DNA methylation and expression characterization of FOXD2 and FOXD2-AS1 genes in colorectal cancer patients

Final Master's Degree Project

Master's Degree in Clinical Laboratory (2020-2021)

MARIA LUCAS GASCÓN

June 25th, 2021

Tutor: Dr. Miguel A. Peinado

Coordinador: Dr. Mariano Sentí

ACKNOWLEDGEMENTS

First of all, I would like to thank all the laboratory of Cancer Epigenetics for giving me the opportunity to work in the laboratory during my practises and being in direct contact with the world of research.

Mar, thank you very much for the time you have spent in my scientific training, your patience, your advice, but especially for your kindness and dedication. You have explained to me some techniques as best as you know but leaving enough room for me to improve for myself, make mistakes and learn from them. During these months you have helped me every time I have needed you, ever solving all my doubts.

Júlia, thank you for your time and your dedication. Thank you for sharing your enthusiasm towards science. Thank you for taking care of me and for receiving me as part of the group from the first day. I could not feel more satisfied.

Berta, Marta, Núria and Sheila thank you for your friendliness. It has been so nice to work with all of you, I cannot imagine anyone better to share this experience with. I will never forget the days in the laboratory, it would not have been the same without you. I have lived a very rewarding experience I will never forget.

Last but not least, Miquel Angel, thank you for giving me the opportunity to work in your research group, thank you for your dedication and care. Nothing of this would have been possible without your trust in me.

TUTOR CERTIFICATE

I, Miquel A. Peinado, authorize the presentation of this work to the University of Pompeu Fabra.

_____ Date: June 25th, 2021

Signature

ABSTRACT

Cell transcription is highly regulated through complex mechanisms to maintain an adequate cell homeostasis. LncRNAs are one example of transcription factors regulating protein expression, and their deregulation can contribute to the acquisition of diseases like cancer. Methylation of gene promoters is also a mechanism that affects gene expression and has been observed to be altered in most cancers. We studied FOXD2 and FOXD2-AS1 expression and methylation in 300 paired normal-tumour colorectal samples. Analysing their expression by qRT-PCR and their methylation levels by bisulfite conversion and Sanger sequencing, we determined that their locus is hypermethylated in cancer samples in comparison to normal samples, and a decreased expression is also found in tumoral samples. Methylation of the locus was negatively correlated with FOXD2 and FOXD2-AS1 expression. We studied the association between these findings and clinicopathological features of the tumours, the only remarkable correlation found was that higher methylation levels of region 1 are related to bigger tumour sizes. Further understanding the molecular differences between normal and tumoral tissues is the first step in finding a treatment. In summary, our results revealed FOXD2 as a new possible target for CRC treatment, as well as a further understanding FOXD2-AS1 dysregulation in CRC.

TABLE OF CONTENT

1. INTRODUCTION

Cancer

Cancer is the group of diseases involving abnormal cell growth with the potential to invade and spread to other parts of the body. Over 100 types of cancers can affect humans (NIH, 2015a). Cancer is the second leading cause of death globally, accounting for an estimated 9.6 million deaths in 2018. Lung, prostate, colorectal, stomach and liver cancer are the most common types of cancer in men, while breast, colorectal, lung, cervical and thyroid cancer are the most common among women (WHO, 2018).

Hallmarks of cancer

The six hallmarks of cancer refer to the capabilities that enable tumour growth and metastatic dissemination (Hanahan and Weinberg, 2000). Probably the most fundamental trait of cancer cells is the ability to chronically sustain their proliferation, but it is also equally important to inhibit growth suppressing and apoptotic genes. In addition to abnormal cell division, cells must also resist cell death. Programmed cell death by apoptosis is known to be a natural barrier to cancer development, but it can be disrupted in some cancers (Lowe and Lin, 2000, Chiu and Harley, 1997). Proper vascularization is needed in order to obtain nutrients and evacuate metabolic waste, for this reason the ability to induce angiogenesis is a cancer hallmark (Hanahan and Folkman, 1996).

Activating invasion and metastasis is a widely known characteristic of cancer cells. When epithelial tissues progress to a higher pathological grade of malignancy, they develop alterations in their shape as well as alterations in the attachment to other cells and to the extracellular matrix (Schmalhofer et al., 2009). All these characteristics can be acquired thanks to an accumulation of mutations in the DNA sequence or through other alterations such as epigenetic changes (Negrini et al., 2010).

Cancer staging systems

The TNM staging system is one of the most widely used cancer staging systems. In this system the T refers to the size and extent of the main tumour or primary tumour, the higher the number the larger the tumour. The N refers to the number of nearby lymph nodes that have been invaded by tumoral cells, the higher the number more lymph nodes are affected. The M refers to whether the cancer has metastasized (M1) or not (M0).

The TNM system helps describe cancer in great detail. But, for many cancers, the TNM combinations are grouped into five less-detailed stages: Stage 0, when abnormal cells are present but have not spread to nearby tissues. Stage I-III: Cancer is present, the larger the number the more it has spread into nearby tissues. Stage IV: The cancer has spread to distant parts of the body (NIH, 2015b).

Epigenetics

Epigenetics is the study of chromatin modifications that do not change the DNA sequence but can affect gene expression. In the past five years several studies have focused their attention on epigenetics, as it became clear that understanding epigenetics was essential in order to understand all aspects of genetics. Gene expression in mammals depends on the epigenetic status of the chromatin, so it determines cell development, cell state and fate, cell cycle regulation, etc.

The most studied modifications are DNA methylation, histone modifications and non-coding RNAs (ncRNAs). DNA is highly compacted in three-dimensional chromosomes, DNA coils around histone octamers to form nucleosomes. This organization not only allows the DNA to compact, but also provides selective accessibility of transcription machinery. Histones are subjected to more than 130 posttranslational modifications such as acetylation, methylation, phosphorylation, sumoylation and ubiquitination. This, in conjunction with DNA methylation, decides the transcriptional states of all the genomic regions determining the transcriptional programs (Chen et al., 2017).

Approximately, three-quarters of the genome are transcribed into RNA, but only a small portion of genes are translated into proteins (Djebali et al., 2012). ncRNAs are RNAs that are not translated into proteins. These RNAs comprise another layer of internal signals that control gene expression, they may determine most of our complex characteristics, play a role in disease and in variation between individuals and species (Mattick and Makunin, 2006).

DNA methylation

The most common DNA modification is the cytidine 5'-addition of a methyl group in a CpG dinucleotide. There are CpG-rich regions called CpG islands, which are typically defined as regions of more than 200 nucleotides with a GC content of at least 50%, located in the 5' end of the regulatory region of many genes. They have critical roles in the control of gene activity and the architecture of the nucleus of the cell.

At first, it was thought that this modification was a stable, hereditary trait, but it has been recently discovered that methyl groups can be dynamically added or erased (Wu and Zhang, 2014). This is due to the existence of DNA methyltransferases that introduce the methyl group in the DNA, and DNA demethylases called Ten-Eleven Translocation enzymes (TET) that erase them.

In general, gene promoter CpG-island methylation is associated with gene silencing. It plays a key role in genomic imprinting, where hypermethylation at one of the two parental alleles leads to monoallelic expression. A similar mechanism is observed in gene-dosage reduction in X-chromosome inactivation in females (Li et al., 1993). DNA methylation is not exclusive to CpG islands. Non-CG methylation has recently been described in humans in stem cells and seems to play a critical role in maintenance of pluripotent state (Laurent et al., 2010)

Cancer epigenetics

Epigenetic processes are natural and essential to many organism functions, but if they occur improperly, there can be major adverse health and behavioural effects. Epigenetic mechanisms account one-third to one-half of known genetic alterations (Weinhold, 2006). Among all epigenetic research, the most studied disease is cancer. In addition to classic mutations, cancer also presents aberrant DNA methylation. It was first discovered in 1983 and has since been observed in many other illnesses and health conditions (Feinberg and Tycko, 2004).

The alteration of DNA methylation patterns leads to disease in the cells, normal patterns and its alterations are shown in the figure below (**Figure 1**). DNA methylation can inhibit or enhance gene expression by recruiting histone-modifying and chromatin remodelling complexes to methylated or unmethylated sites.

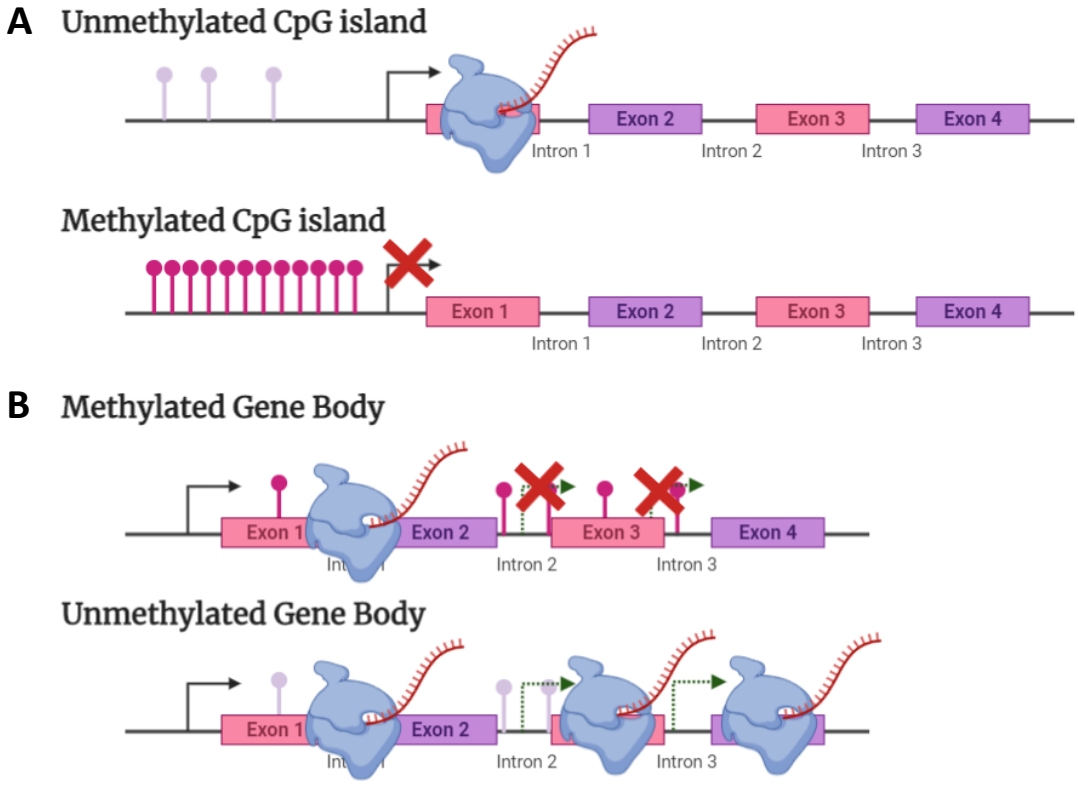


Figure 1: Schematic representation of normal and aberrant methylation patterns. (A) When CpG islands at gene promoters are unmethylated, transcription is allowed. In cancer, aberrant hypermethylation provokes gene silencing and thus lack of transcription. (B) Methylation at the gene body facilitates transcription and prevents aberrant transcription initiation, but in some diseases, gene body is unmethylated allowing transcription to be initiated at incorrect sites. Adapted from Portela and Esteller (2010).

