U.PORTO

MESTRADO

ONCOLOGIA MOLECULAR

Disclosing (epi)genetic mechanisms with potential value as predictive markers of therapy response in Gastric Cancer

Sara Pinto da Silva Teles

M

2018

U.PORTO

Sara Pinto da Silva Teles. Disclosing (epi)genetic mechanisms with potential value as predictive markers of therapy response in Gastric Cancer

Disclosing (epi)genetic mechanisms with potential value as predictive markers of therapy response in Gastric Cancer

Sara Pinto da Silva Teles

M.ICBAS **2018**

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

Sara Pinto da Silva Teles

Disclosing (epi)genetic mechanisms with potential value as predictive markers of therapy response in Gastric Cancer

Tese de candidatura ao grau de Mestre em Oncologia Molecular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Orientador	Doutora Carla Oliveira
Categoria	Investigadora Principal e Professora Afiliada
Afiliação	Instituto de Investigação e Inovação em Saúde/ Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Faculdade de Medicina da Universidade do Porto
Co-orientador	Doutora Patrícia Oliveira
Categoria	Investigadora
Afiliação	Instituto de Investigação e Inovação em Saúde/ Instituto de Patologia e Imunologia Molecular da Universidade do Porto
Co-orientador	Doutora Fátima Gärtner
Categoria	Professora Catedrática e Investigadora
Afiliação	Instituto de Investigação e Inovação em Saúde/ Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Agradecimentos

O trabalho que apresento é o culminar de 3 anos de aprendizagem no grupo de investigação da Dra. Carla Oliveira, no qual tive a oportunidade de desenvolver capacidades importantes como autonomia, comunicação científica, pensamento crítico e trabalho de equipa.

Em primeiro lugar, queria agradecer às minhas orientadoras. À Dra. Carla Oliveira por me ter dado a oportunidade de realizar o estágio da licenciatura em Biologia no seu grupo, e mais tarde a dissertação de mestrado. Se não me tivesse dado essa primeira oportunidade, não teria tido depois esta experiência valiosa de 3 anos e evoluído tanto cientificamente. À Patrícia Oliveira, que estará para sempre na base de todo o meu conhecimento científico. Todas as capacidades que desenvolvi nos últimos 3 anos estão diretamente relacionadas com os ensinamentos que a Patrícia me passou e, sem ela, não teria metade das qualidades que considero ter hoje. Agradeço toda a preocupação e entrega que sempre me dedicou, esperando ser, depois de todo esse esforço, um motivo de orgulho.

Todo o grupo ERiC contribuiu para o meu desenvolvimento nestes anos e aprendi um pouco com todos os seus membros. Em especial, agradeço à Sara Rocha por ter estado sempre ao meu lado e pronta para me transmitir todo o seu conhecimento e experiência, tendo sempre os melhores conselhos para me dar. Agradeço também à Joana Carvalho, por disponibilizar sempre o seu tempo e conhecimento para me ajudar, tendo sido crucial para a minha evolução e a do meu trabalho durante o mestrado. Obrigada à Carla Pereira por todo o apoio que também me deu ao longo dos 3 anos, sempre com a paciência que lhe é caraterística. Obrigada à Marta Ferreira por todo o contributo bioinformático que deu, tão importante neste trabalho, e obrigada à Anabela Ferro por torcer sempre por mim!

Finalmente, quero agradecer aos meus pais e ao João, por acompanharem todas as minhas conquistas e dificuldades, estando intrinsecamente ligados à minha confiança e capacidade de continuar em frente. Sem eles, não seria possível concretizar esta nova etapa.

Table of Contents

Abstract

Gastric Cancer (GC) is one of the most incident and deadliest cancer types. Most cases are diagnosed in advanced stages, with few treatment options available. Several clinical trials have taken place, using antibodies targeting proteins frequently altered in GC, such as the receptor tyrosine kinases HER2, VEGFR2 and FGFR2, or growth factors like VEGFA. *Trastuzumab* (anti-HER2) and *Ramucirumab* (anti-VEGFR2) are clinically approved targeted therapies in GC, however their impact in patient overall survival is poor. Whereas anti-HER2 therapy is chosen based on the overexpression and/or genetic amplification of *ERBB2*/HER2 (15-30% of GC cases), no biomarker has been found that could predict response to anti-VEGFR2, anti-VEGFA or anti-FGFR2 therapies for GC treatment. This is important to select cases for clinical trials and may explain the failure of some of them or the low therapy response to the targeted therapies already tested in these trials.

Although the overexpression of VEGFR2, VEGFA and FGFR2 has been reported in GC at various frequencies, the underlying genetic or epigenetic mechanisms are still unclear. Therefore, our general aim was to explore the potential of the promoter methylation status of *KDR* (codifying VEGFR2) and *FGFR2* as novel predictive biomarkers for anti-VEGFR2 and anti-FGFR2 therapies in GC.

A genome wide preliminary analysis assessing promoter methylation and copy number variants was performed for a cohort of GC and matched normal gastric samples (*n*=47, Cohort 1), by RRBS and WGS, respectively. This study revealed that the majority of tumour samples presented *KDR* (encoding for VEGFR2) promoter hypermethylation, normal *VEGFA* copy number (17% displayed *VEGFA* amplification) and *FGFR2* promoter hypomethylation in comparison with their corresponding normal samples. Moreover, the promoter of *ESRP1*, a known splicing regulator of *FGFR2* responsible for the differential expression of *FGFR2-IIIb* and *FGFR2-IIIc* isoforms, was also hypomethylated in tumour samples. Considering this preliminary data, and to address our general aim, we established two specific aims: in specific aim 1, we investigated the co-occurrence of variations in *KDR* promoter methylation status, *VEGFA* copy number and *KDR* and *VEGFA* RNA expression and; in specific aim 2, we investigated the co-occurrence of variations in *FGFR2* and *ESRP1* promoter methylation status and RNA expression of both genes as well as *FGFR2* isoforms. Both specific aims were explored using selected cases from Cohort 1 for which the promoter methylation status of each gene was validated by bisulfite Sanger sequencing, copy number status validated using copy number qRT-PCR assays and RNA expression was assessed using qRT-PCR. Results were further investigated using tumour samples from two larger TCGA-derived cohorts, referred as Cohorts 2 and 3.

Results from specific aim 1 showed that *KDR* and *VEGFA* were found to be frequently upregulated in tumour samples. However, there was no correlation between *KDR* promoter methylation, *VEGFA* copy number status and *KDR*/*VEGFA* RNA expression in GC tumour samples. These data support the relevance of exploring *KDR* and *VEGFA* RNA expression as putative predictive biomarkers for anti-VEGFR2 and/or anti-VEGFA therapies. Results from specific aim 2 revealed a co-occurrence between *FGFR2* and *ESRP1* promoter hypomethylation, together with the increased RNA expression of *ESRP1* and the epithelial isoform *FGFR2-IIIb* and decreased RNA expression of the mesenchymal isoform FGFR2-IIIc. This was true in gastric tumour samples and *in vitro* models. Therefore, the promoter methylation status of *FGFR2* and *ESRP1*, as well as the RNA expression of *FGFR2-IIIb*, *ESRP1* and *FGFR2-IIIc* could be promising biomarkers for the selection of GC patients for anti-FGFR2/FGFR2-IIIb therapies.

