2006). Germ cell lineage differentiation in non-seminomatous germ cell tumours. *The Journal of Pathology*, *208*(3), 395–400. https://doi.org/10.1002/path.1872

Honorio, S., Agathanggelou, A., Wernert, N., Rothe, M., Maher, E. R., & Latif, F. (2003). Frequent epigenetic inactivation of the RASSF1A tumour suppressor gene in testicular tumours and distinct methylation profiles of seminoma and nonseminoma testicular germ cell tumours. *Oncogene*, *22*(3), 461–466. https://doi.org/10.1038/sj.onc.1206119

Horwich, A., Nicol, D., & Huddart, R. (2013). Testicular germ cell tumours. *BMJ (Clinical Research Ed.)*, *347*, f5526. https://doi.org/10.1136/BMJ.F5526

Huang, G., Nishimoto, K., Zhou, Z., Hughes, D., & Kleinerman, E. S. (2012). miR-20a Encoded by the miR-17-92 Cluster Increases the Metastatic Potential of Osteosarcoma Cells by Regulating Fas Expression. *Cancer Research*, *72*(4), 908–916. https://doi.org/10.1158/0008-5472.CAN-11-1460

Huebner, R. J., & Todaro, G. J. (1969). Oncogenes of RNA tumor viruses as determinants of cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *64*(3), 1087–1094. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5264139

Ibrahim, A. F., Weirauch, U., Thomas, M., Grünweller, A., Hartmann, R. K., & Aigner, A. (2011). MicroRNA replacement therapy for miR-145 and miR-33a is efficacious in a model of colon carcinoma. *Cancer Research*, *71*(15), 5214–5224. https://doi.org/10.1158/0008-5472.CAN-10-4645

Ioka, A., Tsukuma, H., Ajiki, W., & Oshima, A. (2003). Ovarian cancer incidence and survival by histologic type in Osaka, Japan. *Cancer Science*, *94*(3), 292–296. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12824924

Iorio, M. V., Ferracin, M., Liu, C.-G., Veronese, A., Spizzo, R., Sabbioni, S., … Croce, C. M. (2005). MicroRNA Gene Expression Deregulation in Human Breast Cancer. *Cancer Research*, *65*(16), 7065–7070. https://doi.org/10.1158/0008-5472.CAN-05-1783

Iwakawa, H.-O., & Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends in Cell Biology*, *25*, 651–665. https://doi.org/10.1016/j.tcb.2015.07.011

Jeyapalan, J. N., Noor, D. A. M., Lee, S.-H., Tan, C. L., Appleby, V. A., Kilday, J. P., … Scotting, P. J. (2011). Methylator phenotype of malignant germ cell tumours in children identifies strong candidates for chemotherapy resistance. *British Journal of Cancer*, *105*(4), 575–585. https://doi.org/10.1038/bjc.2011.218

Jia, H.-Y., Wang, Y.-X., Yan, W.-T., Li, H.-Y., Tian, Y.-Z., Wang, S.-M., & Zhao, H.-L. (2012). MicroRNA-125b Functions as a Tumor Suppressor in Hepatocellular Carcinoma Cells. *International Journal of Molecular Sciences*, *13*(7), 8762–8774. https://doi.org/10.3390/ijms13078762

Jiang, Q., He, M., Guan, S., Ma, M., Wu, H., Yu, Z., … Wei, M. (2016). MicroRNA-100 suppresses the migration and invasion of breast cancer cells by targeting FZD-8 and inhibiting Wnt/β-catenin signaling pathway. *Tumor Biology*, *37*(4), 5001–5011. https://doi.org/10.1007/s13277-015-4342-x

Jin, H. Y., Gonzalez-Martin, A., Miletic, A. V, Lai, M., Knight, S., Sabouri-Ghomi, M., … Xiao, C. (2015). Transfection of microRNA Mimics Should Be Used with Caution. *Frontiers in Genetics*, *6*, 340. https://doi.org/10.3389/fgene.2015.00340

Jin, L., Zhang, Z., Li, Y., He, T., Hu, J., Liu, J., … Lai, Y. (2017). miR-125b is associated with renal cell carcinoma cell migration, invasion and apoptosis. *Oncology Letters*, *13*(6), 4512–4520. https://doi.org/10.3892/ol.2017.5985

Johnson, C. D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., … Slack, F. J. (2007). The let-7 MicroRNA Represses Cell Proliferation Pathways in Human Cells. *Cancer Research*, *67*(16), 7713–7722. https://doi.org/10.1158/0008-5472.CAN-07-1083

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., … Slack, F. J.

(2005). RAS Is Regulated by the let-7 MicroRNA Family. *Cell*, *120*(5), 635–647. https://doi.org/10.1016/j.cell.2005.01.014

Jonas, S., & Izaurralde, E. (2013). The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes & Development*, *27*(24), 2628–2641. https://doi.org/10.1101/gad.227843.113

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, *13*(7), 484–492. https://doi.org/10.1038/nrg3230

Kaatsch, P., Häfner, C., Calaminus, G., Blettner, M., & Tulla, M. (2015). Pediatric germ cell tumors from 1987 to 2011: incidence rates, time trends, and survival. *Pediatrics*, *135*(1), e136-43. https://doi.org/10.1542/peds.2014-1989

Kandori, S., Kawai, K., Fukuhara, Y., Joraku, A., Miyanaga, N., Shimazui, T., & Akaza, H. (2010). A case of metastatic testicular cancer complicated by pulmonary hemorrhage due to choriocarcinoma syndrome. *International Journal of Clinical Oncology*, *15*(6), 611–614. https://doi.org/10.1007/s10147-010-0092-3

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10592173

Kao, C.-S., Idrees, M. T., Young, R. H., & Ulbright, T. M. (2012). Solid Pattern Yolk Sac Tumor. *The American Journal of Surgical Pathology*, *36*(3), 360–367. https://doi.org/10.1097/PAS.0b013e31823c510b

Katsube, Y., Berg, J. W., & Silverberg, S. G. (1982). Epidemiologic pathology of ovarian tumors: a histopathologic review of primary ovarian neoplasms diagnosed in the Denver Standard Metropolitan Statistical Area, 1 July-31 December 1969 and 1 July-31 December 1979. *International Journal of Gynecological Pathology : Official Journal of the International Society of Gynecological Pathologists*, *1*(1), 3–16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7184889

Kim, K. H., & Lee, M.-S. (2014). Autophagy—a key player in cellular and body metabolism. *Nature Reviews Endocrinology*, *10*(6), 322–337. https://doi.org/10.1038/nrendo.2014.35

Kim, V. N., Han, J., & Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology*, *10*(2), 126–139. https://doi.org/10.1038/nrm2632

Kimmins, S., & Sassone-Corsi, P. (2005). Chromatin remodelling and epigenetic features of germ cells. *Nature*, *434*(7033), 583–589. https://doi.org/10.1038/nature03368

Klein, U., Lia, M., Crespo, M., Siegel, R., Shen, Q., Mo, T., … Dalla-Favera, R. (2010). The DLEU2/miR-15a/16-1 Cluster Controls B Cell Proliferation and Its Deletion Leads to Chronic Lymphocytic Leukemia. *Cancer Cell*, *17*(1), 28–40.

https://doi.org/10.1016/j.ccr.2009.11.019

Kloosterman, W. P., & Hochstenbach, R. (2014). Deciphering the pathogenic consequences of chromosomal aberrations in human genetic disease. *Molecular Cytogenetics*, *7*(1), 100. https://doi.org/10.1186/s13039-014-0100-9

Kollmannsberger, C., Kuzcyk, M., Mayer, F., Hartmann, J. T., Kanz, L., & Bokemeyer, C. (1999). Late toxicity following curative treatment of testicular cancer. *Seminars in Surgical Oncology*, *17*(4), 275–281.