Cancer cells are characterized by global hypomethylation, with approximately a 20-60% overall loss of methylation. It occurs mainly at repetitive sequences, and promotes chromosomal instability, translocation and gene disruption (Goelz et al., 1985). This loss of methylation can lead to the overexpression of oncogenes (genes with the potential to cause cancer) and

facilitate abnormal cell division. In contrast, some gene-specific promoters become hypermethylated, resulting in gene silencing (**Figure 1 A**). When a tumour suppressor gene is silenced because of promoter hypermethylation, it stops inhibiting cell proliferation and tumour development (Llinàs-Arias and Esteller, 2017). On the other hand, gene body methylation facilitates a correct transcription, and it prevents aberrant transcription confining transcription initiation to canonical promoters. However, in some diseases gene body methylation is lost leading to abnormal transcription (**Figure 1 B**). This genetic and epigenetic reprogramming shifts the cell from a normal phenotype to a malignant phenotype, contributing to the acquisition of all hallmarks of cancer cells.

Recent advances in techniques for methylation detection such as sodium bisulfite conversion and CpG island microarrays, have produced a vast amount of knowledge of methylation patterns in diverse human cancers. Tumour-specific methylation changes in different genes have been identified and documented, this information has potential clinical application in cancer diagnosis, prognosis and therapeutics (Das and Singal, 2004).

FOX genes

The FOX family or forkhead family of transcription factors is an expanding group of genes that binds to DNA through a highly conserved region of 105 amino acids. This domain called fork head has been identified in over 40 genes in a vast range of different organisms including humans. To date, the family of forkhead proteins has over 2000 members identified in 108 species of animals and fungi, and 19 subfamilies have been identified. Some experiments suggest that forkhead transcription factors regulate gene expression in an unconventional way for transcription factors, by regulating the local state of chromatin at target loci, independently of their abilities as direct transcriptional activators/repressors (Benayoun et al., 2011).

Although the biological function of these genes has not been fully understood, important roles in tumorigenesis, embryonic development and regulation of tissue-specific gene expression have been investigated. FOX family genes are involved in carcinogenesis as oncogenes and as tumour suppressor genes. Genetic alterations as well as epigenetic alterations in these genes are often found in various types of human cancer (Katoh et al., 2013).

The role of FOX proteins in cancer extends beyond transformation and may be used as prognostic markers. FOX proteins have been implicated in regulating multiple angiogenic factors, as well as interacting with key signalling pathways such as p53. These findings suggest that the FOX families may serve both as direct or indirect therapeutic targets (Myatt and Lam, 2007).

FOXD2

FOXD2 or freac-9 (fkd9) is a protein encoding gene located in chromosome 1 p32-34 and was first identified in 1997. Little is known about FOXD2 gene, its function and association with cancer.

It is known that FOXD2 in lymphocytes is a physiological regulator of the RI1b promoter and works with the protein kinase B to induce cAMP-dependent protein kinase RI expression, which increases cAMP sensitivity and sets the threshold for cAMP-mediated negative modulation of T cell activation (Johansson et al., 2003).

The expression patterns of FOXD2 in zebrafish and mice have been studied, being similar in both cases. This suggests that FOXD2 plays an important role in embryogenesis development. Remarkably, the expression patterns reported for zebrafish fkd9 and the mouse MF-2 (Foxd2) are strikingly similar (Odenthal and Nüsslein-Volhard, 1998; Wu et al., 1998). In both cases, it was concluded that the expression in the developing somites might be correlated with a function in sclerotome formation.

In another study, hybridization panels were used to test 50 different human tissues, FOXD2 mRNA was found only in the kidney, and the mRNA was present in two kidney-derived cell lines. Additionally, they observed that FOXD2 and FREAC-4 DNA binding forkheads motifs are identical, this could be crucial in determining their contribution to transcriptional regulation (Ernstsson et al., 1997).

Long noncoding RNAs

The transcriptional landscape of all organisms is far more complex than was originally imagined, long non-coding RNAs (lncRNA) contribute to this complexity. It is estimated that only 2% of the genes of the whole genome are protein encoding genes, and most of the other are non-coding RNAs (ncRNA) such as lncRNAs. LngRNAs are RNA molecules over 200nt in length that cannot be translated into proteins (Wilusz JE et al., 2009). Most lncRNAs are transcribed by RNA polymerase II and are capped, polyadenylated and spliced. The functions and mechanisms of lncRNAs are diverse, and are associated with their subcellular localization. LncRNAs are diverse and numerous, and the total number continues to increase as the technology advances (Quinn and Chang, 2016).

LncRNAs have regulatory roles in gene expression at both transcriptional and post-transcriptional levels. They are responsible for nuclear structure integrity and can interact with proteins, DNA or RNA in order to regulate gene expression. They are capable of recruiting transcription factors, chromatin-modifying complexes, form complexes with ribonucleoproteins, etc. (Chen, 2016). They also participate in loci genomic imprinting, cell state, disease, development and differentiation.

Although the functions of the majority of lncRNAs are unknown, it is recognised that overexpression, deficiency of mutation of lncRNA has been implicated in numerous human diseases such as cancer (Sun et al., 2018). LncRNAs are closely related to tumour resistance, cancer development, invasion and metastasis, as they can regulate expression of oncogenes, tumour suppressor genes and other cancer-associated genes (Sun et al., 2017).

FOXD2-AS1

The lncRNA known as FOXD2-AS1 is located at chromosome 1p33 and has been recently found aberrantly expressed in a wide range of cancers and associated with cancer progression (Zhang et al., 2020). Its overexpression has been observed in several cancers such as: nasopharyngeal carcinoma, hepatocellular carcinoma, gastric cancer, colorectal cancer, non-small lung cancer, etc. However, the association between FOXD2-AS1 expression and clinicopathological characteristics is controversial, some studies with small sample sizes have been performed. A recent meta-analysis concluded that high FOXD2-AS1 expression is related to poor overall survival and poor disease-free survival in solid tumours. FOXD2-AS1 expression is also associated with bigger tumour size and TNM stage in gastric cancer, non-small cell lung cancer and hepatocellular carcinoma among others. These studies suggest FOXD2-AS1 as a potential biomarker in human cancers (Zhou et al., 2019).

The mechanisms by which FOXD2-AS1 influences cancer development are not fully elucidated. It may be involved in tumour progression through sponging of tumour-suppressive microRNAs (miR) (Zhu et al., 2018). It is also suggested that protein 42 (CDC42) could be a downstream molecule of FOXD2-AS1 in colorectal cancer (Shen et al., 2019). By interacting with several different miR, FOXD2-AS1 modulates proliferation and invasion of colorectal cancer, glioma and other cancers (Dong et al., 2019). It is known that lncRNAs influence the development of tumours by regulating gene expression at the transcriptional or post-transcriptional level (Zhang et al., 2020).

2. PROBLEM INTRODUCTION

According to the World Health Organization, cancer is the second leading cause of death, globally about 1 in 6 deaths is due to cancer, and colorectal cancer (CRC) is one of the more common and deadly cancers. It is the result of an accumulation of genetic and epigenetic changes in colon epithelial cells, which transform them into adenocarcinomas (Portela and Esteller, 2010).

Over the past decades, many studies concluded that nearly all colorectal cancers have aberrant epigenetics, especially aberrant DNA methylation (Das and Singal, 2004). The average colon cancer epigenome has hundreds to thousands of abnormally methylated genes. Accumulation of these alterations as well as genomic alterations allow for more recurrence and metastasis, which are the main cause of death in colorectal cancer patients. Clinical markers for diagnostic, prognostic and therapeutic applications are being developed using this signature alterations (Lao and Grady, 2011).

With this knowledge we wondered which genes could be aberrantly methylated in colorectal cancer. By studying DNA methylation in TCGA data, we observed a hypermethylation of FOXD2-AS1 (**figure 2**). Little was known of FOXD2-AS1 back then when this project started in 2018. In the recent years it has been studied in the context of cancer by different researchers, but no study focused in its relation with FOXD2, which is a gene located in the same locus and is also found hypermethylated in TCGA data (**figure 2**). In order to elucidate their potential associations and implications in CRC we performed a transcriptomic and epigenetic characterization of the region in a series of normal and tumour tissue from colorectal cancer patients.