Overall, our results demonstrate the importance of considering different alterations within GC tumour cells, rather than only genetic amplification or protein overexpression, to improve the selection of patients that would benefit the most from a given therapy. In the future, further studies using larger GC cohorts with paired normal samples and clinical data available should be performed to understand the potential of RNA expression and/or promoter methylation status as novel predictive biomarkers for anti-VEGFR2, anti-VEGFA and anti-FGFR2/FGFR2-IIIb therapies.

Resumo

O Cancro Gástrico (CG) é um dos tipos de cancro com maior incidência e dos que mais mortes provoca. A maioria dos casos são diagnosticados em estádios avançados da doença, com poucas opções terapêuticas disponíveis. Vários ensaios clínicos foram realizados, usando anticorpos direcionados a proteínas frequentemente alteradas em CG, como os recetores tirosina cinase HER2, VEGFR2 e FGFR2, ou fatores de crescimento como o VEGFA. *Trastuzumab* (anti-HER2) e *Ramucirumab* (anti-VEGFR2) são terapias dirigidas aprovadas em CG que, mesmo assim, têm pouco impacto na sobrevivência global dos pacientes. Enquanto que a terapia anti-HER2 é escolhida mediante a sobreexpressão proteica/amplificação genética de *ERBB2*/HER2 (15-30% dos casos de CG), nenhum biomarcador existe ainda para predizer a resposta a terapias anti-VEGFR2, anti-VEGFA ou anti-FGFR2 para o tratamento de CG.

Embora a sobreexpressão de VEGFR2, VEGFA e FGFR2 já tenha sido reportada em CG em várias frequências, o mecanismo genético ou epigenético subjacente permanece por esclarecer. Portanto, o nosso objetivo geral foi explorar o potencial do estado de metilação dos promotores do *KDR* (que codifica o VEGFR2) e *FGFR2* como novos biomarcadores preditivos para terapias anti-VEGFR2 ou anti-FGFR2 em CG.

Uma análise genómica preliminar a avaliar a metilação de promotores e variantes no número de cópias genéticas foi executada usando uma coorte de amostras tumorais de CG pareadas com amostras normais (*n*=47, Coorte 1), por RRBS e WGS, respetivamente. Este estudo revelou que a maioria das amostras tumorais apresentava hipermetilação do promotor do *KDR*, número normal de cópias do *VEGFA* (17% exibia amplificação do *VEGFA*) e hipometilação do promotor do *FGFR2*, comparativamente com as amostras normais correspondentes. Para além disso, o promotor do *ESRP1*, um importante regulador do *splicing* do *FGFR2* responsável pela expressão diferencial das isoformas *FGFR2-IIIb* e *FGFR2-IIIc*, estava também hipometilado nas amostras tumorais. Considerando estes dados preliminares, e para responder ao nosso objetivo geral, dois objetivos específicos foram estabelecidos: no objetivo específico 1, investigámos a co-ocurrência de variações no estado de metilação do promotor do *KDR*, no número de cópias do *VEGFA* e na expressão de ARN do *KDR* e *VEGFA* e; no objetivo específico 2, investigámos a co-ocurrência de variações no estado de metilação do promotor do *FGFR2* e *ESRP1* e na expressão de ARN de ambos os genes, assim como das isoformas do *FGFR2*. Ambos os objetivos específicos foram explorados em casos selecionados da Coorte 1, para os quais o estado de metilação de cada gene foi validado por sequenciação de Sanger de ADN tratado com bissulfito, o número de cópias foi validado por qRT-PCR e

a expressão de ARN foi avaliada por qRT-PCR. Estes resultados foram ainda investigados usando amostras tumorais de duas coortes maiores derivadas do TCGA, referidas como Coortes 2 e 3.

Os resultados do objetivo específico 1 mostraram que os níveis de expressão de ARN do *KDR* e *VEGFA* estavam frequentemente aumentados nas amostras tumorais. No entanto, não havia nenhuma correlação entre a metilação do promotor do *KDR*, o número de cópias do *VEGFA* e a expressão de ARN do *KDR* e *VEGFA* em amostras de CG. Estes resultados suportam a relevância de explorar a expressão de ARN do *KDR* e *VEGFA* como putativos biomarcadores preditivos para terapias anti-VEGFR2 e/ou anti-VEGFA. Os resultados do objetivo específico 2 mostram uma co-ocurrência da hipometilação do promotor do *FGFR2* e *ESRP1*, em conjunto com o aumento da expressão de ARN do *ESRP1* e da isoforma epitelial *FGFR2-IIIb* e a diminuição de expressão de ARN da isoforma mesenquimal *FGFR2-IIIc*. Estas co-ocurrências foram observadas em amostras tumorais gástricas e modelos *in vitro*. Assim, o estado de metilação do promotor do *FGFR2* e *ESRP1*, assim como a expressão de RNA do *FGFR2-IIIb*, *ESRP1* e *FGFR2-IIIc* poderão ser biomarcadores promissores para a seleção de pacientes para terapias anti-FGFR2/FGFR2-IIIb.

Em geral, os nossos resultados demostraram a importância de considerar diferentes alterações nas células tumorais de CG, para além da amplificação genética ou sobreexpressão proteica, para melhorar a seleção dos pacientes que mais irão beneficiar de uma determinada terapia. No futuro, mais estudos usando coortes de CG maiores, com amostras normais pareadas e dados clínicos disponíveis, deverão ser realizados para entender o potencial da expressão de ARN e/ou do estado de metilação dos promotores de genes como novos biomarcadores preditivos para terapias anti-VEGFR2, anti-VEGFA e anti-FGFR2/FGFR2-IIIb.

I. Introduction

1. Gastric cancer

Every year, Cancer threatens millions of people throughout the world, being one of the leading causes of death and one of the biggest concerns of the 21st century. As our lifestyle habits continue to boost cancer risk and our population grows and ages, the number of people affected by this disease is expected to rise [1, 2]. Each type of cancer presents its own therapeutic challenges, as each represents a disease with different molecular and phenotypical characteristics. As a result, new therapies are being developed based on specific molecular alterations observed in each tumour type, providing tailored solutions for each patient and/or type of cancer. In cases such as breast or lung cancer, finding a specific molecular alteration to target led to major improvements in the survival of selected cancer patients [3, 4]. However, many other types of cancer lack alternatives to standard chemotherapy regimens and therefore patients have little chance of survival. Such is the case of Gastric Cancer (GC).

Gastric cancer is the fifth most incident and third deadliest type of cancer in the world [5]. These numbers reflect mostly the late diagnosis of this disease, often due either to the lack or late presentation of specific symptoms over the course of the disease, and high intra- and inter-tumour heterogeneity [6]. As a result, most patients present advanced stages of the disease at diagnosis and available treatment is largely ineffective. Some of the main risk factors for GC are *Helicobacter pylori* infection, smoking, male gender, high intake of salted and smoked foods, obesity and Gastroesophageal Reflux Disease (GERD) [7].

The majority of GC cases are adenocarcinomas of the stomach while a small percentage of patients present adenocarcinoma of the gastroesophageal junction. Although most GC cases are of sporadic nature, approximately 10% of patients present familial aggregation, associated with the syndromes Hereditary Diffuse Gastric Cancer (HDGC), Gastric Adenocarcinoma and Proximal Polyposis of the Stomach (GAPPS) and Familial Intestinal Gastric Cancer (FIGC) [8, 9].