Korkola, J. E., Houldsworth, J., Chadalavada, R. S. V, Olshen, A. B., Dobrzynski, D., Reuter, V. E., … Chaganti, R. S. K. (2006). Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. *Cancer Research*, *66*(2), 820–827. https://doi.org/10.1158/0008-5472.CAN-05-2445

Koster, R., Mitra, N., D'Andrea, K., Vardhanabhuti, S., Chung, C. C., Wang, Z., … Kanetsky, P. A. (2014). Pathway-based analysis of GWAs data identifies association of sex determination genes with susceptibility to testicular germ cell tumors. *Human Molecular Genetics*, *23*(22), 6061–6068. https://doi.org/10.1093/hmg/ddu305

Kota, J., Chivukula, R. R., O'Donnell, K. A., Wentzel, E. A., Montgomery, C. L., Hwang, H.-W., … Mendell, J. T. (2009). Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. *Cell*, *137*(6), 1005–1017. https://doi.org/10.1016/j.cell.2009.04.021

Koul, S., Houldsworth, J., Mansukhani, M. M., Donadio, A., McKiernan, J. M., Reuter, V. E., … Murty, V. V. (2002). Characteristic promoter hypermethylation signatures in male germ cell tumors. *Molecular Cancer*, *1*, 8. https://doi.org/10.1186/1476-4598-1-8

Kraggerud, S. M., Skotheim, R. I., Szymanska, J., Eknaes, M., Fosså, S. D., Stenwig, A. E., … Lothe, R. A. (2002). Genome profiles of familial/bilateral and sporadic testicular germ cell tumors. *Genes, Chromosomes and Cancer*, *34*(2), 168–174. https://doi.org/10.1002/gcc.10058

Krakauer, D. C., & Plotkin, J. B. (2002). Redundancy, antiredundancy, and the robustness of genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(3), 1405–1409. https://doi.org/10.1073/pnas.032668599

Kristensen, D. G., Skakkebaek, N. E., Rajpert-De Meyts, E., & Almstrup, K. (2013). Epigenetic features of testicular germ cell tumours in relation to epigenetic characteristics of foetal germ cells. *The International Journal of Developmental Biology*, *57*(2–4), 309–317. https://doi.org/10.1387/ijdb.130142ka

Kristensen, D. G., Skakkebk, N. E., Rajpert-De Meyts, E., & Almstrup, K. (2013). Epigenetic features of testicular germ cell tumours in relation to epigenetic characteristics of foetal

germ cells. *The International Journal of Developmental Biology*, *57*(2-3–4), 309–317. https://doi.org/10.1387/ijdb.130142ka

Kroemer, G., Mariño, G., & Levine, B. (2010). Autophagy and the Integrated Stress Response. *Molecular Cell*, *40*(2), 280–293. https://doi.org/10.1016/J.MOLCEL.2010.09.023

Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. https://doi.org/10.1038/nrg2843

Kuersten, S., Radek, A., Vogel, C., & Penalva, L. O. F. (2013). Translation regulation gets its "omics" moment. *Wiley Interdisciplinary Reviews. RNA*, *4*(6), 617–630. https://doi.org/10.1002/wrna.1173

Kumar, M. S., Erkeland, S. J., Pester, R. E., Chen, C. Y., Ebert, M. S., Sharp, P. A., & Jacks, T. (2008). Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(10), 3903–3908. https://doi.org/10.1073/pnas.0712321105

Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R., & Jacks, T. (2007). Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genetics*, *39*(5), 673–677. https://doi.org/10.1038/ng2003

Lanford, R. E., Hildebrandt-Eriksen, E. S., Petri, A., Persson, R., Lindow, M., Munk, M. E., … Orum, H. (2010). Therapeutic Silencing of MicroRNA-122 in Primates with Chronic Hepatitis C Virus Infection. *Science*, *327*(5962), 198–201. https://doi.org/10.1126/science.1178178

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, *75*(5), 843–854. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8252621

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., … Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415–419. https://doi.org/10.1038/nature01957

Lee, Y., Jeon, K., Lee, J.-T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, *21*(17), 4663–4670. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12198168

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005a). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell*, *120*(1), 15–20. https://doi.org/10.1016/j.cell.2004.12.035

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005b). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell*, *120*(1), 15–20. https://doi.org/10.1016/j.cell.2004.12.035

Li, D., Liu, X., Lin, L., Hou, J., Li, N., Wang, C., … Cao, X. (2011). MicroRNA-99a inhibits hepatocellular carcinoma growth and correlates with prognosis of patients with hepatocellular carcinoma. *The Journal of Biological Chemistry*, *286*(42), 36677–36685. https://doi.org/10.1074/jbc.M111.270561

Li, H., Rauch, T., Chen, Z. X., Szabó, P. E., Riggs, A. D., & Pfeifer, G. P. (2006). The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M513249200

Li, L.-C., & Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics (Oxford, England)*, *18*(11), 1427–1431. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12424112

Li, Y., Wang, Y., Fan, H., Zhang, Z., & Li, N. (2018). miR-125b-5p inhibits breast cancer cell proliferation, migration and invasion by targeting KIAA1522. *Biochemical and Biophysical Research Communications*, *504*(1), 277–282. https://doi.org/10.1016/j.bbrc.2018.08.172

Li, Z., Huang, H., Chen, P., He, M., Li, Y., Arnovitz, S., … Chen, J. (2011). miR-196b directly targets both HOXA9/MEIS1 oncogenes and FAS tumour suppressor in MLL-rearranged leukaemia. *Nature Communications*, *2*(1), 688. https://doi.org/10.1038/ncomms1681

Li, Z., Owonikoko, T. K., Sun, S.-Y., Ramalingam, S. S., Doetsch, P. W., Xiao, Z.-Q., … Deng, X. (2012). c-Myc suppression of DNA double-strand break repair. *Neoplasia (New York, N.Y.)*, *14*(12), 1190–1202. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23308051

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., … Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, *433*(7027), 769–773. https://doi.org/10.1038/nature03315

Lin, K.-Y., Zhang, X.-J., Feng, D.-D., Zhang, H., Zeng, C.-W., Han, B.-W., … Chen, Y.-Q. (2011). miR-125b, a target of CDX2, regulates cell differentiation through repression of the core binding factor in hematopoietic malignancies. *The Journal of Biological Chemistry*, *286*(44), 38253–38263. https://doi.org/10.1074/jbc.M111.269670

Lin, R.-J., Lin, Y.-C., Chen, J., Kuo, H.-H., Chen, Y.-Y., Diccianni, M. B., … Yu, A. L. (2010). microRNA Signature and Expression of Dicer and Drosha Can Predict Prognosis and Delineate Risk Groups in Neuroblastoma. *Cancer Research*, *70*(20), 7841–7850. https://doi.org/10.1158/0008-5472.CAN-10-0970

Lind, G. E., Skotheim, R. I., & Lothe, R. a. (2007). The epigenome of testicular germ cell tumors. *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica*.

https://doi.org/10.1111/j.1600-0463.2007.apm_660.xml.x

Lindow, M., & Kauppinen, S. (2012). Discovering the first microRNA-targeted drug. *The Journal of Cell Biology*, *199*(3), 407–412. https://doi.org/10.1083/jcb.201208082

Linsley, P. S., Schelter, J., Burchard, J., Kibukawa, M., Martin, M. M., Bartz, S. R., … Lim, L. (2007). Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Molecular and Cellular Biology*, *27*(6), 2240–2252. https://doi.org/10.1128/MCB.02005-06

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., & Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. *Nucleic Acids Research*, *41*(1), 518–532. https://doi.org/10.1093/nar/gks1032

Loginov, V. I., Rykov, S. V., Fridman, M. V., & Braga, E. A. (2015). Methylation of miRNA genes and oncogenesis. *Biochemistry (Moscow)*, *80*(2), 145–162. https://doi.org/10.1134/S0006297915020029

Looijenga, L. H. (2009). Human testicular (non)seminomatous germ cell tumours: the clinical implications of recent pathobiological insights. *The Journal of Pathology*, *218*(2), 146–162. https://doi.org/10.1002/path.2522

Looijenga, L. H., Gillis, A. J., Van Putten, W. L., & Oosterhuis, J. W. (1993). In situ numeric analysis of centromeric regions of chromosomes 1, 12, and 15 of seminomas, nonseminomatous germ cell tumors, and carcinoma in situ of human testis. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *68*(2), 211–219. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8382754

Looijenga, L. H. J., Stoop, H., de Leeuw, H. P. J. C., de Gouveia Brazao, C. A., Gillis, A. J. M., van Roozendaal, K. E. P., … Oosterhuis, J. W. (2003). POU5F1 (OCT3/4) identifies cells with pluripotent potential in human germ cell tumors. *Cancer Research*, *63*(9), 2244–2250. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12727846

Lopez-Serra, P., & Esteller, M. (2012). DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene*, *31*(13), 1609–1622. https://doi.org/10.1038/onc.2011.354

Lorch, A., Kleinhans, A., Kramar, A., Kollmannsberger, C. K., Hartmann, J. T., Bokemeyer, C., … Beyer, J. (2012). Sequential versus single high-dose chemotherapy in patients with relapsed or refractory germ cell tumors: long-term results of a prospective randomized trial. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical*

Oncology, *30*(8), 800–805. https://doi.org/10.1200/JCO.2011.38.6391

Lorente, A., Mueller, W., Urdangarín, E., Lázcoz, P., von Deimling, A., & Castresana, J. S. (2008). Detection of methylation in promoter sequences by melting curve analysis-based semiquantitative real time PCR. *BMC Cancer*, *8*, 61. https://doi.org/10.1186/1471-2407-8-61

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., … Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, *435*(7043), 834–838. https://doi.org/10.1038/nature03702

Lu, Y., Zhao, X., Liu, Q., Li, C., Graves-Deal, R., Cao, Z., … Coffey, R. J. (2017). lncRNA MIR100HG-derived miR-100 and miR-125b mediate cetuximab resistance via Wnt/β-catenin signaling. *Nature Medicine*, *23*(11), 1331–1341. https://doi.org/10.1038/nm.4424

Lujambio, A., Ropero, S., Ballestar, E., Fraga, M. F., Cerrato, C., Setien, F., … Esteller, M. (2007). Genetic Unmasking of an Epigenetically Silenced microRNA in Human Cancer Cells. *Cancer Research*, *67*(4), 1424–1429. https://doi.org/10.1158/0008-5472.CAN-06-4218