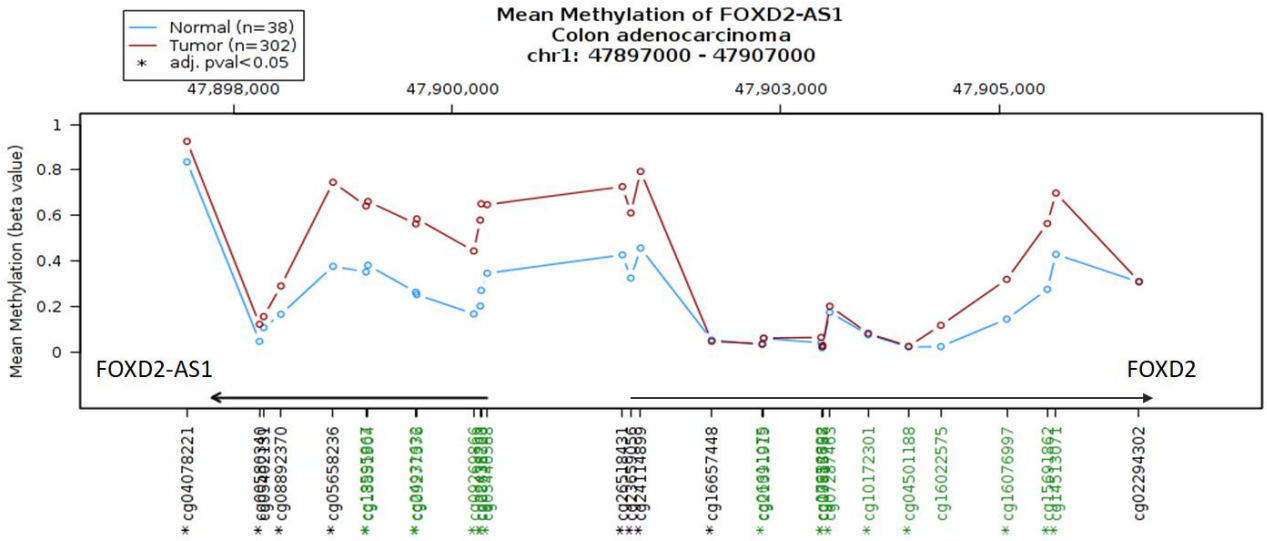


Figure 2: **Mean methylation representation of FOXD2 and FOXD2-AS1 in colon adenocarcinoma**. TCGA data visualized using Wanderer online web tool (Díez-Villanueva et al., 2015). T-test for each CpG shown was performed. * pv<0.05

3. OBJECTIVES

The aim of the project is to explore and characterize the methylation and expression of FOXD2 and FOXD2-AS1 genes in healthy and tumoral colorectal tissue samples and study their associations.

1. Characterization:
    a. Study FOXD2 and FOXD2-AS1 expression in paired healthy and tumoral colorectal samples.
    b. Characterize the DNA methylation profiles of the CpG islands associated to FOXD2 and FOXD2-AS1 in healthy and tumoral paired colorectal samples.
    c. Explore the correlation between the expression of FOXD2 and FOXD2-AS1 and the methylation profile of the associated CpG islands.
2. Clinical analysis
    a. Investigate if the expression and methylation profiles observed are associated with clinicopathological parameters of the patients.

4. MATERIALS AND METHODS

Colorectal tissue

The study includes 150 patients from the Hospital of Bellvitge. For each patient two independent samples were obtained, including colorectal cancer tissue and healthy distal mucosa nearby the tumour.

DNA extraction

Total DNA was extracted from colorectal samples using Genomic DNA Kits PureLink® (Invitrogen) according to manufacturer instructions. DNA was quantified by NanoDrop™ (ND-1000), and the integrity was assessed by 1% agarose gel electrophoresis.

RNA extraction

Total RNA was extracted from colorectal samples using PureLink® RNA Mini Kit (Invitrogen) according to manufacturer instructions with the Phase Separation step with Trizol® Reagent (Triagent- Ambion) and a purification step with On-column PureLink® DNase treatment and purification. RNA was quantified by NanoDrop™ (ND-1000), and the integrity was assessed by 1% agarose gel electrophoresis by visualizing the 28S and 18S bands.

DNA methylation analyses

Bisulfite conversion

Bisulfite conversion was performed using EZ DNA Methylation Kit (Zymo Research) according to the manufacturer instructions, with 300-400 ng of DNA in 50 µL elution volume for each sample. Thermocycler conditions were 15 minutes at 37ºC and 12-16 hours at 50º.

Nested PCR

Nested PCR was performed using Taq DNA Polymerase (Roche). The mix was prepared with 1,5 µL of buffer 10x (Roche), 1,5 µL of DNTPs 1mM (Thermo Scientific), 0,5 µL of forward primer 20 µM, 0,5 µL of reverse primer 20 µM, 0,1 µL of Taq. polymerase 5u/µL (Roche) and 7,4 µL of PCR water. 11,5 µL of the mix was pipetted into a 0,2mL tube and 1 µL of the bisulfited sample was added. For the region 2, a 1/10 dilution of the external PCR was performed. Thermocycler conditions were modified and tested as described in the **supplementary table 1**. PCR products were run in a 2% agarose gel in order to confirm an adequate amplification.

For region 3, the Taq. polymerase used was Zymotaq from Ecogen. Extern PCR conditions are specified in **supplementary table 2**. After the external PCR a 1/10 dilution was performed. Internal PCR was performed with the Lightcycler 384- qPCR machine using Sybergreen mastermix. Hybridization temperature was 56º and 30 cycles were performed. Primers are specified in **table 1**.

Two controls were introduced: A negative control with the mix and no sample, and a positive control with already analysed samples that we know they amplify.

Primers were designed in the laboratory using Methprimer and purchased in Invitrogen.

Gene	Region	PCR	Direction	Sequence
FOXD2-AS1	1	Extern	Fw	TTTTAAGGTTTTGGGTTAGTTT
FOXD2-AS1	1	Extern	Rv	CAACTCATTTATAAAAACCAAAA
FOXD2-AS1	1	Intern	Fw	GGGATTGGGAGAAGGGTTAT
FOXD2-AS1	1	Intern	Rv	ATAAAAAAACCCAACAAACATCC
FOXD2	2	Extern	Fw	AGATAGTTATAGAGATTGAG
FOXD2	2	Extern	Rv	CACCCTATACTCCCTAAA
FOXD2	2	Intern	Fw	ATTTTTTTTAGGTTTAAGGTTG
FOXD2	2	Intern	Rv	CCCTAAATATTAAAACTCACT
FOXD2	3	Extern	Fw	GTTTTGGTTATGTTGATTGT
FOXD2	3	Extern	Rv	TAAAACCTAACCAAACATCT
FOXD2	3	Intern	Fw	GGTTTTAGTGGTTGTTATTTT
FOXD2	3	Intern	Rv	CCCCTACTTTTATTTCTCAA
All primer sequences are presented in 5'->3' orientation.				

Table 1: **Primers needed for the Nested PCR**.

Exosap purification

An exosap purification of the PCR product was performed before sequencing using Thermo Fisher Scientific® enzymes. The mix was prepared with 0,5 µL of Exonuclease (10u/µL) and 1 µL of FastAP (1u/µL). 1,5 µL of the mix was pipetted into a 0,2 mL tube containing 5 µL of the PCR sample. Thermocycler conditions were 37º for 15 minutes, and 85º for 15 minutes.

Sanger sequencing

Purified samples were sent to Macrogen® in order to perform a Sanger sequencing using their Ez-seq service. 1 µL of the purified PCR sample was added to 7 µL of pure water and 2 µL of primer (20µM). Primers for Sanger sequencing were the same as used to amplify the internal PCR (**table 1**). For regions 2 and 3, the FOXD2 intern reverse primer was used, while for region 1, the FOXD2-AS1 intern forward primer was used.

The degree of methylation was analysed by comparing the peak height of cytosine residues with the peak of thymine residues in forward sequences and by comparing the peak of adenine and guanine residues in reverse sequences using the raw data (**supplementary figure 1**). Results were graphed using a methylation plotter (Mallona et al., 2014).

Expression analyses

RNA expression was assessed by quantitative reverse transcription PCR (qRT-PCR).

RT-PCR

500 ng of total RNA were converted to cDNA using the Superscript IV First-Strand Synthesis System (Thermofisher Scientific) according to manufacturer's guidelines.

Two mixes were prepared, the mix 1 with dNTPs (5mM) and 150 ng of pd(N) 6 hexamers and the mix 2 with SSIV buffer (5x), DTT (100mM) , RNase inhibitor (40 u/µL, Promega) and SuperScript IV RT (200U/µL retrotranscriptase) was prepared. Then 3 µL of mix 1 was pipetted into each tube and incubated in the thermocycler at 65ºC for 5 minutes, and in ice for 1 minute. After that, 7 µL of the mix 2 was added into each tube and was briefly centrifuged. SimpliAmp™ Thermal Cycler was programmed with 23º 10 min, 50-55º 10 min and 80º 10 min.

Two negative controls were introduced: one where RNA was not included and one where the reverse transcriptase was not included. After reverse transcription, samples were diluted to obtain 20 ng of DNA in 5 µL of volume.

qPCR

Quantitative PCR was performed using SYBR GREEN I (Roche) according to manufacturer's specifications. A LightCycler® 480 Real-Time PCR System with white Multiwell Plate 384 (Roche) was also used in the procedure.

Briefly, PCR mix was prepared with 5 µL of SYBR Green Master Mix, 0,5 µL of 20 µM primers (**table 2**) and 3µL of water. 1 µL of 1/5 diluted cDNA and 9 µL of mix were loaded into a 384 well plate. Each sample was analysed in triplicate and the expression levels were normalized according to the average of two independent reference genes PMSC4 and PUM1. Efficiencies were calculated using Chainy webtool (Mallona et al. 2017) and LightCycler® 480 SW software version 1.5.1 was used to calculate the relative expression.

Primers	Sequence primer FW	Sequence primer RV
FOXD2 FW	CTGACGTTGAGCGAGATCTG	GGGATCTTGACGAAGCAGTC
FOXD2-AS1	CGTGTAACCCTTCTGAGTCC	CCCTGGCTTTGCTTCTATGAG
PMSC4	TGTTGGCAAAGGCGGTGGCA	TCTCTTGGTGGCGATGGCAT
PUM1	CGGTCGTCCTGAGGATAAAA	CGTACGTGAGGCGTGAGTAA
For all primers the melting temperature was 62º.		

Table 2: Primers used during qPCR procedures. Primers are presented in a 5' -> 3' direction.