1.1. Clinical and histological characterization

The histological classification of GC is mostly governed by the Laurén and World Health Organization (WHO) systems [10, 11]. The Laurén classification divides tumours in two main types: intestinal and diffuse GC. GC cases that do not entirely fit in these two categories are classified as mixed or undetermined types. Tumours of the intestinal type arise generally from intestinal metaplasia, a pre-cancerous lesion characterized by the

replacement of normal gastric cells for differentiated intestinal cells in the stomach epithelium. Tumours of the diffuse type are characterized by undifferentiated and poorly cohesive cells, with no or low gland formation and are considered more aggressive than intestinal GC tumours [12, 13]. The WHO system is composed by four main histological types: tubular and papillary (corresponding to the intestinal type from the Laurén classification), mucinous, poorly cohesive (equivalent to the diffuse type from Laurén classification) and mixed carcinomas [11]. However, this classification does not significantly correlate with clinicopathological or prognostic characteristics [6]. In addition, although the Laurén classification shows clinically relevant GC subgroups, it is still insufficient to improve the current therapeutic landscape. *For this reason, unveiling the molecular complexity of GC tumours could reveal new molecular markers and consequently lead to novel therapies, improving the current outcome of this disease.*

1.2. Molecular characterization

A plethora of molecular alterations contribute to the tumorigenesis of GC and are part of the complex biology of this disease. Like all tumours, GC has several characteristic gene mutations, copy number alterations, epigenetic modifications and transcriptional or translational changes. For example, *TP53* mutations, amplification/overexpression of *ERBB2*/HER2 and E-cadherin expression deregulation are often observed in GC [14-16]. Several studies have aimed at combining different molecular alterations in the hope of finding GC molecular signatures with clinical significance. A seminal paper was published in 2014, by The Cancer Genome Atlas research network (TCGA) combining data derived from whole exome sequencing, array-based DNA methylation and copy number analysis [17], proposing 4 GC subtypes:

- Epstein-Barr virus (EBV) positive tumours, which revealed high levels of DNA promoter hypermethylation (~9%);
- Tumours with microsatellite instability (MSI), with elevated mutational rates and hypermethylation (~22%);
- Genomically stable tumours (~20%);
- Tumours with chromosomal instability (CIN, ~50%).

Although relevant, the clinical significance of this proposed GC molecular stratification has yet to be proven and a consensus on clinically relevant subtypes still needs to be established. Another study aimed at understanding the molecular signatures of GC, was published by the Asian Cancer Research Group (ACRG) which defined 4 GC groups with

different molecular alterations correlated with patient overall survival and recurrence patterns [18]:

- Mesenchymal-like tumours, classified as MSS/EMT, that include those of the diffuse type with worst prognosis and the highest recurrence rates (15%);
- Tumours with microsatellite instability (MSI), presenting hyper-mutated intestinal tumours that also display the best prognosis and lowest recurrence rates (22%);
- Tumours with microsatellite stability and active tumour protein 53, referred as MSS/TP53^{+}, and intermediate prognosis and frequency of recurrence (26%);
- Tumours with microsatellite stability and inactive tumour protein 53, referred as MSS/TP53^{-}, and intermediate prognosis and frequency of recurrence (36%).

These four GC groups derived from distinct expression signatures known to be relevant in the context of GC and were further characterized by somatic alterations. Additionally, the authors verified that the 4 ACRG groups were associated with distinct overall survival benefits, with patients in the MSS/EMT subgroup having the worst prognosis and highest recurrence rates. This study focused on a private GC cohort and later on expanded to two other cohorts to understand the reproducibility of the ACRG proposed GC groups: the TCGA and the Gastric Cancer Project '08 Singapore cohorts. Despite differences between the cohorts, the ACRG proposed groups were considered applicable to all three cohorts. *Overall, the TCGA and ACRG studies enlighten the heterogeneity of GC reflected in the 4 GC subtypes or groups proposed. However, these classifications are not currently used in the clinical context, although proposed groups may encompass novel therapeutic targets relevant for GC patient management.*

1.3. Current management of the disease

Currently, the diagnosis of GC heavily relies on the results obtained from biopsies collected during an endoscopy procedure. After determining the tumour stage, a careful treatment plan is organized by multidisciplinary teams including oncologists, surgeons and radiologists. According to the European Society for Medical Oncology (ESMO) guidelines, patients with early-stage GC are usually fit for surgery with curative potential. In some cases, tumours can be endoscopically removed. However, other patients might need to remove the entire stomach (total gastrectomy) together with pre- and/or postoperative chemotherapy [19]. Nevertheless, the majority of GC cases are diagnosed in advanced stages of the disease. Therefore, surgery is no longer a possibility. In a first-line setting, these patients are treated with a doublet or triplet platinum and fluoropyrimidine regimen.

This type of treatment varies geographically, with the double combination being preferred in Asian countries and the triple regimen being more often used in western countries. This difference likely contributes for the differences observed in the disease outcome between the two populations [20]. In combination with the standard chemotherapy regimen, selected GC patients may be submitted to two targeted therapies: *Trastuzumab* in a first-line setting and *Ramucirumab* in a second-line setting. *Trastuzumab,* the first targeted therapy to be introduced in the treatment plan of GC patients [21], is given to patients with GC tumours expressing the Human Epidermal Growth Factor Receptor 2 (HER2). *Ramucirumab,* approved for the treatment of metastatic GC, is an antibody targeting the Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) given to patients as a second-line therapy without prior selection [22, 23]. *Nevertheless, the survival improvement of GC patients submitted to these targeted therapies is poor, hence GC patients remain without effective long-term solutions.*

2. Targeted therapy in gastric cancer

2.1. Current targeted therapies approved for Gastric Cancer treatment: anti-HER2 and anti-VEGFR2

Over the last decades, several attempts have been made to establish new targeted therapies for GC treatment and improve the current outcome of this disease. However, most clinical trials have failed to present any significant advantage in the addition of these new therapies to the standard chemotherapy regiments. Currently, only two targeted therapies have been approved for GC treatment: the monoclonal antibodies *Trastuzumab* and *Ramucirumab*.

Trastuzumab targets HER2, a known tyrosine kinase receptor involved in many cellular functions important for tumour development [15]. Unlike other members of the HER protein family, HER2 activation is independent of ligand binding. This receptor is either constitutively activated or dimerizes with itself or other HERs, resulting in the initiation of several signalling pathways that contribute to increased cell proliferation, differentiation and invasion [15]. *Trastuzumab* targets the extracellular domain of HER2, impairing receptor dimerization and consequent activation of the tyrosine domain. Moreover, this receptor usually undergoes cleavage, leaving behind a phosphorylated portion that can lengthen the stimulus induced by the HER2 activation. *Trastuzumab* also prevents this cleavage and further pathway stimulation and is believed to induce cell-mediated cytotoxicity and endocytosis [24].