Luo, Y., Na, Z., & Slavoff, S. A. (2018). P-Bodies: Composition, Properties, and Functions. *Biochemistry*, *57*(17), 2424–2431. https://doi.org/10.1021/acs.biochem.7b01162

Luo, Y., Wang, X., Niu, W., Wang, H., Wen, Q., Fan, S., … Zhou, M. (2017). Elevated microRNA-125b levels predict a worse prognosis in HER2-positive breast cancer patients. *Oncology Letters*, *13*(2), 867–874. https://doi.org/10.3892/ol.2016.5482

Lytle, J. R., Yario, T. A., & 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*, *104*(23), 9667–9672. https://doi.org/10.1073/pnas.0703820104

M Dror Michaelson, MD, PhDWilliam K Oh, M. (2018). Epidemiology of and risk factors for testicular germ cell tumors - UpToDate. Retrieved September 9, 2018, from https://www.uptodate.com/contents/epidemiology-of-and-risk-factors-for-testicular-germ-cell-tumors#H4

Malumbres, M., & Barbacid, M. (2001). To cycle or not to cycle: a critical decision in cancer. *Nature Reviews Cancer*, *1*(3), 222–231. https://doi.org/10.1038/35106065

Manecksha, R. P., & Fitzpatrick, J. M. (2009). Epidemiology of testicular cancer. *BJU International*, *104*(9b), 1329–1333. https://doi.org/10.1111/j.1464-410X.2009.08854.x

Mann, J. R., Gray, E. S., Thornton, C., Raafat, F., Robinson, K., Collins, G. S., … UK Children's Cancer Study Group Experience. (2008). Mature and Immature Extracranial Teratomas in Children: The UK Children's Cancer Study Group Experience. *Journal of Clinical Oncology*, *26*(21), 3590–3597. https://doi.org/10.1200/JCO.2008.16.0622

Mann, J. R., Raafat, F., Robinson, K., Imeson, J., Gornall, P., Sokal, M., … Oakhill, A. (2000). The United Kingdom Children's Cancer Study Group's Second Germ Cell Tumor Study: Carboplatin, Etoposide, and Bleomycin Are Effective Treatment for Children With Malignant Extracranial Germ Cell Tumors, With Acceptable Toxicity. *Journal of Clinical Oncology*, *18*(22), 3809–3818. https://doi.org/10.1200/JCO.2000.18.22.3809

Matsuda, A., Chikashige, Y., Ding, D.-Q., Ohtsuki, C., Mori, C., Asakawa, H., … Hiraoka, Y. (2015). Highly condensed chromatins are formed adjacent to subtelomeric and decondensed silent chromatin in fission yeast. *Nature Communications*, *6*(1), 7753. https://doi.org/10.1038/ncomms8753

Matsui, Y., & Okamura, D. (2005). Mechanisms of germ-cell specification in mouse embryos. *BioEssays*, *27*(2), 136–143. https://doi.org/10.1002/bies.20178

Mattie, M. D., Benz, C. C., Bowers, J., Sensinger, K., Wong, L., Scott, G. K., … Haqq, C. (2006). Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Molecular Cancer*, *5*(1), 24. https://doi.org/10.1186/1476-4598-5-24

McCabe, M. T., Brandes, J. C., & Vertino, P. M. (2009). Cancer DNA methylation: molecular mechanisms and clinical implications. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *15*(12), 3927–3937. https://doi.org/10.1158/1078-0432.CCR-08-2784

McDougal, W. S. (William S. (2012). *Campbell-Walsh urology tenth edition review*. Elsevier/Saunders.

McGlynn, K. A., & Cook, M. B. (2009). Etiologic factors in testicular germ-cell tumors. *Future Oncology*, *5*(9), 1389–1402. https://doi.org/10.2217/fon.09.116

McGlynn, K. A., & Trabert, B. (2012). Adolescent and adult risk factors for testicular cancer. *Nature Reviews Urology*, *9*(6), 339–349. https://doi.org/10.1038/nrurol.2012.61

McIntyre, A., Summersgill, B., Lu, Y. J., Missiaglia, E., Kitazawa, S., Oosterhuis, J. W., … Shipley, J. (2007). Genomic copy number and expression patterns in testicular germ cell tumours. *British Journal of Cancer*, *97*(12), 1707–1712. https://doi.org/10.1038/sj.bjc.6604079

Melo, S. A., Moutinho, C., Ropero, S., Calin, G. A., Rossi, S., Spizzo, R., … Esteller, M. (2010). A Genetic Defect in Exportin-5 Traps Precursor MicroRNAs in the Nucleus of Cancer Cells. *Cancer Cell*, *18*(4), 303–315. https://doi.org/10.1016/j.ccr.2010.09.007

Menck, K., Sönmezer, C., Worst, T. S., Schulz, M., Dihazi, G. H., Streit, F., … Gross, J. C. (2017). Neutral sphingomyelinases control extracellular vesicles budding from the plasma membrane. *Journal of Extracellular Vesicles*, *6*(1), 1378056.

https://doi.org/10.1080/20013078.2017.1378056

Mestdagh, P., Lefever, S., Pattyn, F., Ridzon, D., Fredlund, E., Fieuw, A., … Vandesompele, J. (2011). The microRNA body map: dissecting microRNA function through integrative genomics. *Nucleic Acids Research*, *39*(20), e136–e136. https://doi.org/10.1093/nar/gkr646

Milne, T. A., Zhao, K., & Hess, J. L. (2009). Chromatin immunoprecipitation (ChIP) for analysis of histone modifications and chromatin-associated proteins. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-59745-418-6_21

Mori, N., Yokobori, T., Mimori, K., Sudo, T., Tanaka, F., Shibata, K., … Mori, M. (2011). MicroRNA miR-125b is a prognostic marker in human colorectal cancer. *International Journal of Oncology*, *38*(5), 1437–1443. https://doi.org/10.3892/ijo.2011.969

Motzer, R. J., Cooper, K., Geller, N. L., Pfister, D. G., Lin, S. Y., Bajorin, D., … Morse, M. (1990). Carboplatin, etoposide, and bleomycin for patients with poor-risk germ cell tumors. *Cancer*, *65*(11), 2465–2470.

Mullany, L. E., Herrick, J. S., Wolff, R. K., & Slattery, M. L. (2016). MicroRNA Seed Region Length Impact on Target Messenger RNA Expression and Survival in Colorectal Cancer. *PloS One*, *11*(4), e0154177. https://doi.org/10.1371/journal.pone.0154177

Murphy, B. L., Obad, S., Bihannic, L., Ayrault, O., Zindy, F., Kauppinen, S., & Roussel, M. F. (2013). Silencing of the miR-17~92 cluster family inhibits medulloblastoma progression. *Cancer Research*, *73*(23), 7068–7078. https://doi.org/10.1158/0008-5472.CAN-13-0927

Murray, M. J., Bailey, S., Raby, K. L., Saini, H. K., de Kock, L., Burke, G. A. A., … Tischkowitz, M. (2014). Serum levels of mature microRNAs in DICER1-mutated pleuropulmonary blastoma. *Oncogenesis*, *3*(2), e87–e87. https://doi.org/10.1038/oncsis.2014.1

Murray, M. J., Fern, L. A., Stark, D. P., Eden, T. O., & Nicholson, J. C. (2009). Breaking down barriers: improving outcomes for teenagers and young adults with germ cell tumours. *Oncology Reviews*, *3*(4), 201–206. https://doi.org/10.1007/s12156-009-0030-7

Murray, M. J., & Nicholson, J. C. (2010). Germ cell tumours in children and adolescents. *Paediatrics and Child Health*, *20*(3), 109–116. https://doi.org/10.1016/j.paed.2009.10.006

Murray, M. J., Nicholson, J. C., & Coleman, N. (2015a). Biology of childhood germ cell tumours, focussing on the significance of microRNAs. *Andrology*, *3*(1), 129–139. https://doi.org/10.1111/andr.277

Murray, M. J., Nicholson, J. C., & Coleman, N. (2015b). Biology of childhood germ cell

tumours, focussing on the significance of microRNAs. *Andrology*, *3*(1), 129–139. https://doi.org/10.1111/andr.277

Murray, M. J., Nicholson, J. C., & Coleman, N. (2015c). Biology of childhood germ cell tumours, focussing on the significance of microRNAs. *Andrology*, *3*(1), 129–139. https://doi.org/10.1111/andr.277

Murray, M. J., Saini, H. K., Siegler, C. A., Hanning, J. E., Barker, E. M., van Dongen, S., … Coleman, N. (2013). LIN28 Expression in malignant germ cell tumors downregulates let-7 and increases oncogene levels. *Cancer Research*, *73*(15), 4872–4884. https://doi.org/10.1158/0008-5472.CAN-12-2085

Murray MJ, Fern LA, Stark DP, Eden TO, N. J. (2009). Breaking Down Barriers: Improving Outcomes for Teenagers and Young Adults with Germ Cell Tumours. *Oncology Reviews*, *3*, 201–206.