Statistics

Data was analysed and graphed using GraphPad Prism version 9.0.0. For all data analysed normality tests were applied. Paired experimental data sets were compared by two-tailed Wilcoxon signed rank test. Unpaired experimental data sets were compared by two-tailed Mann-Whitney tests comparing ranks. For multiple comparisons, a Kruskal-Wallis test and Dunn's multiple comparisons test were performed. To analyse survival curves, a Mantel-Cox test with Gehan-Breslow-Wilcoxon test were performed. For ROC curves Wilson/Brown method was used to graph the results. Values of * $p<0.05$, ** $p<0.01$, *** $p<0.001$ and **** $p<0.0001$ were considered statistically significant.

Methylation levels at each site were transformed to M values following this formula: M value=$\log_2$(Beta value/1-Beta value) to perform statistical analysis. Methylation results were graphed using beta values, expressing the estimated methylation levels using the ratio of methylated and unmethylated alleles.

Heatmaps were performed using ComplexHeatmap package in R Studio (Gu et al., 2016). Clusters were calculated using Ward.D2 method.

5. RESULTS

As mentioned above, FOXD2 is located on the positive strand of chromosome 1, while FOXD2-AS1 is located in the same locus but in the opposite direction, being on head to head disposition. As shown in **figure 3**, there are several CpG islands along the region containing the FOXD2 and FOXD2-AS1 genes. In order to explore their methylation, three independent regions have been determined, named regions 1, 2, and 3. Region 1 is located at the 5' of FOXD2-AS1 (according to RefSeq annotation) and is 525 bp. Region 2 is located on one side of the CpG island between and near the TSS of both genes (according to FANTOM CAT annotation) and is 495 bp long. Finally, region 3 is located at 3' of FOXD2 and includes the last 479 bp of the CpG Island.

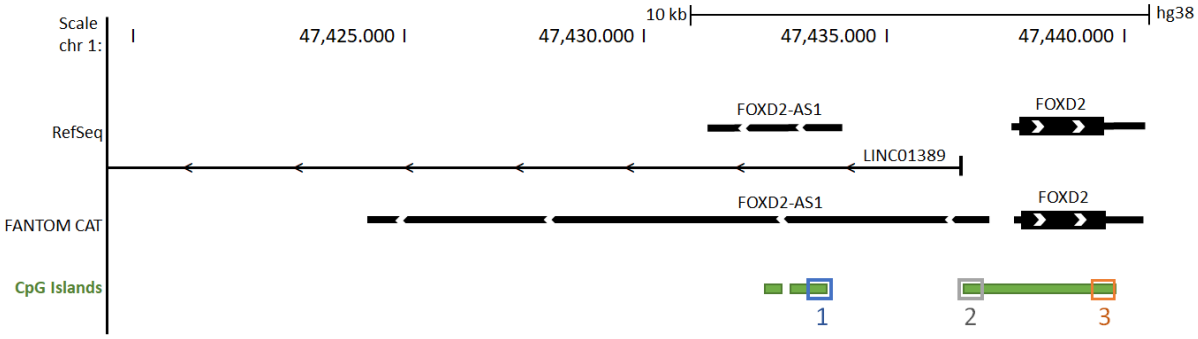


Figure 3: **Representation of FOXD2 and FOXD-AS1 locus in chromosome 1** with two independent gene annotations (FantomCat and RefSeq) as well as the CpG islands. Three regions experimentally determined are highlighted in blue, grey and orange squares. Image extracted from: UCSC Genome browser.

Tumours were found hypermethylated in FOXD2 and FOXD2-AS1 locus in comparison to normal samples.

First, methylation of FOXD2 and FOXD2-AS1 locus was studied by bisulfite conversion and Sanger sequencing in 50 paired normal-tumour samples. A significant hypermethylation of regions 1 and 3 was found in tumour samples when compared to normal samples ($p<0.0001$) (**figure 4 A-B**). Mean methylation of each region was also found significantly higher in tumour samples when compared to normal samples ($p<0.0001$) (**figure 4 C-D**).

On the other hand, region 2 was found completely unmethylated in all tumour and normal samples (data not shown).

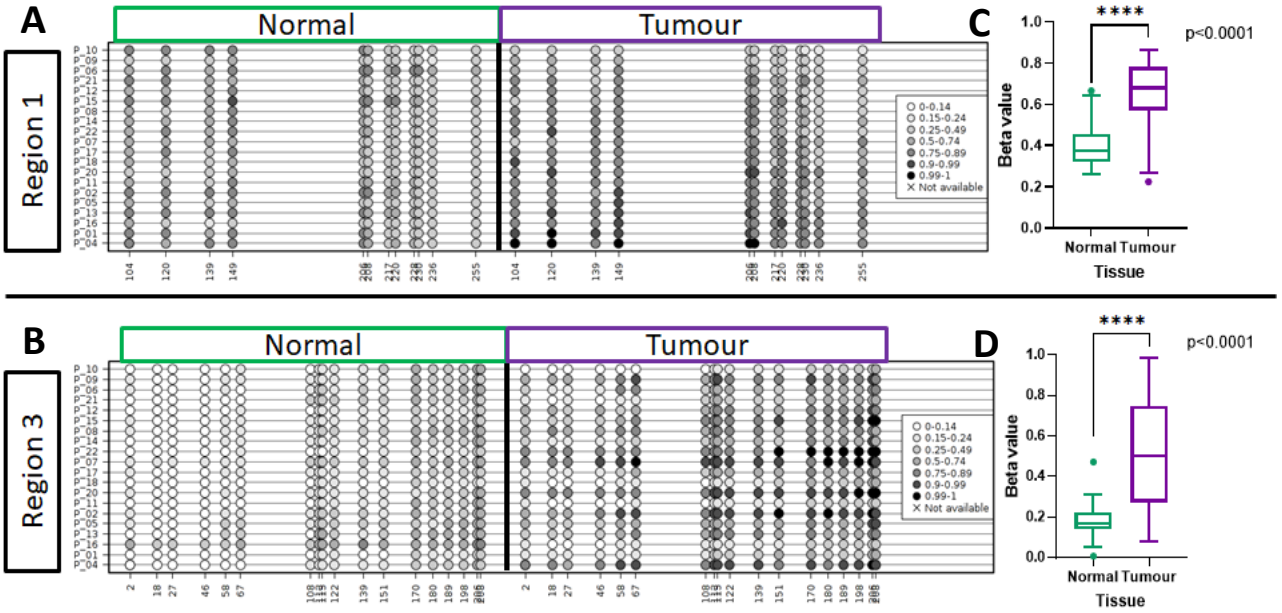


Figure 4**: FOXD2 and FOXD2-AS1 CpG methylation.** A) and B) Methylation plots where each dot represents a CpG methylation fraction (the darker the more methylated). The X axis determines the position of a given CpG and each line represents a different patient. C) and D) boxplots represent the comparison of mean methylation of regions 1 (n=49) and 3 (n=48) between normal and tumour tissues. Wilcoxon matched pairs signed rank test. **** p<0.0001

Following the methylation differential analysis, we explored the correlation of the methylation between all CpGs of the region of study in 48 patients (96 paired normal-tumour samples). A correlation matrix of all the individual CpGs comprised in regions 1 and 3 was performed, and a general significant positive correlation was found between regions with some exceptions (**Figure 5 B**). As expected, CpGs within the same region have a strong positive correlation, with a Spearman r value between 0.52 and 0.95. However, CpGs within regions 1 and 3, which are more than 3kb apart, have a mild positive correlation, with a Spearman r between 0.21 and 0.51 (**Figure 5 A**).

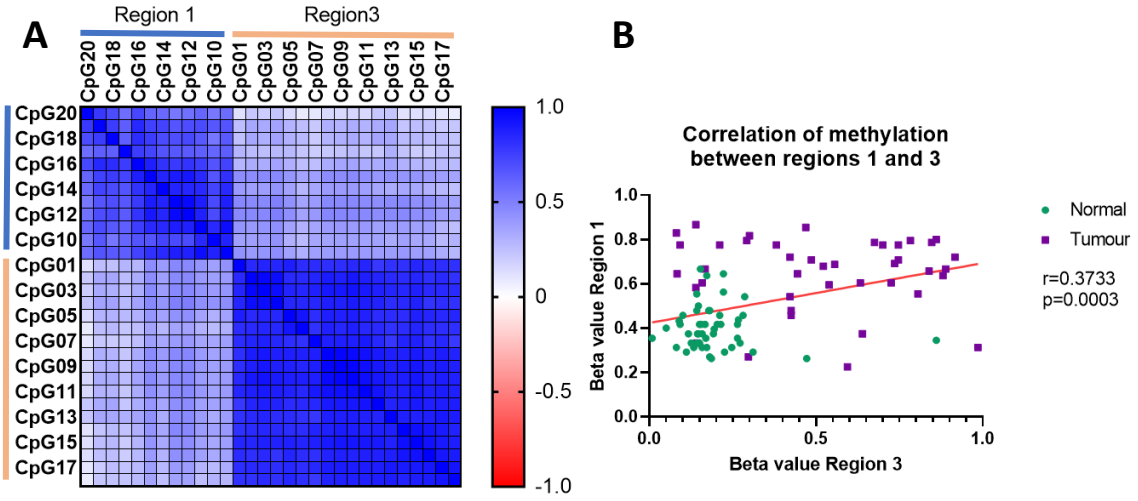


Figure 5: **Correlation between CpGs methylation in normal and tumoral colorectal tissue**. a) Correlation matrix between single CpGs of regions 1 (n=49) and 3 (n=48) in both normal and tumoral colorectal tissue. Methylation values were transformed to M values and analysed using a Spearman correlation test. b) Mean methylation CpG values correlation between regions 1 (n=49) and 3 (n=48) for both normal and tumoral colorectal tissue was studied by Spearman correlation test.

Normal and tumour samples can be differentiated by studying their DNA methylation levels.

In addition, a heatmap study was performed in order to clarify if methylation status could differentiate between normal and tumoral samples. Three patient clusters were determined studying the CpG methylation of both regions (**figure 6**). The first cluster is composed of 17 tumoral samples (100% tumoral samples), the second cluster is composed of 38 normal samples and 1 tumoral sample (97% normal samples and 3% tumoral samples) and the third cluster is composed of 25 tumoral samples and 6 normal ones (81% tumoral and 19% normal samples). Overall, the clustering was able to classify the samples correctly with a sensitivity of 97% and a specificity of 86%.