Trastuzumab was firstly introduced to the clinical practice as a new targeted therapy for advanced breast cancer patients with tumours overexpressing HER2 [25]. Years later, the focus turned to GC given the similar rates of HER2 overexpression (15-30%), which is mainly caused by gene amplification, resulting in the worse prognosis of patients carrying these alterations [15, 26]. Consequently, *Trastuzumab* became the first approved targeted therapy for GC treatment after completion of the *Trastuzumab* for Gastric Cancer (ToGA) clinical trial [21]. In this trial, patients with inoperable locally advanced tumours were randomly assigned to two treatment groups: chemotherapy plus *Trastuzumab* and; chemotherapy alone. The authors observed that the first group presented a significant increase of 2.7 months in median overall survival. Next, a stratification of tumours based on HER2 expression was performed using an adaptation of the score described by Hofmann *et al* [27]. The *post-hoc* analysis showed that by dividing patients with tumours with HER2 high *versus* patients with HER2 low expression, the difference in survival was even stronger: 4.2 months [21]. Therefore, the current consensus is to screen GC patients diagnosed with inoperable advanced disease for HER2 protein expression by immunohistochemistry (IHC) and *ERBB2*/HER2 gene amplification by fluorescence *in-situ* hybridisation (FISH). For IHC, tumours are scored from 0 to 3+, with tumours classified as 2+ or 3+ presenting either a weak to moderate (2+) or strong (3+) staining for HER2 in 10% or more tumour cells. Tumours that are classified as HER2 IHC 3+ or tumours presenting both HER2 IHC 2+ and FISH-assay positive results are considered eligible for a combination of chemotherapy and *Trastuzumab* [19].

Ramucirumab is a monoclonal antibody that targets VEGFR2 and has been vastly investigated as a new anti-angiogenic therapy for a variety of cancer types [28]. VEGFR2 is one of the main regulators of angiogenesis, *i.e.* the formation of new blood vessels from pre-existing ones and one of the most important factors driving tumour growth [29]. This receptor, like the other VEGFR protein family members, can be activated in two distinct manners, both causing tyrosine phosphorylation and downstream signalling initiation: 1) through the canonical pathway, by VEGF ligand binding and; 2) through the non-canonical pathway, by non-VEGF ligands or mechanical forces, such as shear stress [30]. *Ramucirumab* is highly specific for the extracellular domain of VEGFR2, having a great binding affinity for this receptor, impairing VEGFR2 interaction with ligands [31].

To understand the relevance of targeting VEGFR2 in the setting of GC, a pre-clinical trial was performed using a mouse-specific anti-VEGFR2 antibody in mouse GC xenograft models [32]. This study revealed that by blocking VEGFR2, tumour growth *in vivo* was inhibited, strengthening the therapeutic importance of targeting VEGFR2 [32]. Subsequent safety and dose-finding trials were performed in cancer patients and *Ramucirumab* was

eventually approved for GC treatment after completion of two phase III studies: the REGARD and RAINBOW trials [32]. In the REGARD trial, 355 patients with metastatic or inoperable gastric or gastroesophageal cancer with disease progression after first-line treatment were randomly assigned to receive *Ramucirumab* or placebo. Patients treated with the antibody presented an increase of 1.4 months in overall survival [22]. Simultaneously, the RAINBOW trial initiated the safety and efficacy assessment of *Ramucirumab* combined with Paclitaxel in patients with gastric or gastroesophageal adenocarcinoma who had progressed after first-line chemotherapy. Patients randomly assigned to the *Ramucirumab* plus Paclitaxel group presented a significantly improved median overall survival in comparison with the Paclitaxel alone group: 9.6 months *vs.* 7.4 months, respectively [23]. The results from these two trials culminated in the approval of *Ramucirumab* as a second-line therapy, combined with chemotherapy, for patients with gastric or gastroesophageal cancer after disease progression. Nevertheless, contrarily to *Trastuzumab*, there are currently no biomarkers to predict response to *Ramucirumab* therapy. The authors from both trials performed exploratory analyses to find potential predictive biomarkers [33, 34]. However, the level of assessed VEGFRs, VEGFs or other cytokines in either tumour or serum samples failed to present any significant association with *Ramucirumab* efficacy. Therefore, there is still the need for appropriate predictive biomarkers to select patients that will benefit the most from *Ramucirumab* therapy.

2.2. Targeted therapies tested in Gastric Cancer without approval

Although *Trastuzumab* and *Ramucirumab* are currently approved for GC treatment, the poor survival benefit underlies the continuous effort for development of novel targeted therapies. However, most have yet to present any significant survival benefit for GC patients. The majority of these unsuccessful therapies target known cancer-associated receptors/ligands such as the Vascular Endothelial Growth Factor A (VEGFA), Epidermal Growth Factor Receptor (EGFR), Hepatocyte Growth Factor Receptor (MET) and Fibroblast Growth Factor Receptor 2 (FGFR2).

2.2.1 Targeting VEGFA in GC

VEGFA was the first cytokine and member of the VEGF family to be discovered as a tumour-associated angiogenic factor, being the principal regulator of this complex process [35]. This family also encompasses VEGFB, VEGFC, VEGFD and placental growth factor (PlGF), which have different affinities to the VEGF-receptors VEGFR1, VEGFR2 and VEGFR3,

which in turn have distinct kinase activities. Together, VEGFs and VEGFRs are key regulators of important processes such as vasculogenesis, angiogenesis, lymphangiogenesis and vascular permeability, both in the physiological and pathological context [36]. In GC, VEGFA overexpression and amplification was seen in up to 58% and 7% of cases, respectively [37]. Many studies have tried to correlate VEGFA protein expression with patient prognosis, however presented contradictory results [38]. More recently, a meta-analysis showed that VEGFA expression was associated with poor overall survival and disease-free survival of GC patients [38]. Hence, successfully targeting VEGFA in GC continues to be an objective of researchers that try to implement new effective treatments for these patients.

Bevacizumab was the first monoclonal antibody designed to target VEGFA and is currently approved for treatment of several cancer types, such as advanced colon, lung, breast, ovarian, endometrial and clear cell renal carcinoma [39-43]. In advanced GC, *Bevacizumab* was the first anti-angiogenic therapy tested: phase II trials supported the anti-angiogenic effect of *Bevacizumab* in GC patients, reporting a 42-67% overall response rate (ORR) in the test group, as well as a median progression free survival (PFS) of 6.6-12 months and an overall survival (OS) of 8.9-16.2 months [44]. For this reason, the Avastin (or *Bevacizumab*) for Advanced Gastric Cancer (AVAGAST) trial was initiated [45]. In this study, patients with unresectable advanced adenocarcinoma of the stomach or gastroesophageal junction were randomly assigned to receive either chemotherapy (fluoropyrimidine-cisplatin) plus *Bevacizumab* or chemotherapy alone, as a first-line therapy. Despite the first group presenting a significant ORR and PFS, there was no overall survival benefit in the combination of *Bevacizumab* with standard chemotherapy. Further analyses showed there were geographical differences in the efficacy of this therapy, with Asian patients presenting no advantage from this treatment. Moreover, the same authors published *a posteriori* exploratory study to find possible biomarkers for *Bevacizumab* efficacy [46]. Circulating VEGFA in blood samples was evaluated, as well as the expression of VEGFA, VEGFR1, VEGFR2 and the co-receptor neuropilin-1 in tumour samples. In addition, HER1 and HER2 were also assessed as potentially prognostic markers, given the fact that these patients did not receive *Trastuzumab*. The authors concluded that high baseline VEGFA plasma levels and low baseline neuropilin-1 tumour expression were candidate biomarkers for *Bevacizumab* efficacy after observing an increased OS in advanced GC patients who had been given this therapy. However, there was no correlation between VEGFA plasma and tumour expression levels and the observed trends for OS and PFS were again mostly seen in non-asian patients.