Nagaraja, A. K., Creighton, C. J., Yu, Z., Zhu, H., Gunaratne, P. H., Reid, J. G., … Matzuk, M. M. (2010). A Link between mir-100 and FRAP1/mTOR in Clear Cell Ovarian Cancer. *Molecular Endocrinology*, *24*(2), 447–463. https://doi.org/10.1210/me.2009-0295

Nettersheim, D., Westernströer, B., Haas, N., Leinhaas, A., Brüstle, O., Schlatt, S., & Schorle, H. (2012). Establishment of a versatile seminoma model indicates cellular plasticity of germ cell tumor cells. *Genes Chromosomes and Cancer*. https://doi.org/10.1002/gcc.21958

Niu, M., Dai, X., Zou, W., Yu, X., Teng, W., Chen, Q., … Liu, P. (2017). Autophagy, Endoplasmic Reticulum Stress and the Unfolded Protein Response in Intracerebral Hemorrhage. *Translational Neuroscience*, *8*, 37–48. https://doi.org/10.1515/tnsci-2017-0008

Nonaka, D. (2009). Differential Expression of SOX2 and SOX17 in Testicular Germ Cell Tumors. *American Journal of Clinical Pathology*, *131*(5), 731–736. https://doi.org/10.1309/AJCP7MNCNBCRN8NO

Noor, D. A. M., Jeyapalan, J. N., Alhazmi, S., Carr, M., Squibb, B., Wallace, C., … Scotting, P. J. (2016). Genome-wide methylation analysis identifies genes silenced in non-seminoma cell lines. *Npj Genomic Medicine*, *1*(1), 15009. https://doi.org/10.1038/npjgenmed.2015.9

Norbury, C., & Nurse, P. (1992). *ANIMAL CELL CYCLES AND THEIR CONTROL. Annu. Rev. Biochem* (Vol. 61). Retrieved from www.annualreviews.org

O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., & Mendell, J. T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, *435*(7043), 839–843. https://doi.org/10.1038/nature03677

Obad, S., dos Santos, C. O., Petri, A., Heidenblad, M., Broom, O., Ruse, C., … Kauppinen, S. (2011). Silencing of microRNA families by seed-targeting tiny LNAs. *Nature Genetics*, *43*(4), 371–378. https://doi.org/10.1038/ng.786

Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., … Bestor, T. H. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*. https://doi.org/10.1038/nature05987

Ørom, U. A., Nielsen, F. C., & Lund, A. H. (2008). MicroRNA-10a Binds the 5′UTR of Ribosomal Protein mRNAs and Enhances Their Translation. *Molecular Cell*, *30*(4), 460–471. https://doi.org/10.1016/j.molcel.2008.05.001

Ottaviani, S., Stebbing, J., Frampton, A. E., Zagorac, S., Krell, J., de Giorgio, A., … Castellano, L. (2018). TGF-β induces miR-100 and miR-125b but blocks let-7a through LIN28B controlling PDAC progression. *Nature Communications*, *9*(1), 1845. https://doi.org/10.1038/s41467-018-03962-x

Pallesen, G., & Hamilton-Dutoit, S. J. (1988). Ki-1 (CD30) antigen is regularly expressed by tumor cells of embryonal carcinoma. *The American Journal of Pathology*, *133*(3), 446–450. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2849300

Palmer, R. D., Barbosa-Morais, N. L., Gooding, E. L., Muralidhar, B., Thornton, C. M., Pett, M. R., … Children's Cancer and Leukaemia Group. (2008). Pediatric Malignant Germ Cell Tumors Show Characteristic Transcriptome Profiles. *Cancer Research*, *68*(11), 4239–4247. https://doi.org/10.1158/0008-5472.CAN-07-5560

Palmer, R. D., Foster, N. A., Vowler, S. L., Roberts, I., Thornton, C. M., Hale, J. P., … Coleman, N. (2007). Malignant germ cell tumours of childhood: new associations of genomic imbalance. *British Journal of Cancer*, *96*(4), 667–676. https://doi.org/10.1038/sj.bjc.6603602

Palmer, R. D., Murray, M. J., Saini, H. K., van Dongen, S., Abreu-Goodger, C., Muralidhar, B., … Coleman, N. (2010). Malignant germ cell tumors display common microRNA profiles resulting in global changes in expression of messenger RNA targets. *Cancer Research*, *70*(7), 2911–2923. https://doi.org/10.1158/0008-5472.CAN-09-3301

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., … Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, *408*(6808), 86–89. https://doi.org/10.1038/35040556

Patricia M Baker, E. O. (2009). *Germ Cell Tumors of the Ovary-Gynecologic Pathology*. Elsevier. https://doi.org/10.1016/B978-0-443-06920-8.X5001-9

Pauniaho, S.-L., Salonen, J., Helminen, M., Vettenranta, K., Heikinheimo, M., & Heikinheimo,

O. (2012). The incidences of malignant gonadal and extragonadal germ cell tumors in males and females: a population-based study covering over 40 years in Finland. *Cancer Causes & Control*, *23*(12), 1921–1927. https://doi.org/10.1007/s10552-012-0069-9

Peng, Y., & Croce, C. M. (2016). The role of MicroRNAs in human cancer. *Signal Transduction and Targeted Therapy*, *1*(1), 15004. https://doi.org/10.1038/sigtrans.2015.4

Penn, A., Jenney, M. E. M., & Nicholson, J. C. (2014). Germ cell tumours in children and adolescents. *Paediatrics and Child Health*, *24*(4), 148–154. https://doi.org/10.1016/j.paed.2014.01.003

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), e45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11328886

Pierson, J., Hostager, B., Fan, R., & Vibhakar, R. (2008). Regulation of cyclin dependent kinase 6 by microRNA 124 in medulloblastoma. *Journal of Neuro-Oncology*, *90*(1), 1–7. https://doi.org/10.1007/s11060-008-9624-3

Pigazzi, M., Manara, E., Bresolin, S., Tregnago, C., Beghin, A., Baron, E., … Basso, G. (2013). MicroRNA-34b promoter hypermethylation induces CREB overexpression and contributes to myeloid transformation. *Haematologica*, *98*(4), 602–610. https://doi.org/10.3324/haematol.2012.070664

Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nature Biotechnology*, *28*(10), 1057–1068. https://doi.org/10.1038/nbt.1685

Poulos, C., Cheng, L., Zhang, S., Gersell, D. J., & Ulbright, T. M. (2006). Analysis of ovarian teratomas for isochromosome 12p: evidence supporting a dual histogenetic pathway for teratomatous elements. *Modern Pathology*, *19*(6), 766–771. https://doi.org/10.1038/modpathol.3800596

Poynter, J. N., Amatruda, J. F., & Ross, J. A. (2010). Trends in incidence and survival of pediatric and adolescent patients with germ cell tumors in the United States, 1975 to 2006. *Cancer*, *116*(20), 4882–4891. https://doi.org/10.1002/cncr.25454

Preusse, M., Theis, F. J., & Mueller, N. S. (2016). miTALOS v2: Analyzing tissue specific microRNA function. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0151771

Puissegur, M. P., Eichner, R., Quelen, C., Coyaud, E., Mari, B., Lebrigand, K., … Brousset, P. (2012). B-cell regulator of immunoglobulin heavy-chain transcription (Bright)/ARID3a is a direct target of the oncomir microRNA-125b in progenitor B-cells. *Leukemia*, *26*(10), 2224–2232. https://doi.org/10.1038/leu.2012.95

Qin, C., Huang, R.-Y., & Wang, Z.-X. (2015). Potential role of miR-100 in cancer diagnosis, prognosis, and therapy. *Tumor Biology*, *36*(3), 1403–1409.

https://doi.org/10.1007/s13277-015-3267-8

Qin, D., Wang, X., Li, Y., Yang, L., Wang, R., Peng, J., … Fan, G.-C. (2016). MicroRNA-223-5p and -3p Cooperatively Suppress Necroptosis in Ischemic/Reperfused Hearts. *The Journal of Biological Chemistry*, *291*(38), 20247–20259. https://doi.org/10.1074/jbc.M116.732735

Rabinowits, M., Souhami, L., Gil, R. A., Andrade, C. A., & Paiva, H. C. (1990). Increased pulmonary toxicity with bleomycin and cisplatin chemotherapy combinations. *American Journal of Clinical Oncology*, *13*(2), 132–138. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1690503

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., & Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, *470*(7333), 279–283. https://doi.org/10.1038/nature09692

Ramassone, A., Pagotto, S., Veronese, A., & Visone, R. (2018). Epigenetics and MicroRNAs in Cancer. *International Journal of Molecular Sciences*, *19*(2), 459. https://doi.org/10.3390/ijms19020459

Rashid, H.-O., Yadav, R. K., Kim, H.-R., & Chae, H.-J. (2015). ER stress: Autophagy induction, inhibition and selection. *Autophagy*, *11*(11), 1956–1977. https://doi.org/10.1080/15548627.2015.1091141

Reid, G., Kao, S. C., Pavlakis, N., Brahmbhatt, H., MacDiarmid, J., Clarke, S., … van Zandwijk, N. (2016). Clinical development of TargomiRs, a miRNA mimic-based treatment for patients with recurrent thoracic cancer. *Epigenomics*, *8*(8), 1079–1085. https://doi.org/10.2217/epi-2016-0035