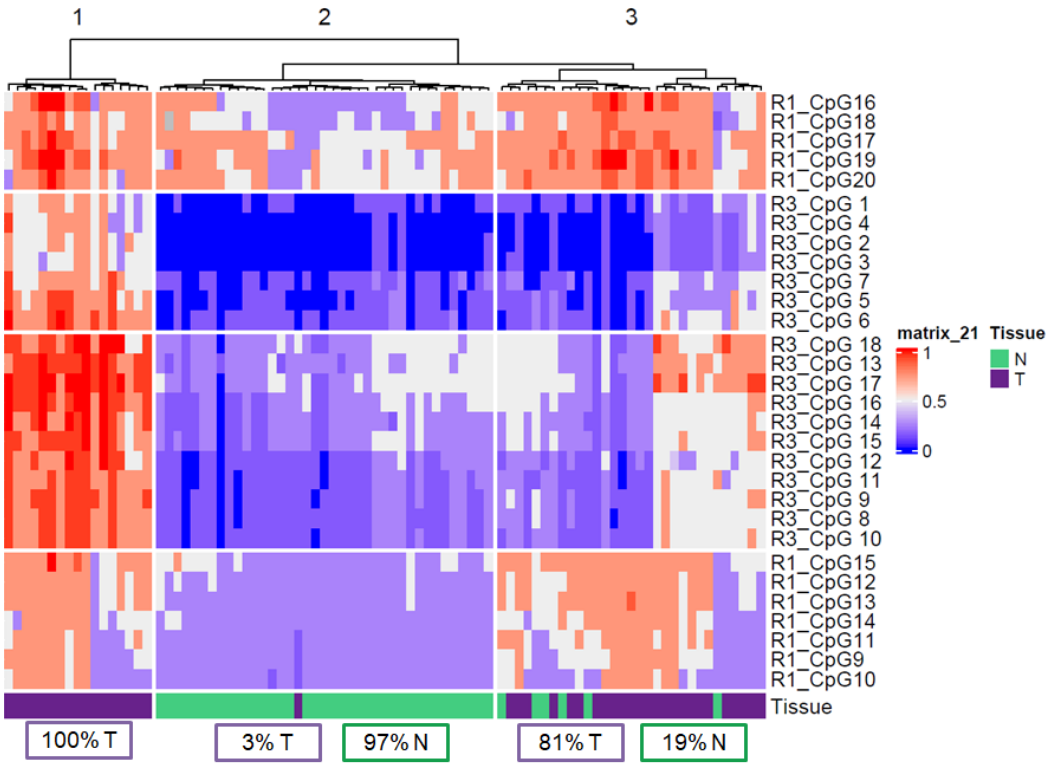


Figure 6: **Heatmap methylation with sample and CpG clustering**. Methylation of regions 1 (n=49) and 3 (n=48) were determined by bisulfite conversion and Sanger sequencing in normal and tumoral colorectal tissue. Clusters were calculated using Ward.D2 method in ComplexHeatmap R program.

Methylation levels of tumoral samples were not significantly associated with clinicopathological features of the tumours in our patient cohort.

Being observed the methylation profiles, we wondered if the methylation levels were associated with clinicopathological features of the tumours. Therefore, correlation between tumoral methylation levels and clinical data of patients was studied using the average methylation of

regions 1 and 3. No significant correlation was found between the methylation levels and clinicopathological features (cancer stage, lymph node involvement nor relapse). However, higher methylation levels of region 1 were significantly associated with bigger tumour sizes, in contrast, lower methylation levels of region 3 were associated with biopsies from the distal colon (**Table 3**).

		Region 3			Region 1		
Variables	**Cathegories**	**n**	**mean ± SD**	**p value**	**n**	**mean ± SD**	**p value**
Cancer Stage	Stage II	25	0.50 ± 0.28	0.8650	28	0.65± 0.17	0.9083
	Stage III	19	0.51 ± 0.26		21	0.65 ± 0.15	
Tumor size and invasion	T1 and T2	1	0.13 ± 0.0	0.3090	1	0.22 ± 0.00	0.003
	T3	32	0.50 ± 0.28		36	0.68 ± 0.13	*
	T4	11	0.56 ± 0.24		12	0.58 ± 0.17	
Lymph node involvement	N0	25	0.50 ± 0.28	0.8361	28	0.65 ± 0.17	0.5022
	N1	13	0.54 ± 0.27		14	0.67 ± 0.13	
	N2	6	0.46 ± 0.26		7	0.59 ± 0.16	
Relapse	Yes	28	0.48 ± 0.28	0.6943	10	0.62 ± 0.17	0.272
	No	90	0.51 ± 0.27		95	0.54 ± 0.32	
Tumour site	Proximal	25	0.60 ± 0.26	0.0112	26	0.62 ± 0.17	0.2884
	Distal	19	0.38 ± 0.25	*	23	0.68 ± 0.13	

Table 3: **Correlation between region 1 and region 3 methylation and the clinical pathological parameters in tumoral colorectal tissue.** Associations between tumour average methylation of regions 1 (n=49) and 3 (n=48) and patient clinicopathological data. Data was analysed using Mann-Whitney test or Krustal-Wallis test with Dunn's multiple comparisons test.

FOXD2 and FOXD2-AS1 expression was found decreased in tumoral samples in comparison to normal samples.

In addition, FOXD2 and FOXD2-AS1 expression was studied in 108 patients (216 paired normal-tumour samples). A significant decrease in both FOXD2 and FOXD2-AS1 expression was found in tumoral samples compared to normal samples (p<0.0001) (**Figure 6 A**). In addition, the ratio between FOXD2-AS1 and FOXD2 expression was found to be significantly different in normal and tumoral samples (p<0.0001), being higher in tumours (**Figure 6 B**). Specifically, the average expression ratio in normal samples was 0.59 ± 0.26 in contrast to 1.17 ± 1.24 in tumoral samples.

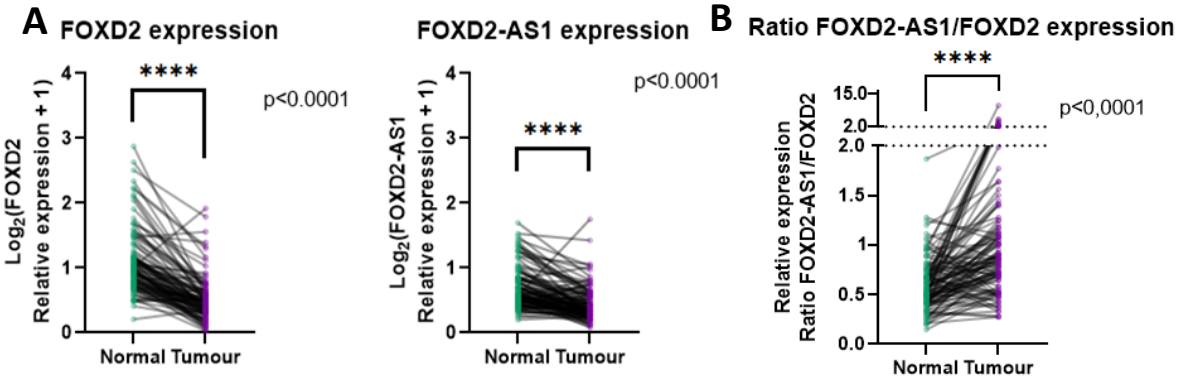


Figure 7: **FOXD2 and FOXD2-AS1 expression determined by qPCR in normal and tumoral colorectal cancer tissue in 108 patients.** qPCR data was normalized using PMSC4 and PUM1 expression and was analysed using Wilcoxon matched pairs signed rank test. **** p< 0.0001

As the transcription start site (TSS) of FOXD2 and FOXD2-AS1 are in close proximity (according to FantomCAT annotation), we wondered if they coexpressed. Interestingly, a significant positive correlation between the expression of FOXD2 and FOXD2-AS1 was found when studying both normal and tumoral samples (p<0.0001) (**Figure 8**).

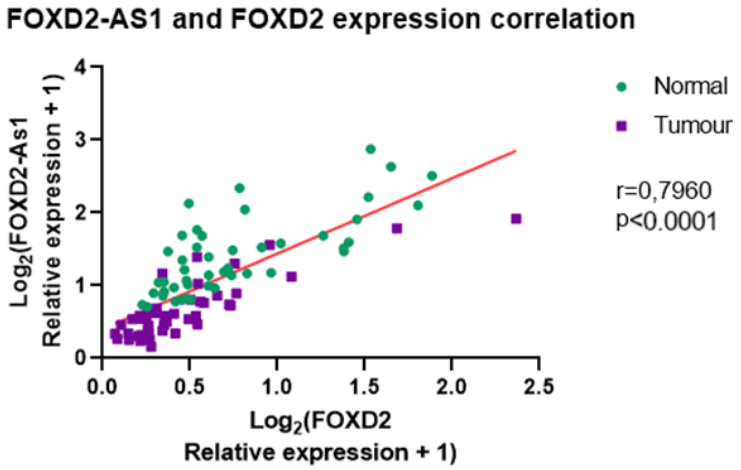


Figure 8: **FOXD2 and FOXD2-AS1 expression correlation determined by qPCR in normal and tumoral colorectal cancer tissue in 108 patients.** qPCR data was normalized using PMSC4 and PUM1 expression and was analysed applying Spearman correlation. (p< 0.0001).

Samples can be differentiated between normal and tumoral by studying their FOXD2 and FOXD2-AS1 expression levels.

After characterizing FOXD2 and FOXD2-AS1 expression between paired tissues, we wondered if the samples could be classified between normal and tumoral by determining their gene expression. We performed ROC curves to stablish cut points for FOXD2 and FOXD2-AS1 expression and the ratio FOXD2-AS1/FOXD2 based on sensibility and specificity. All three curves were significative (p<0.0001) and with good areas under de curve, especially for FOXD2 (A=0.87) (**figure 9**). Thus, FOXD2 would be the best candidate to differentiate between normal and tumoral samples, setting the cut point at 0,5433 with a sensitivity of 76,85% and specificity of 89,81%.