Further exploring the efficacy of *Bevacizumab*, the AVATAR clinical trial investigated the addition of this antibody to first-line chemotherapy in 202 Chinese patients with advanced GC [47]. The authors obtained a similar result to the AVAGAST study: the addition of this antibody to the standard chemotherapy presented no advantage for GC patients. Interestingly, the AVATAR Chinese population therapy response was closer to European and Pan-American GC patients from the AVAGAST trial than to its Asian subgroup, constituted by 90% of Japanese GC patients. This highlighted the differences between the Chinese and Japanese populations, which have very distinct GC clinical management strategies. Overall, these studies demonstrate the importance of elucidating the heterogeneity associated with Asian and non-Asian populations in GC that may confound the results obtained in these trials, as well as the need for proper biomarkers to better predict Bevacizumab efficacy [46]. *In summary, no predictive biomarkers for anti-VEGFA therapies currently exist for GC patient stratification and treatment.*

2.2.2 Targeting EGFR in GC

EGFR is part of the HER receptor family, encoded by *HER1* gene, being overexpressed in 2.3-40% and amplified in up to 10% of GC cases [17, 37, 48]. Two major monoclonal antibodies have been tested in GC patients as potential new targeted therapies: *Cetuximab* and *Panitumumab* [49]. *Cetuximab* binds to the extracellular domain of EGFR with higher affinity than the natural ligands, preventing the initiation of signal transduction. In the EXPAND trial, 904 advanced gastric or gastroesophageal cancer patients were randomly assigned to receive *Cetuximab* plus chemotherapy or chemotherapy alone [49]. The addition of this antibody to the chemotherapy regimen brought no benefit in terms of OS, PFS or ORR. Concurrently, the REAL3 trial tested the efficacy of *Panitumumab* (added to chemotherapy) in a group of advanced or inoperable esophagogastric cancer patients [50]. *Panitumumab* targets the extracellular portion of EGFR and had previously shown a survival benefit in advanced colorectal cancer patients. However, in the REAL3 study *Panitumumab* was shown to be associated with lower OS and PFS. The authors associated this unexpected effect with the need to lower the chemotherapy dosage and not with the presence of the antibody being tested. In addition, a retrospective analysis of tumour EGFR expression was also performed by immunohistochemistry. EGFR presented consistently low expression across samples and there was no difference in OS or PFS in patients between the two treatment groups. Other molecular events could be associated with this lack of efficacy: for example, in metastatic colorectal cancer, *KRAS* mutations have been shown to confer resistance to anti-EGFR antibodies [51]. However, in GC, KRAS mutations

occur in low frequency (3-10%) [50]. Nevertheless, mutations in other genes may have similar significance in GC and remain unexplored. *In summary, there is a clear lack of patient stratification that is likely contributing to the inefficiency of anti-EGFR therapies.*

2.2.3 Targeting MET/HGF in GC

MET is a receptor mainly expressed in epithelial tissues that, after binding of its ligand Hepatocyte Growth Factor (HGF) or autodimerization, is activated and commands a signalling pathway responsible for many functions important for morphogenesis, such as cell proliferation, survival and migration [52]. In cancer, MET and/or HGF can be overactivated, resulting in increased tumour growth and invasion. In GC, MET can be amplified and overexpressed in 4-10% and approximately 50% of cases, respectively [44]. Three promising monoclonal antibodies addressing MET/HGF-related activity in GC were tested: *Rilotumumab*, targeting HGF; *Onartuzumab* and *ABT-700* both targeting MET [44].

After a phase II study demonstrating a correlation between higher MET expression and *Rilotumumab* efficacy, two phase III studies were initiated (the RILOMET-1 and RILOMET-2) testing the addition of this antibody to chemotherapy in advanced GC patients [53, 54]. However, RILOMET-1 reported an increased number of deaths in the *Rilotumumab* plus chemotherapy group and both trials were stopped.

Onartuzumab targets the extracellular portion of MET, impairing HGF binding. One phase II trial reported high toxicity levels related to the presence of this antibody [55]. Consequently, the METGastric phase III trial that had already begun was stopped [56]. Nonetheless, subgroup analyses from both trials suggested that *Ornatuzumab* did not add any benefit to the intent-to-treat or MET-overexpressing patient population.

ABT-700 has been shown to have anti-tumour activity in four patients with gastric or gastroesophageal cancer in a phase I study [57]. More recently, another study demonstrated that this antibody prevents MET activation by both preventing HGF binding and MET autodimerization. Furthermore, *ABT-700* induced apoptosis and suppressed tumour growth in tumour cell lines and cancers harbouring MET amplification [58]. Although the sample size was small, this antibody was considered a good candidate for pre-screened MET-amplified GC patients, however no phase II/III trials are currently underway [59]. *Overall, anti-MET and anti-HGF therapies have lacked success. Nevertheless, these targets remain biologically relevant for GC therapy as novel clinical trials are expected to be developed.*

2.2.4. Targeting FGFR2 in GC

Another popular target in GC is FGFR2, a member of the FGFR family that includes the Receptor Tyrosine Kinases (RTKs) FGFR1, FGFR3 and FGFR4. Together with their large group of FGF ligands, FGFRs are in the centre of command of many important developmental programmes such as gastrulation and organogenesis, regulating functions such as cell proliferation, survival and migration [60]. With different affinities, FGFs bind to FGFRs which dimerize and induce the phosphorylation of their tyrosine kinase domains, activating downstream signalling pathways. In cancer, this FGF/FGFR signalling is dysregulated, owning to molecular alterations such as activating mutations, amplifications, fusion genes and FGFs or FGFRs overexpression, creating autocrine or paracrine loops that influence tumour growth by promoting cell proliferation, survival and angiogenesis [60, 61]. Therefore, FGFRs are desirable targets for the treatment of many types of cancer.

The role of FGFR2, initially referred as *K-sam*, has been studied in GC since its amplification was first detected in the gastric cancer cell line KATO-III [62]. Subsequent studies demonstrated that *FGFR2*/*K-sam* codified two different isoforms that differed in the third immunoglobulin domain and had distinct binding affinities to FGFs and presented tissue-specific expression [63-65]. The FGFR2-IIIb isoform (also called KGFR) was mainly expressed in carcinoma cell lines, while the FGFR2-IIIc isoform (also called Bek) was only expressed in nonepithelial cancer cells [65]. Later, high FGFR2/K-sam protein expression was confirmed in gastric cancer tumours [66]. However, reports addressing FGFR2 protein expression in GC have been controversial, with studies detecting FGFR2 overexpression from 0% up to 51% of GC cases [67-69]. This disparity could be due to high intra-tumour heterogeneity, differences in protein detection techniques, regional differences between the populations being tested or other underlying genetic events: in fact, *FGFR2* is amplified in less than 10% of GC cases [17, 70, 71]. However, the correlation between FGFR2 copy number status and clinicopathological data of patients is still debated [70, 72, 73]. Some studies have tried to correlate *FGFR2* gene amplification with FGFR2 or FGFR2-IIIb protein overexpression [74, 75]. In one report, *FGFR2* was amplified in 2.7% of GC cases and all cases presented FGFR2-IIIb protein overexpression [74]. In another study, 4% of GC cases displayed FGFR2-IIIb overexpression and 92% of these also presented *FGFR2* amplification [75]. Although *FGFR2* genetic amplification could explain some GC cases with FGFR2 overexpression, these constitute less than 10% of all GC cases. Therefore, the molecular mechanisms underlying FGFR2 overexpression in GC remain largely unclear.