Reid, G., Pel, M. E., Kirschner, M. B., Cheng, Y. Y., Mugridge, N., Weiss, J., … van Zandwijk, N. (2013). Restoring expression of miR-16: a novel approach to therapy for malignant pleural mesothelioma. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO*, *24*(12), 3128–3135. https://doi.org/10.1093/annonc/mdt412

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., … Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, *403*(6772), 901–906. https://doi.org/10.1038/35002607

Richiardi, L., Bellocco, R., Adami, H.-O., Torrång, A., Barlow, L., Hakulinen, T., … Akre, O. (2004). Testicular cancer incidence in eight northern European countries: secular and recent trends. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *13*(12), 2157–2166. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15598775

Rijlaarsdam, M. A., Tax, D. M. J., Gillis, A. J. M., Dorssers, L. C. J., Koestler, D. C., de Ridder, J., & Looijenga, L. H. J. (2015). Genome wide DNA methylation profiles provide clues to the origin and pathogenesis of germ cell tumors. *PloS One*, *10*(4), e0122146. https://doi.org/10.1371/journal.pone.0122146

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb Protoc*, *2010*(6), pdb.prot5439. https://doi.org/10.1101/pdb.prot5439

Ririe, K. M., Rasmussen, R. P., & Wittwer, C. T. (1997). Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Analytical Biochemistry*, *245*(2), 154–160. https://doi.org/10.1006/abio.1996.9916

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. https://doi.org/10.1093/nar/gkv007

Rivas, F. V, Tolia, N. H., Song, J.-J., Aragon, J. P., Liu, J., Hannon, G. J., & Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Structural & Molecular Biology*, *12*(4), 340–349. https://doi.org/10.1038/nsmb918

Roberts, A. P. E., Lewis, A. P., & Jopling, C. L. (2011). miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Research*, *39*(17), 7716–7729. https://doi.org/10.1093/nar/gkr426

Roden, C., Gaillard, J., Kanoria, S., Rennie, W., Barish, S., Cheng, J., … Lu, J. (2017). Novel determinants of mammalian primary microRNA processing revealed by systematic evaluation of hairpin-containing transcripts and human genetic variation. *Genome Research*, *27*(3), 374–384. https://doi.org/10.1101/gr.208900.116

Rodriguez, S., Jafer, O., Goker, H., Summersgill, B. M., Zafarana, G., Gillis, A. J. M., … Shipley, J. M. (2003). Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i(12p) and amplification at 12p11.2-p12.1. *Oncogene*, *22*(12), 1880–1891. https://doi.org/10.1038/sj.onc.1206302

Rosa, A., Ballarino, M., Sorrentino, A., Sthandier, O., De Angelis, F. G., Marchioni, M., … Bozzoni, I. (2007). The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(50), 19849–19854. https://doi.org/10.1073/pnas.0706963104

Rose, N. R., & Klose, R. J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. https://doi.org/10.1016/j.bbagrm.2014.02.007

Rosen, A., Jayram, G., Drazer, M., & Eggener, S. E. (2011). Global Trends in Testicular Cancer

Incidence and Mortality. *European Urology*, *60*(2), 374–379. https://doi.org/10.1016/j.eururo.2011.05.004

Ruby, J. G., Jan, C. H., & Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature*, *448*(7149), 83–86. https://doi.org/10.1038/nature05983

Rupaimoole, R., Calin, G. A., Lopez-Berestein, G., & Sood, A. K. (2016). miRNA Deregulation in Cancer Cells and the Tumor Microenvironment. *Cancer Discovery*, *6*(3), 235–246. https://doi.org/10.1158/2159-8290.CD-15-0893

Sacconi, A., Biagioni, F., Canu, V., Mori, F., Di Benedetto, A., Lorenzon, L., … Blandino, G. (2012). miR-204 targets Bcl-2 expression and enhances responsiveness of gastric cancer. *Cell Death & Disease*, *3*(11), e423. https://doi.org/10.1038/cddis.2012.160

Saito, Y., & Jones, P. M. (2006). Epigenetic Activation of Tumor Suppressor MicroRNAs in Human Cancer Cells. *Cell Cycle*, *5*(19), 2220–2222. https://doi.org/10.4161/cc.5.19.3340

Saito, Y., Liang, G., Egger, G., Friedman, J. M., Chuang, J. C., Coetzee, G. A., & Jones, P. A. (2006a). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell*, *9*(6), 435–443. https://doi.org/10.1016/j.ccr.2006.04.020

Saito, Y., Liang, G., Egger, G., Friedman, J. M., Chuang, J. C., Coetzee, G. A., & Jones, P. A. (2006b). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell*, *9*(6), 435–443. https://doi.org/10.1016/J.CCR.2006.04.020

Sakai, E., Miura, Y., Suzuki-Kouyama, E., Oka, K., Tachibana, M., Kawabata, K., … Mizuguchi, H. (2017). A mammalian mirtron miR-1224 promotes tube-formation of human primary endothelial cells by targeting anti-angiogenic factor epsin2. *Scientific Reports*, *7*(1), 5541. https://doi.org/10.1038/s41598-017-05782-3

Saltzman, A. F., & Cost, N. G. (2018). Adolescent and Young Adult Testicular Germ Cell Tumors: Special Considerations. *Advances in Urology*, *2018*, 1–8. https://doi.org/10.1155/2018/2375176

Samuel, P., Mulcahy, L. A., Furlong, F., McCarthy, H. O., Brooks, S. A., Fabbri, M., … Carter, D. R. F. (2018). Cisplatin induces the release of extracellular vesicles from ovarian cancer cells that can induce invasiveness and drug resistance in bystander cells. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1737), 20170065. https://doi.org/10.1098/rstb.2017.0065

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, *94*(3), 441–448. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1100841

Sarraf, S. A., & Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2004.06.043

Sasidharan Nair, V., El Salhat, H., Taha, R. Z., John, A., Ali, B. R., & Elkord, E. (2018). DNA methylation and repressive H3K9 and H3K27 trimethylation in the promoter regions of PD-1, CTLA-4, TIM-3, LAG-3, TIGIT, and PD-L1 genes in human primary breast cancer. *Clinical Epigenetics*. https://doi.org/10.1186/s13148-018-0512-1

Sax, K., & Enzmann, E. V. (1939). The Effect of Temperature on X-Ray Induced Chromosome Aberrations. *Proceedings of the National Academy of Sciences of the United States of America*, *25*(8), 397–405. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16577923

Schabel, F. M., Trader, M. W., Laster, W. R., Corbett, T. H., & Griswold, D. P. (1979). cis-Dichlorodiammineplatinum(II): combination chemotherapy and cross-resistance studies with tumors of mice. *Undefined*. Retrieved from https://www.semanticscholar.org/paper/cis-Dichlorodiammineplatinum(II)%3A-combination-and-Schabel-Trader/4c22a17b6f66012815be161945eba4bca70e4e84

Schonberger, S., Okpanyi, V., Alemazkour, K., Looijenga, L. H. J., Nicholson, J., Borkhardt, A., & Schneider, D. T. (2010). Extracellular regulators of the Wnt signaling pathway in childhood germ cell tumors: Methylation of the SFRP2 promoter leads to Wnt activation and beta-catenin accumulation. *Pediatric Blood & Cancer*.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., … Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & Development*, *18*(11), 1251–1262. https://doi.org/10.1101/gad.300704

Schwarzenbach, H., da Silva, A. M., Calin, G., & Pantel, K. (2015). Data Normalization Strategies for MicroRNA Quantification. *Clinical Chemistry*, *61*(11), 1333–1342. https://doi.org/10.1373/clinchem.2015.239459

Scott, G. K., Goga, A., Bhaumik, D., Berger, C. E., Sullivan, C. S., & Benz, C. C. (2007). Coordinate Suppression of *ERBB2* and *ERBB3* by Enforced Expression of Micro-RNA *miR-125a* or *miR-125b*. *Journal of Biological Chemistry*, *282*(2), 1479–1486. https://doi.org/10.1074/jbc.M609383200

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature*, *455*(7209), 58–63. https://doi.org/10.1038/nature07228

Shen, H., Shih, J., Hollern, D. P., Wang, L., Bowlby, R., Tickoo, S. K., … Hoadley, K. A. (2018). Integrated Molecular Characterization of Testicular Germ Cell Tumors. *Cell*

Reports, *23*(11), 3392–3406. https://doi.org/10.1016/j.celrep.2018.05.039

Shi, L., Zhang, J., Pan, T., Zhou, J., Gong, W., Liu, N., … You, Y. (2010). MiR-125b is critical for the suppression of human U251 glioma stem cell proliferation. *Brain Research*, *1312*, 120–126. https://doi.org/10.1016/J.BRAINRES.2009.11.056

Sievers, S., Alemazkour, K., Zahn, S., Perlman, E. J., Gillis, A. J. M., Looijenga, L. H. J., … Schneider, D. T. (2005). IGF2/H19 imprinting analysis of human germ cell tumors (GCTs) using the methylation-sensitive single-nucleotide primer extension method reflects the origin of GCTs in different stages of primordial germ cell development. *Genes Chromosomes and Cancer*. https://doi.org/10.1002/gcc.20237