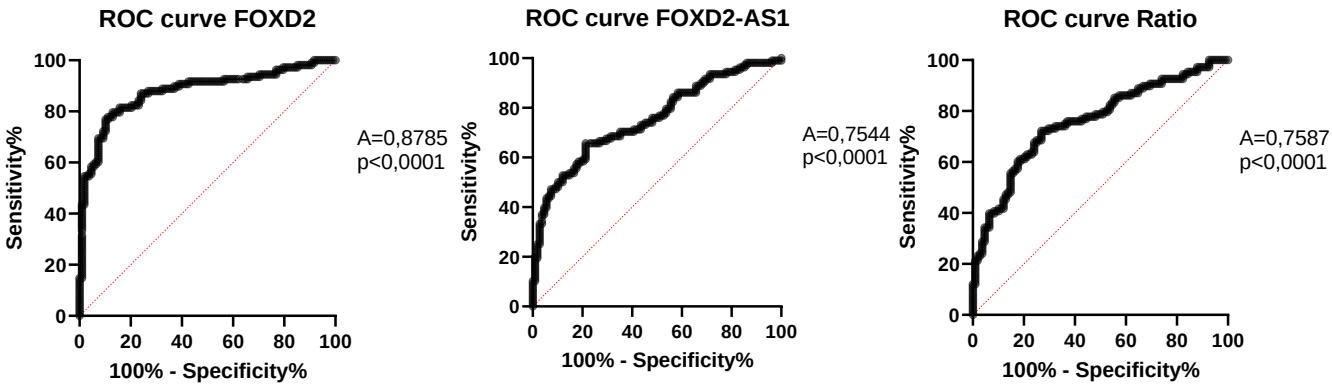


Figure 9: **ROC curves of FOXD2 and FOXD2-AS1 expression and their ratio.** Gene expression was determined by qRT-PCR and data was normalized using PMSC4 and PUM1 expression and was analysed utilizing ROC curves analysis with Wilson/Brown test ($p<0.0001$).

Expression levels of FOXD2 and FOXD2-AS1 in tumour samples were not associated with clinicopathological features of the tumours in our patient cohort.

Similar to the studies performed regarding methylation, we studied if the expression levels of FOXD2 and FOXD2-AS1 were associated with clinicopathological features of the tumours. Unlike methylation, no significant association between gene expression and clinicopathological data from patients was found (**table 4**).

			FOXD2		FOXD2-AS1		FOXD2-AS1/FOXD2	
Variables	**Cathegories**	**n**	**mean ± SD**	**p value**	**mean ± SD**	**p value**	**mean ± SD**	**p value**
Cancer Stage	Stage II	60	0.52 ± 0.53	0.4815	0.43± 0.38	0.2798	1.01 ± 0.82	0.7372
	Stage III	48	0.40 ± 0.29		0.33 ± 0.20		1,27 ± 1,62	
Tumor size and invasion	T1 and T2	5	0.38 ± 0.10	0.3831	0.31 ± 0.16	0.911	0.81 ± 0.35	0.2466
	T3	78	0.51 ± 0.50		0.40 ± 0.34		1.16 ± 1.34	
	T4	25	0.34 ± 0.21		0.36 ± 0.22		1.30 ± 0.98	
Lymph node involvement	N0	60	0.52 ± 0.53	0.3925	0.43 ± 0.38	0.4852	1.10 ± 0.82	0.9109
	N1	29	0.43 ± 0.26		0.34 ± 0.17		1.24 ± 1.83	
	N2	19	0.37 ± 0.34		0.33 ± 0.25		1.31 ± 1.26	
Relapse	Yes	28	0.34 ± 0.18	0.4019	0.30 ± 0.16	0.272	1.22 ± 1.10	0.8992
	No	90	0.49 ± 0.47		0.40 ± 0.33		1.16 ± 1.27	
Tumour site	Proximal	47	0.48 ± 0.56	0.1643	0.38 ± 0.41	0.1112	1.34 ± 1.63	0.537
	Distal	61	0.46 ± 0.32		0.39 ± 0.22		1.05 ± 0.81	

Table 4: **Correlation between FOXD2 and FOXD2-AS1 expression and the clinical pathological parameters**. FOXD2 and FOXD2-AS1 expression was determined by qPCR in tumoral colorectal cancer tissue in 108 patients. qPCR data was normalized using PMSC4 and PUM1 expression and was analysed using Mann-Whitney test or Krustal-Wallis test with Dunn's multiple comparisons test.

The possible prognostic function of these genes was also determined by studying the correlation between FOXD2 and FOXD2-AS1 expression and overall survival. First, patients were classified based on their survival, whether they were dead (n=11) or alive (n=97). Secondly, several ROC curves were performed to establish an expression cut point based on sensibility and specificity, but a low area and no significant p value were obtained. However, we

continued the study with the best cut point, and performed three survival curves based on FOXD2 expression, FOXD2-AS1 expression and the ratio between the two genes (data not shown). No significant differences in overall survival could be found in our patient cohort in relation to gene expression. In addition, disease free survival was studied using the same method, but no correlation was found either.

Correlation between gene expression and methylation was found significant.

Finally, in order to determine if the levels of methylation were related to gene expression, correlation between average methylation levels of regions 1 and 3 and FOXD2 and FOXD2-AS1 expression levels were studied in 47 patients. Interestingly, the correlation analysis results revealed that methylation of regions 1 and 3 was negative correlated with FOXD2 and FOXD2-AS1 (**figure 10**).

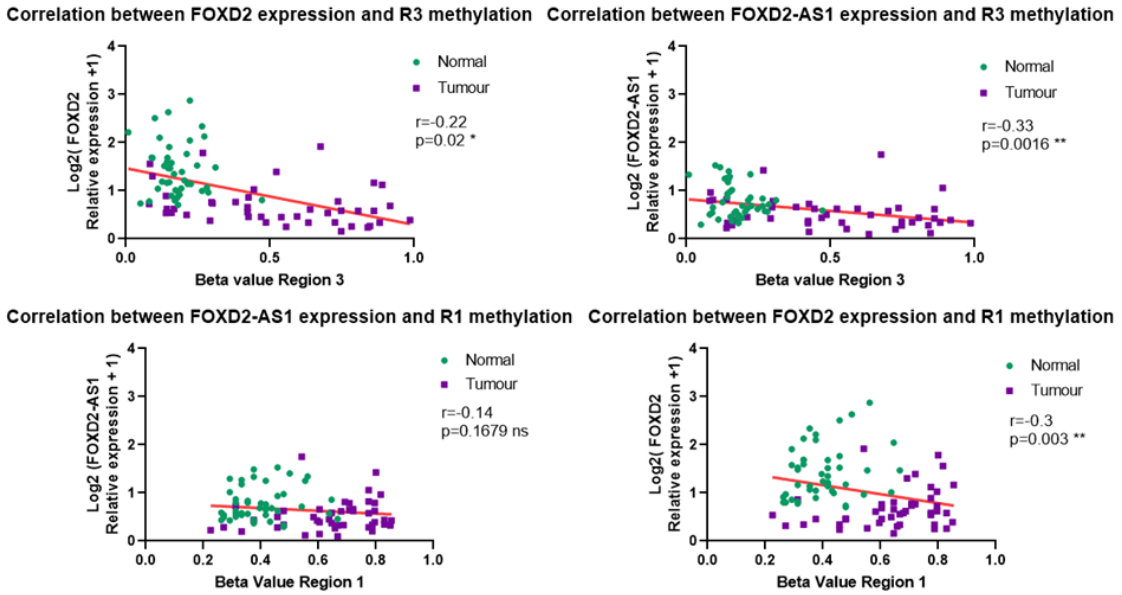


Figure 10: **Correlation between FOXD2 and FOXD2-AS1 expression and Region 1 and Region 3 methylation**. Expression of FOXD2 and FOXD2-AS1 of was determined by qPCR in normal and tumoral colorectal tissue of 47 patients. qPCR data was normalized using PMSC4 and PUM1 expression. Methylation of both FOXD2 CpG regions was assessed by bisulfite conversion and Sanger sequencing. Data was analysed using Spearman correlation.

6. DISCUSSION

Cancer cells acquire some biological capabilities during neoplastic transformation, known as the hallmarks of cancer. Some examples of them are a sustained proliferative signalling, activated invasion and metastasis, genome instability and mutations, and cellular energetics deregulation (Hanahan & Weinberg, 2011). Classical genetics, epigenetics and epitranscriptomics have been studied in the context of cancer, and their alteration contributes to the acquisition of these hallmarks. lncRNAs are extremely important for proper cell function, as they engage in numerous biological processes across every branch of life such as regulating gene expression at both transcriptional and post-transcriptional levels. For many years, little was known about ncRNAs, only in the recent years with the advance of technology they are starting to be characterized, so their regulatory functions are not usually well established. However, recent studies have shown that dysregulation of lncRNAs is usually involved in proliferation, metastasis, invasiveness of cancer cells, tumour growth and angiogenesis (Sun et al., 2017). Some data that supports this idea are the findings of the interplay between lncRNAs and microRNAs in normal and cancer cells, as well as the importance of these interactions during the tumorigenic process (Liz and Esteller, 2015).

Epigenetic alterations in human CRC have been recognized and investigated for a long time. Abnormalities in epigenetics, including DNA methylation, histone acetylation and methylation, non-coding RNAs and chromatin remodelling, have been reported in every aspect of tumour development. In our study, we focused on lncRNA FOXD2-AS1, which has been found to be aberrantly expressed in some cancers, as well as FOXD2, the coding protein gene located in the same region in the opposite strand. The lncRNA known as FOXD2-AS1 is a novel cancer-related lncRNA (Zhang et al., 2020). Its overexpression has been observed in several cancers such as: nasopharyngeal carcinoma, hepatocellular carcinoma, gastric cancer, colorectal cancer, non-small lung cancer, etc. Specifically, in colorectal cancer it was found upregulated affecting functions such as proliferation, migration and invasion (Hu et al., 2019). On the other hand, very little is known about FOXD2 and its implications in human disease, especially with cancer.