Anti-FGFR2 therapies in GC have yet to reveal clinical benefits. Several inhibitors have been tested, from compounds targeting all FGFRs (pan-FGFR), to specific antibodies

designed to recognize FGFR2 or the FGFR2-IIIb isoform [76]. The last promising clinical trial to report results was the SHINE study that tested the efficacy of AZD4547 in combination with chemotherapy, as a second-line therapy, in patients with locally advanced or metastatic GC [77]. Although AZD4547 recognized FGFR1, FGFR2 and FGFR3, tumours were required to have *FGFR2* amplification for patients to enter the study. However, neither the overall nor the stratified population presented any significant response to this therapy. Moreover, a biomarker analysis performed by the same authors showed that both amplified and non-amplified tumour samples presented similar FGFR2 expression levels, with only 6 out of 24 amplified GC tumours showing FGFR2 overexpression. This could be due to the intra-tumour heterogeneity of FGFR2 expression in these cases, as reported by the authors. Moreover, setting a correct cut-off for FGFR2 amplification is still needed for patients that present this alteration and respond to therapy [78]. *In summary, anti-FGFR2 therapies have yet to prove their clinical relevance, a problem likely associated with the lack of relevant biomarkers to predict efficient therapy response.*

Overall, the findings on currently unapproved targeted therapies in GC highlight the unmet need to find proper biomarkers to select available and developing therapies. This could be achieved by uncovering new molecular mechanisms regulating the expression of currently GC actionable genes, such as the previously described *VEGFA*, *EGFR*, *MET* and *FGFR2*.

3. FGFR2 and VEGFR2 signalling in Gastric Cancer

As occurred for clinical trials testing anti-VEGFA, anti-EGFR, anti-MET and anti-FGFR2 targeted therapies, most other targeted therapies are tested in GC patients without their stratification based on any predictive markers of therapeutic response. Moreover, studies that search for a correlation between a given molecular feature and therapy response, recurrently consider only genetic amplification and/or protein overexpression of the molecule being targeted, ignoring other mechanisms. Until now, only *Trastuzumab* has a predictive biomarker associated with its higher efficacy in the treatment of GC: *ERBB2*/HER2 amplification and/or protein overexpression. Therefore, other molecular alterations or partners regulating the expression of popular targets in GC, such as VEGFR2, VEGFA, EGFR, MET and FGFR2, should be studied as potential predictive biomarkers of therapy response. In the next subchapter, we explored FGFR2 and VEGFR2 signalling pathways as well as mechanisms of expression control in GC, as examples.

3.1. FGFR2

3.1.1. FGFR2 signalling in the normal and cancer contexts

FGFR2 and the other FGFR family members have been deeply studied in several types of cancer, including GC. There are five FGFRs belonging to the same family, with FGFR1, FGFR2, FGFR3 and FGFR4 being RTKs constituted by an intracellular tyrosine kinase domain, a transmembrane domain and an extracellular ligand-binding domain composed by three immunoglobulin-like structures (IgI-III) [79]. FGFs are a family of 23 proteins, with 18 constituting active FGFR ligands (3 hormone-like and 15 canonical FGFs) [80]. The majority of FGFs (canonical FGFs) are seized at the extracellular matrix by Heparan Sulphate Proteoglycans (HPSG). After being released, FGFs bind to FGFRs present at the membrane in a ternary complex with HPSGs. Hormone-like FGFs bind poorly to HPSGs, thus depending on the help of Klotho proteins to bind to FGFRs. After binding, these receptors usually undergo dimerization and the tyrosine residues of the FGFRs are transphosphorylated. This allows adaptor proteins such as FGFR substrate 2 (FRS2) to bind. FRS2 acts as a docking site to several other proteins, including SOS (Son of Sevenless) and GRB2 (growth factor receptor-bound 2) that activate the RAS-MAPK-ERK pathway. Furthermore, another GRB2-related protein, GRB2-associated binding protein 1 (GAB1) can bind to FRS2 and activate the PI3K-AKT signalling pathway. In a different portion of the intracellular domain of FGFRs, phospholipase Cy (PLCy) also binds to the phosphorylated carboxy-terminal tail of the receptors and eventually reinforces the RAS-MAPK-ERK pathway [60].

In cancer, the enhanced activation of FGFRs results from multiple molecular alterations, from gene amplification to mutations. *FGFR1* is the most frequently amplified gene in the FGFR family: amplifications have been reported in 6-17% and 5-15% of lung and breast cancer cases, respectively [81-83]. *FGFR2* is amplified in 5-10% and 2-4% of gastric and breast cancer cases, respectively [73, 81, 84]. Somatic activating mutations are more frequently seen in *FGFR2* and *FGFR3*. *FGFR2* presents mutations in non-small-cell carcinomas and endometrial, gastric and urothelial cancer, while *FGFR3* is mutated in 75% of non-muscle invasive urothelial cell carcinomas and in 15% and 5% of high-grade invasive urothelial cancers and cervical cancers, respectively. FGFR oncogenic fusions have also been reported and most of the fusion partners encompass dimerization domains which lead to ligand-independent dimerization of the receptors and activation of downstream pathways. These mechanisms can lead to FGF or FGFR overexpression, causing an overactivation of the signalling pathways important for tumour growth.

The intricacy of FGF/FGFR signalling arises from the different homo- and heterodimers that can be formed, which associate with different affinities to FGFs, as well as several alternative splicing events that add variability to this process [85, 86]. FGFRs are able to dimerize in the presence or absence of ligands and, regardless of ligand binding FGFRs can become phosphorylated [86-88]. For example, FGFR2 has been shown to form homodimers, as well as heterodimers with FGFR1 and FGFR3 in the presence or absence of ligands, leading up to its activation via phosphorylation [86, 87]. This ability to be in a phosphorylated state without the need for a ligand, supports the pathogenicity associated with FGFR overexpression. In addition, the heterodimerization capacity of FGFRs supports the need for assessment of all family members in a disease context for an accurate inference of FGF/FGFR pathway status. Furthermore, the relevance of FGFRs is likely different between cancer types and hence further investigation is required.