Slap, G. B., Lehmann, C., & Biro, F. M. (2008). Testicular and Scrotal Disorders. *Adolescent Medicine*, 140–145. https://doi.org/10.1016/B978-032304073-0.10020-2

Smiraglia, D. J., Rush, L. J., Frühwald, M. C., Dai, Z., Held, W. a, Costello, J. F., … Plass, C. (2001). Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/10.13.1413

Smiraglia, D. J., Szymanska, J., Kraggerud, S. M., Lothe, R. A., Peltomäki, P., & Plass, C. (2002). Distinct epigenetic phenotypes in seminomatous and nonseminomatous testicular germ cell tumors. *Oncogene*, *21*(24), 3909–3916. https://doi.org/10.1038/sj.onc.1205488

Smith-Sørensen, B., Lind, G. E., Skotheim, R. I., Fosså, S. D., Fodstad, Ø., Stenwig, A.-E., … Lothe, R. A. (2002). Frequent promoter hypermethylation of the O 6 -Methylguanine-DNA Methyltransferase (MGMT) gene in testicular cancer. *Oncogene*, *21*(57), 8878–8884. https://doi.org/10.1038/sj.onc.1205978

Song, T., Xia, W., Shao, N., Zhang, X., Wang, C., Wu, Y., … Li, H. (2010). Differential miRNA expression profiles in bladder urothelial carcinomas. *Asian Pacific Journal of Cancer Prevention*.

Srivastava, N., Manvati, S., Srivastava, A., Pal, R., Kalaiarasan, P., Chattopadhyay, S., … Bamezai, R. N. (2011). miR-24-2 controls H2AFX expression regardless of gene copy number alteration and induces apoptosis by targeting antiapoptotic gene BCL-2: a potential for therapeutic intervention. *Breast Cancer Research*, *13*(2), R39. https://doi.org/10.1186/bcr2861

Stang, A., Trabert, B., Wentzensen, N., Cook, M. B., Rusner, C., Oosterhuis, J. W., & McGlynn, K. A. (2012). Gonadal and extragonadal germ cell tumours in the United States, 1973-2007. *International Journal of Andrology*, *35*(4), 616–625. https://doi.org/10.1111/j.1365-2605.2011.01245.x

Stewart, M. D., Li, J., & Wong, J. (2005). Relationship between Histone H3 Lysine 9 Methylation, Transcription Repression, and Heterochromatin Protein 1 Recruitment. *Molecular and Cellular Biology*, *25*(7), 2525–2538.

https://doi.org/10.1128/MCB.25.7.2525-2538.2005

Stresemann, C., & Lyko, F. (2008). Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *International Journal of Cancer*, *123*(1), 8–13. https://doi.org/10.1002/ijc.23607

Su, Z., Yang, Z., Xu, Y., Chen, Y., Yu, Q., Su, Z., … Yu, Q. (2015). MicroRNAs in apoptosis, autophagy and necroptosis. *Oncotarget*, *6*(11), 8474–8490. https://doi.org/10.18632/oncotarget.3523

Sui, M., Jiao, A., Zhai, H., Wang, Y., Wang, Y., Sun, D., & Li, P. (2017). Upregulation of miR-125b is associated with poor prognosis and trastuzumab resistance in HER2-positive gastric cancer. *Experimental and Therapeutic Medicine*, *14*(1), 657–663. https://doi.org/10.3892/etm.2017.4548

Sun, F., Fu, H., Liu, Q., Tie, Y., Zhu, J., Xing, R., … Zheng, X. (2008). Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS Letters*, *582*(10), 1564–1568. https://doi.org/10.1016/j.febslet.2008.03.057

Suzuki, H., Maruyama, R., Yamamoto, E., & Kai, M. (2012). DNA methylation and microRNA dysregulation in cancer. *Molecular Oncology*, *6*(6), 567–578. https://doi.org/10.1016/j.molonc.2012.07.007

Suzuki, S., Murakami, Y., & Takahata, S. (2017). H3K36 methylation state and associated silencing mechanisms. *Transcription*, *8*(1), 26–31. https://doi.org/10.1080/21541264.2016.1246076

Tagawa, H., & Seto, M. (2005). A microRNA cluster as a target of genomic amplification in malignant lymphoma. *Leukemia*, *19*(11), 2013–2016. https://doi.org/10.1038/sj.leu.2403942

Takahashi, R.-U., Miyazaki, H., & Ochiya, T. (2014). The role of microRNAs in the regulation of cancer stem cells. *Frontiers in Genetics*, *4*, 295. https://doi.org/10.3389/fgene.2013.00295

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., … Takahashi, T. (2004). Reduced Expression of the *let-7* MicroRNAs in Human Lung Cancers in Association with Shortened Postoperative Survival. *Cancer Research*, *64*(11), 3753–3756. https://doi.org/10.1158/0008-5472.CAN-04-0637

Takeshige, K., Baba, M., Tsuboi, S., Noda, T., & Ohsumi, Y. (n.d.). Autophagy in Yeast Demonstrated with Proteinase-deficient Mutants and Conditions for its Induction. https://doi.org/10.1083/jcb.119.2.301

Takeshita, F., Patrawala, L., Osaki, M., Takahashi, R., Yamamoto, Y., Kosaka, N., … Ochiya, T. (2010). Systemic Delivery of Synthetic MicroRNA-16 Inhibits the Growth of

Metastatic Prostate Tumors via Downregulation of Multiple Cell-cycle Genes. *Molecular Therapy*, *18*(1), 181–187. https://doi.org/10.1038/MT.2009.207

Tamaru, H., & Selker, E. U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. *Nature*. https://doi.org/10.1038/35104508

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, *45*(W1), W98–W102. https://doi.org/10.1093/nar/gkx247

Tassano, E., Acquila, M., Tavella, E., Micalizzi, C., Panarello, C., & Morerio, C. (2010). MicroRNA-125b-1 and BLID upregulation resulting from a novel IGH translocation in childhood B-Cell precursor acute lymphoblastic leukemia. *Genes, Chromosomes and Cancer*, *49*(8), 682–687. https://doi.org/10.1002/gcc.20776

Tay, Y., Zhang, J., Thomson, A. M., Lim, B., & Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, *455*(7216), 1124–1128. https://doi.org/10.1038/nature07299

Teilum, G. (1965). Classification of endodermal sinus tumour (mesoblatoma vitellinum) and so-called "embryonal carcinoma" of the ovary. *Acta Pathologica et Microbiologica Scandinavica*, *64*(4), 407–429. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5318716

Terashima, K., Yu, A., Chow, W.-Y. T., Hsu, W. J., Chen, P., Wong, S., … Lau, C. C. (2014a). Genome-wide analysis of DNA copy number alterations and loss of heterozygosity in intracranial germ cell tumors. *Pediatric Blood & Cancer*, *61*(4), 593–600. https://doi.org/10.1002/pbc.24833

Terashima, K., Yu, A., Chow, W.-Y. T., Hsu, W. J., Chen, P., Wong, S., … Lau, C. C. (2014b). Genome-wide analysis of DNA copy number alterations and loss of heterozygosity in intracranial germ cell tumors. *Pediatric Blood & Cancer*, *61*(4), 593–600. https://doi.org/10.1002/pbc.24833

Tivnan, A., Orr, W. S., Gubala, V., Nooney, R., Williams, D. E., McDonagh, C., … Stallings, R. L. (2012). Inhibition of Neuroblastoma Tumor Growth by Targeted Delivery of MicroRNA-34a Using Anti-Disialoganglioside GD2 Coated Nanoparticles. *PLoS ONE*, *7*(5), e38129. https://doi.org/10.1371/journal.pone.0038129

Torres, A. G., Fabani, M. M., Vigorito, E., Williams, D., Al-Obaidi, N., Wojciechowski, F., … Gait, M. J. (2012). Chemical structure requirements and cellular targeting of microRNA-122 by peptide nucleic acids anti-miRs. *Nucleic Acids Res*, *40*(5), 2152–2167. https://doi.org/10.1093/nar/gkr885

Toyota, M., Suzuki, H., Sasaki, Y., Maruyama, R., Imai, K., Shinomura, Y., & Tokino, T. (2008). Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is

associated with CpG island methylation in colorectal cancer. *Cancer Research*, *68*(11), 4123–4132. https://doi.org/10.1158/0008-5472.CAN-08-0325

Trama, A., & Berrino, F. (2017). The Epidemiology of Malignant Germ Cell Tumors: The EUROCARE Study. *Pathology and Biology of Human Germ Cell Tumors*. https://doi.org/10.1007/978-3-662-53775-6_2

Trang, P., Medina, P. P., Wiggins, J. F., Ruffino, L., Kelnar, K., Omotola, M., … Slack, F. J. (2010). Regression of murine lung tumors by the let-7 microRNA. *Oncogene*, *29*(11), 1580–1587. https://doi.org/10.1038/onc.2009.445