While Yang (2017) found that FOXD2-AS1 expression was significantly increased in CRC tissues and cell lines suggesting an oncogene role, we found that FOXD2-AS1 expression was decreased in tumoral samples in comparison to normal samples. In their study they concluded that down-regulation of FOXD2-AS1 suppressed cell proliferation, migration and invasion in vitro (Yang et al., 2017). In addition, Zhu (2018) also reported that expression of FOXD2-AS1 was increased in tumoral samples, and its upregulation promoted proliferation, migration, and invasion in colon cancer cells.

Colorectal cancer patients usually exhibit a survival rate of less than 5 years due to early metastasis. Although treatment methods (such as surgery, radiation therapy, chemotherapy and targeted therapy) have advanced, the high recurrence and unfavourable prognosis remain

the disturbing problems. Consequently, it is vital to research the molecular mechanisms of CRC occurrence and development for exploring new therapeutic targets for CRC patients. Regarding the survival of the patients, Zhu (2018) performed a 12 years follow-up study that reported that high expression levels of FOXD2-AS1 in CRC were associated with shorter survival rates (Zhu et al., 2018). We did not observe any correlation between FOXD2-AS1 expression and survival rates in our patient cohort, this is probably due to the small and uniform patient cohort we studied, where only 11 patients of the 110 patients studied had died.

This upregulation of FOXD2-AS1 and its correlation to poor survival has also been observed in other types of cancer such as: gastric cancer, non-small lung cancer, bladder cancer, nasopharyngeal carcinoma, oesophageal cancer, hepatocellular cancer, thyroid cancer and cutaneous melanoma. In most of these cancers FOXD2-AS1 plays a role in proliferation and is associated with advanced TNM stage and shorter overall survival. In cutaneous melanoma it was also related to distant metastasis (Hu et al., 2019). One possible reason for the discordance of our results with the previous ones could be the limitations in number and variety of our sample cohort. More studies with larger sample sizes are required. Additionally, the majority of studies regarding FOXD2-AS1 were performed in Asia, while our study was performed in Europe with samples obtained from patients of the Hospital of Bellvitge located in Barcelona. Genetic and epigenetic landscape of individuals can vary within countries. The techniques used for the study may vary as well, possibly affecting the results.

On the other hand, very little is known about FOXD2 expression, and only three papers published study FOXD2 in the context of cancer. Ernstsson (1997) determined that FOXD2 was only expressed in kidney tissue using a hybridization panel. We analysed FOXD2 expression in colorectal tissue by qRT-PCR and found that FOXD2 was also expressed there. In addition, we determined that FOXD2 expression was significantly decreased in tumoral samples when compared to tumoral samples. Sulman (2004) discovered a deletion on some meningiomas that included the FOXD2 locus, this fact together with our finding suggests that FOXD2 may play a tumour suppressor gene role. But contrarily, van der Heul-Nieuwenhuijsen (2009) found an overexpression of FOXD2 in prostate cancer and lymph node metastases. Moreover, we observed a significative and strong correlation between FOXD2 and FOXD2-AS1 expression. We speculate that, because they are found in the same locus in opposite strands, their expression may be regulated through the same mechanism; this phenomenon has been described before in other genes (Fraser and Bickmore, 2007). Further functional studies are needed to determine the transcription mechanisms that regulate the expression of FOXD2 and FOXD2-AS1.

On the other hand, FOXD2-AS1 and FOXD2 locus methylation has not been well characterized. Our studies show that the locus is found hypermethylated in tumoral samples. We observed a mild positive correlation between the different CpGs within the region, with a stronger correlation between closer CpGs. With the heatmap study, we determined that methylation could serve as a marker to distinguish between tumoral and normal samples. Additionally, we

determined a negative correlation between region 3 methylation and FOXD2-AS1 expression, but we couldn't find any significant correlation between region 1 and FOXD2-AS1 expression. Li (2019) also observed that the 15 CpGs sites they studied within the FOXD2-AS1 region were all hypermethylated in colorectal cancer (Li et al., 2019). But contrarily, they observed an increased expression of FOXD2-AS1, indicating a positive correlation between methylation and expression. Their findings suggest that FOXD2-AS1 may be upregulated by hypermethylation, while our findings suggest the opposite. Moreover, we observed a negative correlation between regions 1 and 3 methylation and FOXD2 expression. This is a common phenomenon, as usually hypermethylated genes are silenced because of the inability of the transcription machinery to bind to the promoter region (Herman and Baylin, 2003). Zamora (2015) also studied the 3'UTR region of FOXD2 (region 3 in our study) in colorectal cancer and adenocarcinoma and found out that methylation and FOXD2 expression were negatively correlated; these findings concur with our results. A differential methylation between normal and tumoral tissues is commonly observed amongst all cancers, and it is known to have important roles in tumour progression (Robertson, 2015). As we found an association between region 1 hypermethylation and higher tumour sizes, we speculate that targeting DNA methylation may be a possible treatment for colorectal cancer. Some treatments with DNA inhibitors for cancer therapy are already approved for clinical use in some countries (Issa, 2007). Survival of the patients in relation to methylation levels was not analysed in our study because of the limited sample size (44 patients alive and 6 dead).

In conclusion, although the overexpression of FOXD2-AS1 and its link to poor survival is generally accepted, our studies suggest that FOXD2-AS1 is decreased in tumoral samples and is not significantly linked to poor survival. Further studies are needed to determine if FOXD2-AS1 might function as a potential tumour biomarker for diagnosis and prognosis. Research on FOXD2-AS1 is just getting started; further specific functions and mechanisms of FOXD2-AS1 should be deeply explored in order to fully elucidate its clinical application.

7. CONCLUSIONS

- Regions 1 and 3 get hypermethylated in colorectal cancer tissues compared to normal, while region 2 remains completely unmethylated in all samples.
- Methylation levels of regions 1 and 3 could potentially be used to differentiate between normal and colorectal tumour samples.
- FOXD2 and FOXD2-AS1 are co-expressed and downregulated in tumours compared to normal.
- Expression levels of FOXD2 and FOXD2-AS1 could potentially be used to differentiate between normal and tumoral colorectal samples.
- Neither DNA methylation nor expression levels of FOXD2 and FOXD2-AS1 are strongly associated with clinicopathological features of the tumours. Only higher levels of region 1 methylation were associated to bigger tumour sizes.
- Higher levels of methylation in regions 1 and 3 were correlated with a decreased FOXD2 and FOXD2-AS1 expression.

8. BIBLIOGRAPHY

Benayoun, B. A., Caburet, S., & Veitia, R. A. (2011). Forkhead transcription factors: key players in health and disease. *Trends in Genetics*, *27*(6), 224-232.

Chen, L. L. (2016). Linking long noncoding RNA localization and function. *Trends in biochemical sciences*, *41*(9), 761-772.

Chen, Z., Li, S., Subramaniam, S., Shyy, J. Y. J., & Chien, S. (2017). Epigenetic regulation: a new frontier for biomedical engineers. Annual review of biomedical engineering, 19, 195-219.

Chiu, C. P., & Harley, C. B. (1997). Replicative senescence and cell immortality: the role of telomeres and telomerase. Proceedings of the Society for Experimental Biology and Medicine, 214(2), 99-106.

Conesa-Zamora, P., García-Solano, J., del Carmen Turpin, M., Sebastián-León, P., Torres-Moreno, D., Estrada, E., & Conesa, A. (2015). Methylome profiling reveals functions and genes which are differentially methylated in serrated compared to conventional colorectal carcinoma. *Clinical epigenetics*, *7*(1), 1-14.

Das, P. M., & Singal, R. (2004). DNA methylation and cancer. Journal of clinical oncology, 22(22), 4632-4642.

Díez-Villanueva, A., Mallona, I., & Peinado, M. A. (2015). Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer. *Epigenetics & chromatin*, *8*(1), 1-8.

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., & Xue, C. (2012). Landscape of transcription in human cells. Nature, 489(7414), 101-108.

Dong, H., Cao, W., & Xue, J. (2019). Long noncoding FOXD2-AS1 is activated by CREB1 and promotes cell proliferation and metastasis in glioma by sponging miR-185 through targeting AKT1. *Biochemical and biophysical research communications*, *508*(4), 1074-1081.

Ernstsson, S., Betz, R., Lagercrantz, S., Larsson, C., Ericksson, S., Cederberg, A., & Enerbäck, S. (1997). Cloning and Characterization offreac-9 (FKHL171), a Novel Kidney-Expressed Human Forkhead Gene That Maps to Chromosome 1p32–p34. *Genomics*, *46*(1), 78-85.

Feinberg, A. P., & Tycko, B. (2004). The history of cancer epigenetics. Nature Reviews Cancer, 4(2), 143-153.

Fraser, P., & Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature*, *447*(7143), 413-417.

Goelz, S. E., Vogelstein, B., & Feinberg, A. P. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. Science, 228(4696), 187-190.

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847-2849.

Hanahan, D., & Folkman, J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. cell, 86(3), 353-364.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. cell, 100(1), 57-70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. cell, 144(5), 646-674.

Herman, J. G., & Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. *New England Journal of Medicine*, *349*(21), 2042-2054.

Hu, Q., Tai, S., & Wang, J. (2019). Oncogenicity of lncRNA FOXD2-AS1 and its molecular mechanisms in human cancers. *Pathology-Research and Practice*, *215*(5), 843-848.

Issa, J. P. J. (2007). DNA methylation as a therapeutic target in cancer. Clinical Cancer Research, 13(6), 1634-1637.

Johansson, C. C., Dahle, M. K., Blomqvist, S. R., Grønning, L. M., Aandahl, E. M., Enerbäck, S., & Taskén, K. (2003). A winged helix forkhead (FOXD2) tunes sensitivity to cAMP in T lymphocytes through regulation of cAMP-dependent protein kinase RIα. *Journal of Biological Chemistry*, *278*(19), 17573-17579.

Katoh, M., Igarashi, M., Fukuda, H., Nakagama, H., & Katoh, M. (2013). Cancer genetics and genomics of human FOX family genes. *Cancer letters*, *328*(2), 198-206.

Lao, V. V., & Grady, W. M. (2011). Epigenetics and colorectal cancer. Nature reviews Gastroenterology & hepatology, 8(12), 686-700.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C. T., ... & Wei, C. L. (2010). Dynamic changes in the human methylome during differentiation. Genome research, 20(3), 320-331.