Another important feature of FGF/FGFR signalling is FGFR alternative splicing events. In particular, the alternative splicing of exons 7, 8 and 9, which encode the extracellular Ig-III region of *FGFR1, FGFR2 and FGFR3*, has been shown to be relevant in different cancer types [89]. Exons 7, 8 and 9 correspond to the IIIa, IIIb and IIIc protein domains, respectively. The IIIb and IIIc regions encode for the C-terminal portion of the receptors and the alternative inclusion of either IIIb or IIIc originates different isoforms of each receptor with distinct biological consequences [90]. For example, the deletion of the IIIc region in *FGFR1* and the deletion of IIIb in *FGFR2* are both embryonically lethal. These isoforms are of particular importance during embryogenesis and organogenesis given that each participates in tissue-specific programs. The IIIb isoform is mainly expressed in epithelial tissues, while the IIIc isoform is mostly expressed in mesenchymal tissues. This is particularly true for *FGFR2*, whose isoforms strictly follow this expression pattern [91]. FGFs are key partners in this expression diversity, given the different binding affinities to each isoform, with FGFR-IIIb variants having the strictest binding-pattern [90]. In GC, FGF7/FGFR2-IIIb signalling has been shown to be associated with disease progression. FGF7 exerts its effect in a paracrine manner being produced by fibroblasts and activating FGFR2-IIIb-expressing gastric cancer cells [92]. However, the relevance of the correlation between expression of FGFs and that of FGFR2 or its isoforms for the prediction of therapy response has yet to be proven. Moreover, and as previously mentioned, it has not yet been revealed whether expression or molecular alterations affecting FGFRs are predictive biomarkers for targeted therapy efficacy in GC [70, 74]. In fact, *FGFR2* genetic amplification, alternative splicing and/or protein overexpression have proven not to be adequate biomarkers in GC. Nevertheless, FGFR2 and FGFR2-IIIb protein expression has been found deregulated in GC using different cohorts [74, 75]. This observation supports the

study of other genes/proteins known to be involved in the alternative splicing of FGFR2. A key molecule in this process is the RNA binding protein ESRP1 (Epithelial Splicing Regulatory Protein 1), a protein seldomly explored in the context of cancer.

3.1.2. ESRP1 and the alternative splicing of FGFR2

A decade ago two new important epithelial cell-type specific hnRNPs were found to regulate the splicing of a variety of transcripts with tissue-specific expression: The Epithelial Splicing Regulatory Protein 1 and 2 (ESRP1 and ESRP2) [93, 94]. The authors first showed that expression of ESRP1 and ESRP2 was necessary for the expression of the epithelial FGFR2-IIIb isoform. Knocking down ESRPs led to an isoform switch favouring FGFR2-IIIc expression, with ESRP1 showing great relevance in this process [93]. Subsequent studies showed that ESRPs regulated the alternative splicing of *CD44*, *CTNND1* and *ENAH*, genes with differential expression during the Epithelial-to-Mesenchymal Transition (EMT). EMT is a known process important for embryogenesis, organogenesis and wound healing in the normal context, as well as tumour invasion and migration in the cancer context [95]. ESRPs are considered key regulators of EMT, as they regulate the genetic switch between epithelial or mesenchymal isoforms of proteins with functions such as regulation of actin cytoskeleton, cell polarity, adhesion and migration [96]. This discovery contributed for the revelation of a major alternative splicing program that occurs during EMT, involving not only RBPs promoting an epithelial splicing pattern, such as ESRPs and RBM47, but also RBPs promoting a mesenchymal splicing program, such as QKI and RBFOX2 [97, 98].

With such a responsibility in EMT and given the association between EMT and cancer, it was hypothesized that ESRPs could also have an important role in cancer progression. Following up on this, some studies have reported increased ESRPs expression in carcinomas [99, 100]. However, the correlations found between ESRPs overexpression and overall survival of patients were contradictory, with ESRPs being associated both with cancer progression and anti-tumorigenic effects [100, 101]. Moreover, the authors of one study showed that ESRPs presented a plastic expression during breast cancer progression, being expressed in carcinomas *in situ* and in metastases, while losing their expression at the invasive front, thus regulating negatively cell motility [99]. Further studies were conducted exploring this role in cell motility but presented conflicting results as well [96, 100, 102]. Overall, ESPRs may have opposite roles in cancer progression depending on cancer type, presenting a favourable prognostic in some cases, while aiding cell invasion in others [103].

The expression regulation of ESRPs and ESRP-promoted alternative splicing of *FGFR2* is therefore worth further investigation, particularly in the setting of FGFR2-deregulated cancers, such as GC. Recently, one study showed that ESRP1-overexpressing colorectal cancer cells had also increased FGFR1 and FGFR2 expression and promoted the formation of macrometastases in mice [104]. In GC, there is only one hint of ESRP1-related malignancy: a region in chromosome 8 that included the *ESRP1* locus presented an increased copy number, an observation that was correlated with poor overall survival of patients [105]. Given the known FGFR2 deregulation in GC, it would be interesting to understand whether ESRP1 expression deregulation is contributing to FGFR2/FGFR2-IIIb-related malignancy in gastric tumours.

3.2. VEGFR2 signalling in the normal and cancer contexts

As previously mentioned, VEGFR2 is one of only two molecules with a targeted therapy approved for GC treatment. Throughout the years, VEGF/VEGFR signalling has been a popular target for the development of therapies across different cancer types and often successful due to its central role in neoangiogenesis. VEGFs encompass five growth factors, VEGFA, VEGFB, VEGFC, VEGFD and PlGF, which bind with different affinities to three receptor tyrosine kinases VEGFR1, VEGFR2 and VEGFR3 that contain an intracellular tyrosine kinase domain, a transmembrane domain and seven extracellular immunoglobulin homology domain repeats [106]. In general, VEGFRs are activated after binding of a VEGF ligand. Similarly to other RTKs, this induces receptor dimerization that leads to phosphorylation of the tyrosine kinase domain which then acts as binding site for other molecules involved in several signalling pathways. For VEGF/VEGFR mediated signalling several factors are at play: besides the fact that these receptors and ligands can be regulated by alternative splicing, forming different protein isoforms, VEGFRs are able to form homo- and heterodimers, changing their binding affinities and the transduced signal. Moreover, VEGFR activation and signalling can also be modulated by: binding of co-receptors such as Neuropilin proteins 1 and 2 (NRP1, NRP2); binding of non-VEGF ligands and; mechanical forces sensed by the cell [30, 107-110]. Consequently, VEGF/VEGFR signalling is greatly diverse, being able to induce different biological outcomes. Exploring this mechanistic diversity has brought new insight into anti-angiogenic therapy.

VEGFR2 is known to be the principal regulator of VEGF/VEGFR signalling in blood cells. As a result, the mechanisms behind VEGFR2 expression regulation in the normal and pathological context have been thoroughly explored. Through the canonical pathway, VEGFR2 is mainly activated by binding of VEGFA, VEGFC or VEGFD. However, the