Travis, L. B., Fosså, S. D., Schonfeld, S. J., McMaster, M. L., Lynch, C. F., Storm, H., … Gilbert, E. S. (2005). Second Cancers Among 40 576 Testicular Cancer Patients: Focus on Long-term Survivors. *JNCI: Journal of the National Cancer Institute*, *97*(18), 1354–1365. https://doi.org/10.1093/jnci/dji278

Tyson, J. J., & Novak, B. (2014). Control of cell growth, division and death: information processing in living cells. *Interface Focus*, *4*(3), 20130070. https://doi.org/10.1098/rsfs.2013.0070

Ulbright, T. M. (2005). Germ cell tumors of the gonads: a selective review emphasizing problems in differential diagnosis, newly appreciated, and controversial issues. *Modern Pathology*, *18*(S2), S61–S79. https://doi.org/10.1038/modpathol.3800310

Vaksman, O., Hetland, T. E., Trope', C. G., Reich, R., & Davidson, B. (2012). Argonaute, Dicer, and Drosha are up-regulated along tumor progression in serous ovarian carcinoma. *Human Pathology*, *43*(11), 2062–2069. https://doi.org/10.1016/J.HUMPATH.2012.02.016

Van Der Ree, M., de Vree, M., Stelma, F., Willemse, S., van der Valk, M., Rietdijk, S., … Reesink, H. (2015). *LO7: A single subcutaneous dose of 2mg/kg or 4mg/kg of rg-101, a galnac-conjugated oligonucleotide with antagonist activity against mir-122, results in significant viral load reductions in chronic hepatitis C patients*. *Journal of Hepatology* (Vol. 62). https://doi.org/10.1016/S0168-8278(15)30153-7

van der Ree, M. H., van der Meer, A. J., van Nuenen, A. C., de Bruijne, J., Ottosen, S., Janssen, H. L., … Reesink, H. W. (2016). Miravirsen dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Alimentary Pharmacology & Therapeutics*, *43*(1), 102–113. https://doi.org/10.1111/apt.13432

van der Zwan, Y. G., Rijlaarsdam, M. A., Rossello, F. J., Notini, A. J., de Boer, S., Watkins, D. N., … Looijenga, L. H. J. (2014). Seminoma and embryonal carcinoma footprints identified by analysis of integrated genome-wide epigenetic and expression profiles of germ cell cancer cell lines. *PloS One*, *9*(6), e98330.

https://doi.org/10.1371/journal.pone.0098330

van Dongen, S., Abreu-Goodger, C., & Enright, A. J. (2008a). Detecting microRNA binding and siRNA off-target effects from expression data. *Nature Methods*, *5*(12), 1023–1025. https://doi.org/10.1038/nmeth.1267

van Dongen, S., Abreu-Goodger, C., & Enright, A. J. (2008b). Detecting microRNA binding and siRNA off-target effects from expression data. *Nature Methods*, *5*(12), 1023–1025. https://doi.org/10.1038/nmeth.1267

van Dongen, S., Abreu-Goodger, C., & Enright, A. J. (2008c). Detecting microRNA binding and siRNA off-target effects from expression data. *Nature Methods*, *5*(12), 1023–1025. https://doi.org/10.1038/nmeth.1267

van Zandwijk, N., Pavlakis, N., Kao, S. C., Linton, A., Boyer, M. J., Clarke, S., … Reid, G. (2017a). Safety and activity of microRNA-loaded minicells in patients with recurrent malignant pleural mesothelioma: a first-in-man, phase 1, open-label, dose-escalation study. *The Lancet Oncology*, *18*(10), 1386–1396. https://doi.org/10.1016/S1470-2045(17)30621-6

van Zandwijk, N., Pavlakis, N., Kao, S. C., Linton, A., Boyer, M. J., Clarke, S., … Reid, G. (2017b). Safety and activity of microRNA-loaded minicells in patients with recurrent malignant pleural mesothelioma: a first-in-man, phase 1, open-label, dose-escalation study. *The Lancet Oncology*, *18*(10), 1386–1396. https://doi.org/10.1016/S1470-2045(17)30621-6

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), RESEARCH0034. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12184808

Vazquez, A., Bond, E. E., Levine, A. J., & Bond, G. L. (2008). The genetics of the p53 pathway, apoptosis and cancer therapy. *Nature Reviews Drug Discovery*, *7*(12), 979–987. https://doi.org/10.1038/nrd2656

Vermeulen, K., Van Bockstaele, D. R., & Berneman, Z. N. (2003). The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Proliferation*, *36*(3), 131–149. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12814430

Viswanathan, S. R., & Daley, G. Q. (2010a). Lin28: A MicroRNA Regulator with a Macro Role. *Cell*, *140*(4), 445–449. https://doi.org/10.1016/j.cell.2010.02.007

Viswanathan, S. R., & Daley, G. Q. (2010b). Lin28: A MicroRNA Regulator with a Macro Role. *Cell*, *140*(4), 445–449. https://doi.org/10.1016/j.cell.2010.02.007

Viswanathan, S. R., Daley, G. Q., & Gregory, R. I. (2008). Selective blockade of microRNA

processing by Lin28. *Science (New York, N.Y.)*, *320*(5872), 97–100. https://doi.org/10.1126/science.1154040

Volinia, S., Calin, G. A., Liu, C.-G., Ambs, S., Cimmino, A., Petrocca, F., … Croce, C. M. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2257–2261. https://doi.org/10.1073/pnas.0510565103

Volinia, S., Galasso, M., Costinean, S., Tagliavini, L., Gamberoni, G., Drusco, A., … Croce, C. M. (2010). Reprogramming of miRNA networks in cancer and leukemia. *Genome Research*, *20*(5), 589–599. https://doi.org/10.1101/gr.098046.109

Vrba, L., Muñoz-Rodríguez, J. L., Stampfer, M. R., & Futscher, B. W. (2013). miRNA Gene Promoters Are Frequent Targets of Aberrant DNA Methylation in Human Breast Cancer. *PLoS ONE*, *8*(1), e54398. https://doi.org/10.1371/journal.pone.0054398

Wahle, E., & Winkler, G. S. (2013). RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6–7), 561–570. https://doi.org/10.1016/J.BBAGRM.2013.01.003

Walz, A. L., Ooms, A., Gadd, S., Gerhard, D. S., Smith, M. A., Guidry Auvil, J. M., … Perlman, E. J. (2015). Recurrent DGCR8, DROSHA, and SIX Homeodomain Mutations in Favorable Histology Wilms Tumors. *Cancer Cell*, *27*(2), 286–297. https://doi.org/10.1016/j.ccell.2015.01.003

Wang, B., Hsu, S., Wang, X., Kutay, H., Bid, H. K., Yu, J., … Ghoshal, K. (2014). Reciprocal regulation of microRNA-122 and c-Myc in hepatocellular cancer: Role of E2F1 and transcription factor dimerization partner 2. *Hepatology*, *59*(2), 555–566. https://doi.org/10.1002/hep.26712

Wang, F., Fu, X.-D., Zhou, Y., & Zhang, Y. (2009). Down-regulation of the cyclin E1 oncogene expression by microRNA-16-1 induces cell cycle arrest in human cancer cells. *BMB Reports*, *42*(11), 725–730. https://doi.org/10.5483/BMBRep.2009.42.11.725

Wang, S., Wu, J., Ren, J., Vlantis, A. C., Li, M., Liu, S. Y. W., … Chen, G. G. (2018). MicroRNA-125b Interacts with Foxp3 to Induce Autophagy in Thyroid Cancer. *Molecular Therapy*. https://doi.org/10.1016/j.ymthe.2018.06.015

Warnefors, M., Liechti, A., Halbert, J., Valloton, D., & Kaessmann, H. (2014). Conserved microRNA editing in mammalian evolution, development and disease. *Genome Biology*, *15*(6), R83. https://doi.org/10.1186/gb-2014-15-6-r83

Watanabe, K., Emoto, N., Hamano, E., Sunohara, M., Kawakami, M., Kage, H., … Takai, D. (2012). Genome structure-based screening identified epigenetically silenced microRNA associated with invasiveness in non-small-cell lung cancer. *International Journal of Cancer*, *130*(11), 2580–2590. https://doi.org/10.1002/ijc.26254

Weaver, K. E., Forsythe, L. P., Reeve, B. B., Alfano, C. M., Rodriguez, J. L., Sabatino, S. A., … Rowland, J. H. (2012). Mental and Physical Health-Related Quality of Life among U.S. Cancer Survivors: Population Estimates from the 2010 National Health Interview Survey. *Cancer Epidemiology Biomarkers & Prevention*, *21*(11), 2108–2117. https://doi.org/10.1158/1055-9965.EPI-12-0740

Weber, B., Stresemann, C., Brueckner, B., & Lyko, F. (2007). Methylation of Human MicroRNA Genes in Normal and Neoplastic Cells. *Cell Cycle*, *6*(9), 1001–1005. https://doi.org/10.4161/cc.6.9.4209