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature, 366(6453), 362-365.

Li, Z., Tan, H., Zhao, W., Xu, Y., Zhang, Z., Wang, M., & Zhou, X. (2019). Integrative analysis of DNA methylation and gene expression profiles identifies MIR4435-2HG as an oncogenic lncRNA for glioma progression. *Gene*, *715*, 144012.

Liz, J., & Esteller, M. (2016). lncRNAs and microRNAs with a role in cancer development. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1859*(1), 169-176.

Llinas-Arias, P., & Esteller, M. (2017). Epigenetic inactivation of tumour suppressor coding and non-coding genes in human cancer: an update. Open biology, 7(9), 170152.

Lowe, S. W., & Lin, A. W. (2000). Apoptosis in cancer. Carcinogenesis, 21(3), 485-495.

Mallona, I., Díez-Villanueva, A., & Peinado, M. A. (2014). Methylation plotter: a web tool for dynamic visualization of DNA methylation data. *Source code for biology and medicine*, *9*(1), 1-5.

Mallona, I., Díez-Villanueva, A., Martín, B., & Peinado, M. A. (2017). Chainy: an universal tool for standardized relative quantification in real-time PCR. *Bioinformatics*, *33*(9), 1411-1413.

Mattick, J. S., & Makunin, I. V. (2006). Non-coding RNA. Human molecular genetics, 15(suppl_1), R17-R29.

Myatt, S. S., & Lam, E. W. F. (2007). The emerging roles of forkhead box (Fox) proteins in cancer. *Nature Reviews Cancer*, *7*(11), 847-859.

National Cancer Institute. (2015). Cancer Staging. Retrieved from: https://www.cancer.gov/about-cancer/diagnosis-staging/staging. Retrieval date: June, 2020.

National Cancer Institute. (2015). What is cancer? Retrieved from: https://www.cancer.gov/about-cancer/understanding/what-is-cancer. Retrieval date: June, 2020.

Negrini, S., Gorgoulis, V. G., & Halazonetis, T. D. (2010). Genomic instability—an evolving hallmark of cancer. Nature reviews Molecular cell biology, 11(3), 220-228.

Odenthal, J., & Nüsslein-Volhard, C. (1998). Fork head domain genes in zebrafish. *Development genes and evolution*, *208*(5), 245-258.

Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. Nature biotechnology, 28(10), 1057.

Quinn, J. J., & Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nature Reviews Genetics*, *17*(1), 47.

Robertson, K. D. (2005). DNA methylation and human disease. Nature Reviews Genetics, 6(8), 597-610.

Schmalhofer, O., Brabletz, S., & Brabletz, T. (2009). E-cadherin, β-catenin, and ZEB1 in malignant progression of cancer. Cancer and Metastasis Reviews, 28(1-2), 151-166.

Shen, F., Chang, H., Gao, G., Zhang, B., Li, X., & Jin, B. (2019). Long noncoding RNA FOXD2-AS1 promotes glioma malignancy and tumorigenesis via targeting miR-185-5p/CCND2 axis. *Journal of cellular biochemistry*, *120*(6), 9324-9336.

Sulman, E. P., White, P. S., & Brodeur, G. M. (2004). Genomic annotation of the meningioma tumor suppressor locus on chromosome 1p34. *Oncogene*, *23*(4), 1014-1020.

Sun, L. Y., Li, X. J., Sun, Y. M., Huang, W., Fang, K., Han, C., ... & Wang, W. T. (2018). LncRNA ANRIL regulates AML development through modulating the glucose metabolism pathway of AdipoR1/AMPK/SIRT1. *Molecular cancer*, *17*(1), 1-6.

Sun, W., Yang, Y., Xu, C., & Guo, J. (2017). Regulatory mechanisms of long noncoding RNAs on gene expression in cancers. *Cancer genetics*, *216*, 105-110.

Van der Heul-Nieuwenhuijsen, L., Dits, N. F., & Jenster, G. (2009). Gene expression of forkhead transcription factors in the normal and diseased human prostate. *Cancer related gene expression in the human prostate zones*, 41.

Weinhold, B. (2006). Epigenetics: the science of change.

Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes & development*, *23*(13), 1494-1504.

World Health Organization. (2018). Cancer. Retrieved from: https://www.who.int/health-topics/cancer#tab=tab_3. Retrieved date: June, 2020.

Wu H, Zhang Y. 2014. Reversing DNA methylation: mechanisms, genomics, and biological functions.Cell 156:45–68

Wu, S. C. Y., Grindley, J., Winnier, G. E., Hargett, L., & Hogan, B. L. (1998). Mouse mesenchyme forkhead 2 (Mf2): expression, DNA binding and induction by sonic hedgehog during somitogenesis. *Mechanisms of development*, *70*(1-2), 3-13.

Yang, X., Duan, B., & Zhou, X. (2017). Long non-coding RNA FOXD2-AS1 functions as a tumor promoter in colorectal cancer by regulating EMT and Notch signaling pathway. *Eur Rev Med Pharmacol Sci*, *21*(16), 3586-3591.

Zhang, Y., Liang, C., Zhang, Y., Wang, Z., Li, R., Wei, Z., & Guo, J. (2020). The role of FOXD2-AS1 in cancer: a comprehensive study based on data mining and published articles. *Bioscience Reports*, *40*(11).

Zhou, L., Li, Z., Shao, X., Yang, B., Feng, J., Xu, L., & Teng, Y. (2019). Prognostic value of long non-coding RNA FOXD2-AS1 expression in patients with solid tumors. *Pathology-Research and Practice*, *215*(9), 152449.

Zhu, Y., Qiao, L., Zhou, Y., Ma, N., Wang, C., & Zhou, J. (2018). Long non-coding RNA FOXD 2-AS 1 contributes to colorectal cancer proliferation through its interaction with micro RNA-185-5p. *Cancer science*, *109*(7), 2235-2242.

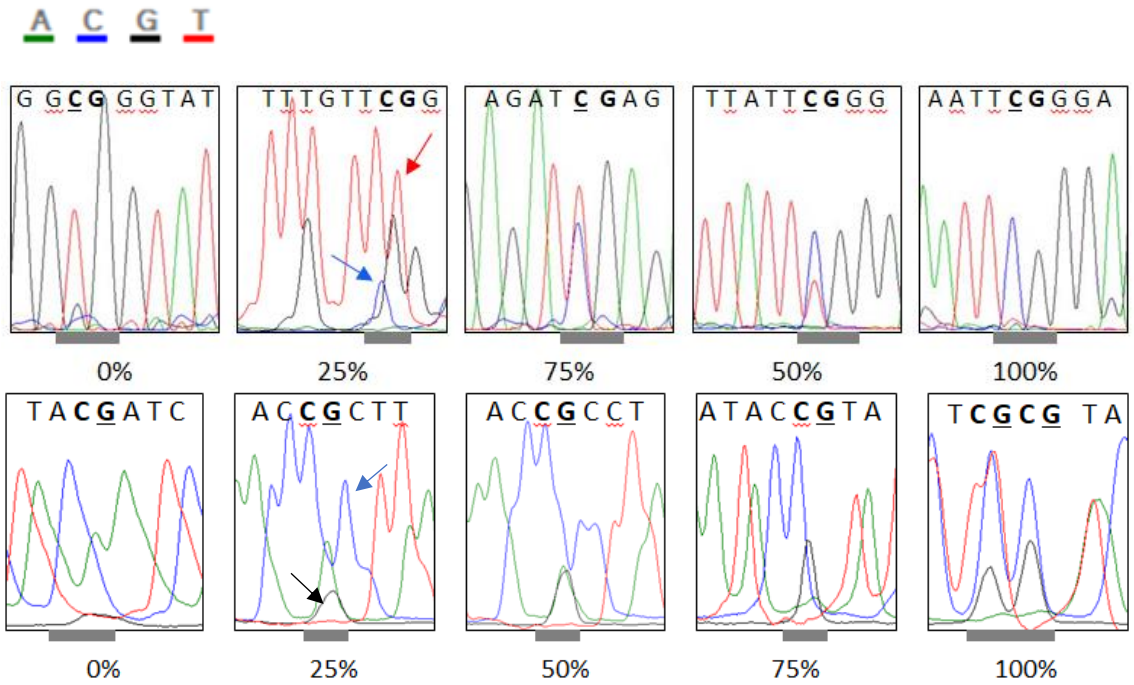
SUPPLEMENTARY INFORMATION

		External PCR			Internal PCR		
		ºC	time (min)	cycles	ºC	time (min)	cycles
Denaturalization		95	3	1	95	3	1
Low astringency amplification	Denaturalization	95	1	5	95	1	30
	Hybridisation	T1*	2		T3*	1	
	Elongation	72	3		72	1,5	
High astringency amplification	Denaturalization	95	1	30			
	Hybridisation	T2*	2				
	Elongation	72	1,5				
Elongation		72	5	1	72	5	1
*Hybridisation temperature varies between genes and regions. For region 2 T1= 48º, T2=50º and T3=50º. For region 1 T1= 54º, T2=56º and T3=60º.							

Supplementary table 1: **Nested PCR thermocycler conditions for region 1 and 2.**

		External PCR		
		ºC	time	cycles
Denaturalization		95	10’	1
Low astringency amplification	Denaturalization	95	30’’	5
	Hybridisation	50	40’’	
	Elongation	72	1’	
High astringency amplification	Denaturalization	95	30’’	30
	Hybridisation	52	40’’’	
	Elongation	72	7’	
Elongation		72	7’	1

Supplementary table 2**: External PCR conditions for region 3.**



Supplementary figure 1: **Sanger sequencing raw data after a bisulfite conversion**. a) Forward Sanger sequencing of a CpG island showing different degrees of methylation. The red arrow points a thymine peak and the blue arrow indicates a cytosine peak, indicating thus that the CpGs is 25% methylated. b) Reverse Sanger sequencing of a CpG island showing different degrees of methylation. The blue arrow points a cytosine peak and the black arrow indicates a guanine peak, indicating thus that the CpGs is 25% methylated.