Wermann, H., Stoop, H., Gillis, A. J., Honecker, F., van Gurp, R. J., Ammerpohl, O., … Looijenga, L. H. (2010). Global DNA methylation in fetal human germ cells and germ cell tumours: association with differentiation and cisplatin resistance. *The Journal of Pathology*, *221*(4), n/a-n/a. https://doi.org/10.1002/path.2725

Westholm, J. O., & Lai, E. C. (2011). Mirtrons: microRNA biogenesis via splicing. *Biochimie*, *93*(11), 1897–1904. https://doi.org/10.1016/j.biochi.2011.06.017

Wetherell, D., Weerakoon, M., Williams, D., Koonj Beharry, B., Sliwinski, A., Ow, D., … Lawrentschuk, N. (2014). Mature and Immature Teratoma: A Review of Pathological Characteristics and Treatment Options. *Med Surg Urol*, *3*, 124. https://doi.org/10.4172/2168-9857.1000124

Wiederschain, D., Wee, S., Chen, L., Loo, A., Yang, G., Huang, A., … Benson, J. D. (2009). Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle*. https://doi.org/10.4161/cc.8.3.7701

Williamson, S. R., Delahunt, B., Magi-Galluzzi, C., Algaba, F., Egevad, L., Ulbright, T. M., … Members of the ISUP Testicular Tumour Panel. (2017). The World Health Organization 2016 classification of testicular germ cell tumours: a review and update from the International Society of Urological Pathology Testis Consultation Panel. *Histopathology*, *70*(3), 335–346. https://doi.org/10.1111/his.13102

Winans, S., Flynn, A., Malhotra, S., Balagopal, V., & Beemon, K. L. (2017). Integration of ALV into CTDSPL and CTDSPL2 genes in B-cell lymphomas promotes cell immortalization, migration and survival. *Oncotarget*. https://doi.org/10.18632/oncotarget.19328

Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G., & Pryor, R. J. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry*, *49*(6 Pt 1), 853–860. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12765979

Wojdacz, T. K., & Dobrovic, A. (2007). Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Research*, *35*(6), e41–e41. https://doi.org/10.1093/nar/gkm013

Wu, H., Wang, F., Hu, S., Yin, C., Li, X., Zhao, S., … Yan, X. (2012). MiR-20a and miR-106b negatively regulate autophagy induced by leucine deprivation via suppression of ULK1 expression in C2C12 myoblasts. *Cellular Signalling*, *24*(11), 2179–2186. https://doi.org/10.1016/j.cellsig.2012.07.001

Wu, K., He, J., Pu, W., & Peng, Y. (2018). The Role of Exportin-5 in MicroRNA Biogenesis and Cancer. *Genomics, Proteomics & Bioinformatics*, *16*(2), 120–126. https://doi.org/10.1016/j.gpb.2017.09.004

Wu, N., Xiao, L., Zhao, X., Zhao, J., Wang, J., Wang, F., … Lin, X. (2012). miR-125b regulates the proliferation of glioblastoma stem cells by targeting E2F2. *FEBS Letters*, *586*(21), 3831–3839. https://doi.org/10.1016/J.FEBSLET.2012.08.023

Xu, T., Zhu, Y., Xiong, Y., Ge, Y.-Y., Yun, J.-P., & Zhuang, S.-M. (2009). MicroRNA-195 suppresses tumorigenicity and regulates G_1/S transition of human hepatocellular carcinoma cells. *Hepatology*, *50*(1), 113–121. https://doi.org/10.1002/hep.22919

Yamakuchi, M., & Lowenstein, C. J. (2009). MiR-34, SIRT1, and p53: The feedback loop. *Cell Cycle*, *8*(5), 712–715. https://doi.org/10.4161/cc.8.5.7753

Yang, G., Gong, Y., Wang, Q., Wang, Y., & Zhang, X. (2015). The role of miR-100-mediated Notch pathway in apoptosis of gastric tumor cells. *Cellular Signalling*, *27*(6), 1087–1101. https://doi.org/10.1016/j.cellsig.2015.02.013

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, *17*(24), 3011–3016. https://doi.org/10.1101/gad.1158803

Ying, Y., Liu, X.-M., Marble, A., Lawson, K. A., & Zhao, G.-Q. (2000). Requirement of *Bmp8b* for the Generation of Primordial Germ Cells in the Mouse. *Molecular Endocrinology*, *14*(7), 1053–1063. https://doi.org/10.1210/mend.14.7.0479

Yu, X., Shi, W., Zhang, Y., Wang, X., Sun, S., Song, Z., … Qu, X. (2017). CXCL12/CXCR4 axis induced miR-125b promotes invasion and confers 5-fluorouracil resistance through enhancing autophagy in colorectal cancer. *Scientific Reports*, *7*(1), 42226. https://doi.org/10.1038/srep42226

Yuan, M., Da Silva, A. C. A. L., Arnold, A., Okeke, L., Ames, H., Correa-Cerro, L. S., … Rodriguez, F. J. (2018). MicroRNA (miR) 125b regulates cell growth and invasion in pediatric low grade glioma. *Scientific Reports*, *8*(1), 12506. https://doi.org/10.1038/s41598-018-30942-4

Yuan, X., Liu, C., Yang, P., He, S., Liao, Q., Kang, S., & Zhao, Y. (2009). Clustered microRNAs' coordination in regulating protein-protein interaction network. *BMC Systems Biology*. https://doi.org/10.1186/1752-0509-3-65

Zhang, H., Li, Y., Huang, Q., Ren, X., Hu, H., Sheng, H., & Lai, M. (2011). MiR-148a promotes apoptosis by targeting Bcl-2 in colorectal cancer. *Cell Death & Differentiation*, *18*(11), 1702–1710. https://doi.org/10.1038/cdd.2011.28

Zhang, J., Zhang, H., Yao, Y.-F., Zhong, S.-L., Zhao, J. H., & Tang, J. H. (2015). β-Elemene Reverses Chemoresistance of Breast Cancer Cells by Reducing Resistance Transmission via Exosomes. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *36*(6), 2274–2286. https://doi.org/10.1159/000430191

Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M. S., Giannakakis, A., … Coukos, G. (2006a). microRNAs exhibit high frequency genomic alterations in human cancer. *Proceedings of the National Academy of Sciences*, *103*(24), 9136–9141. https://doi.org/10.1073/pnas.0508889103

Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M. S., Giannakakis, A., … Coukos, G. (2006b). microRNAs exhibit high frequency genomic alterations in human cancer. *Proceedings of the National Academy of Sciences*, *103*(24), 9136–9141. https://doi.org/10.1073/pnas.0508889103

Zhang, Y., Roccaro, A. M., Rombaoa, C., Flores, L., Obad, S., Fernandes, S. M., … Ghobrial, I. M. (2012). LNA-mediated anti-miR-155 silencing in low-grade B-cell lymphomas. *Blood*, *120*(8), 1678–1686. https://doi.org/10.1182/blood-2012-02-410647

Zhang, Y., Yan, L.-X., Wu, Q.-N., Du, Z.-M., Chen, J., Liao, D.-Z., … Shao, J.-Y. (2011a). miR-125b Is Methylated and Functions as a Tumor Suppressor by Regulating the ETS1 Proto-oncogene in Human Invasive Breast Cancer. *Cancer Research*, *71*(10), 3552–3562. https://doi.org/10.1158/0008-5472.CAN-10-2435

Zhang, Y., Yan, L.-X., Wu, Q.-N., Du, Z.-M., Chen, J., Liao, D.-Z., … Shao, J.-Y. (2011b). miR-125b Is Methylated and Functions as a Tumor Suppressor by Regulating the ETS1 Proto-oncogene in Human Invasive Breast Cancer. *Cancer Research*, *71*(10), 3552–3562. https://doi.org/10.1158/0008-5472.CAN-10-2435

Zhao, T., Liu, Y., Jiang, H., Zhang, H., & Lu, Y. (2016). Management of bilateral malignant ovarian germ cell tumors: Experience of a single institute. *Molecular and Clinical Oncology*, *5*(2), 383–387. https://doi.org/10.3892/mco.2016.915

Zheng, Y. S., Zhang, H., Zhang, X. J., Feng, D. D., Luo, X. Q., Zeng, C. W., … Chen, Y. Q. (2012). MiR-100 regulates cell differentiation and survival by targeting RBSP3, a phosphatase-like tumor suppressor in acute myeloid leukemia. *Oncogene*. https://doi.org/10.1038/onc.2011.208

Zhou, H., & Rigoutsos, I. (2014). MiR-103a-3p targets the 5′ UTR of *GPRC5A* in pancreatic cells. *RNA*, *20*(9), 1431–1439. https://doi.org/10.1261/rna.045757.114

Zhou, M., Liu, Z., Zhao, Y., Ding, Y., Liu, H., Xi, Y., … Tan, M. (2010). MicroRNA-125b Confers the Resistance of Breast Cancer Cells to Paclitaxel through Suppression of Pro-apoptotic Bcl-2 Antagonist Killer 1 (Bak1) Expression. *Journal of Biological Chemistry*, *285*(28), 21496–21507. https://doi.org/10.1074/jbc.M109.083337

Some more appendix

And even more

And some more text on a new page

And even more text

References

Some more appendix

And even more

And some more text on a new page

And even more text