binding affinity of each growth factor to VEGFR2 is influenced by the receptor dimerization, being able to form both homodimers as well as heterodimers with VEGFR1 and VEGFR3 [30]. Moreover, VEGFR co-receptors NRP1 and NRP2 form distinct complexes that are able to modulate VEGFR activation and signalling. In the non-canonical pathway, VEGFR2 can be activated after binding of different non-VEGF ligands. In addition, by forming a mechanosensory protein complex, mechanical forces can induce VEGFR2 phosphorylation and activate the receptor. Regardless of how VEGFR2 is activated, different key signalling pathways for endothelial biology are initiated: the phospholipase Cy (PLCy)-ERK1/2, PI3K-AKT-mTOR and small GTPases pathway [30]. These pathways regulate important functions such as cell survival, migration, polarization and vascular development [30, 110]. These processes, overlapping known hallmarks of cancer [29], underly the importance of VEGFR2 deregulation in cancer progression. Nevertheless, finding a biomarker within VEGFR2 deregulation in GC has proven to be difficult. In GC, some studies have reported VEGFR2 expression, although its clinical significance has been contradictory. In one study, which evaluated the expression of several growth factors and VEGFRs in GC by immunocytochemistry, VEGFR2 was expressed in just 4.4% of gastric adenocarcinomas and was associated with poor overall survival [111]. In another study, VEGFR2 expression was not detected in cancer cells, however it was present in 53% of stromal cells [112]. Moreover, no association between VEGFR2 protein expression and therapy efficacy was observed in a biomarker analysis performed following up on the REGARD trial, one of the studies responsible for the approval of *Ramucirumab* for GC treatment [33]. In fact, VEGFR2 expression in tumour cells was minimal, in line with the previously mentioned studies. Even with almost 87% of GC samples presenting VEGFR2 expression in tumour vessels, patients with high or low expression had no significant difference in survival. Therefore, a proper biomarker for anti-VEGFR2 therapies is still lacking.

Several studies have highlighted the relevance of FGFR2 and VEGFR2 signalling in GC. Although only anti-VEGFR2 therapy is currently approved for GC treatment, strong evidences support the continued relevance of FGFR2 and its isoforms as important actionable molecules. Nevertheless, strong predictive biomarkers, beyond protein expression or gene amplification, are still lacking to effectively stratify GC patients and improve therapy response. For example, the methylation status of the promoter of these genes and the presence of known alternative isoforms could represent novel FGFR2 and/or VEGFR2 molecular biomarkers.

4. Epigenetic mechanisms controlling gene expression in gastric cancer

Cancer results mostly from structural alterations in the DNA sequence of several genes that induce abnormal expression patterns and disturb normal cell phenotype and behaviour [29]. However, there are other modifications that can occur without altering the DNA sequence and still regulate expression patterns, both in the normal and cancer context. Epigenetics studies these expression-based alterations that result from a complex collaboration between non-coding RNAs, DNA methylation, histone and nucleosome modifications and transcription factors [113]. In cancer, these mechanisms have been thoroughly explored as key molecular events underlying cancer progression and pinpointed novel therapies. For example, recognizing increased DNA methylation and loss of histone acetylation in tumours led to the approval of different therapeutic agents, such as azacitidine and decitabine in leukemia, with encouraging results for cancer patients [113-115]. In GC, genetic and histological profiles have yet to show enough potential to predict therapy response and impact the outcome of this disease. However, like other types of cancer, many epigenetic modifications are involved in the carcinogenesis and malignancy of GC at the 1) RNA; 2) histone and; 3) DNA levels [116]:

1) At the RNA level, the existence of numerous micro-RNAs (miRNAs) with oncogenic or tumour suppressor functions, as well as a histological type-specific miRNA signature have been discovered in GC [116-118]. Moreover, miRNAs have been extensively investigated as potential diagnostic and prognostic biomarkers in liquid biopsies, with increasing sensitivity when compared with other standard cancer biomarkers [117].

2) Several histones have shown to correlate with either active or repressive chromatin and transcriptional states. Currently, studying the global influence of histone modifications is possible using either microarray or next-generation sequencing technologies. For example, one study found over 100 *loci* affected by one suppressive histone modification in GC samples [119]. Therefore, exploring histone alterations in GC could open new avenues to stratify GC patients and uncover new therapeutic targets.

3) Studying DNA methylation in GC has provided a deeper understanding on the expression regulation of several genes important for carcinogenesis such as tumour suppressors or cell cycle regulators [120, 121]. One example of the importance of this type of analysis was the discovery of the CpG Island Methylator Phenotype (CIMP) in GC, originally described in colorectal cancer [122, 123]. The CIMP phenotype was shown to be associated with the TCGA-defined EBV and MSI

groups, called EBV-CIMP and Gastric-CIMP tumours, respectively, and presented different molecular characteristics [17]. Therefore, DNA methylation analysis could have an impact in both patient stratification, as well as in the discovery of new predictive biomarkers for therapy response.

4.1. DNA methylation

DNA methylation generally consists in the addition of a methyl group to a cytosine linked to a guanine by one phosphate bond called CpG site/dinucleotide [124]. It is one of the most studied epigenetic deregulations in cancer and is largely focused on CpG site methylation analysis. CpGs are often located near Transcription Start Sites (TSSs), however they can also be found across intergenic regions [125]. DNA regions rich in CpG sites are called CpG islands and are usually located near TSS/gene promoters [124]. In the normal context, DNA methylation is responsible for several important functions such as silencing genetic imprinted genes, X chromosome inactivation and maintenance of genetic stability [126]. These patterns, that can be inherited across cell divisions or established *de novo*, are regulated by the enzymes DNA methyltransferases (DNMTs) [126, 127]. In cancer, loss of DNA methylation in tumours (hypomethylation) across the genome was the first described epigenetic change in cancer [128]. During tumour formation, DNA hypomethylation occurs in many genomic regions such as repetitive sequences and retrotransposons, leading to increased genomic instability [129]. Increased methylation (hypermethylation) was also observed in tumours, in specific regions of the genome such as CpG islands in the promoter of genes. In particular, this promoter hypermethylation was often observed in tumour suppressor genes, important for the regulation of cell proliferation and survival, that led to their transcriptional inactivation, contributing for tumour formation [129]. However, this direct and inverse correlation between DNA methylation and gene expression has been further explored and sometimes contradicted [130, 131]. For example, Guillaumet-Adkins *et al.* revealed that acute myeloid leukemia cell lines presented hypermethylation of the promoter of Wilms' tumour 1 gene (*AWT1)*, however it was associated with increased RNA expression [131]. Moreover, researchers have also investigated the role of hypomethylation in the promoter of genes, in correlation with increased expression, as an activation mechanism of cancer-related genes [132, 133]. For example, Yuille MR *et al.* suggested that promoter hypomethylation of the *TCL1* gene may be an alternative activating mechanism, other than chromosomal rearrangement, of its expression as it correlated with TCL1 increased expression in mature B-cell malignancies [134].

In GC, besides defining a CIMP phenotype, DNA methylation accompanies *Helicobacter pylori* (*H. pylori*) infection, where hypermethylation of several genes such as *p16*, *LOX*, *THBD* and *HAND1* has been observed [135]. Eradicating *H. pylori*, which is thought to reduce GC risk [136], decreases DNA methylation of these genes as well, supporting the role of DNA methylation in GC carcinogenesis and malignancy. Furthermore, promoter DNA methylation of known target genes in GC such as *FGFR2* and *KDR*, (encoding VEGFR2) has been explored as a new regulatory mechanism of gene expression and therapy response in several types of cancer [137, 138]. One study found that the promoters of *VEGFR1* and *VEGFR2* were methylated in GC cell lines, while the *VEGFA* promoter was demethylated [137]. In addition, the authors found that cancer cell lines with demethylated *VEGFRs* showed proliferation inhibition by tyrosine kinase inhibitors, contrarily to cancer cell lines with methylated *VEGFRs*. Another study found that the *FGFR2* promoter was differently methylated in a panel of GC cell lines [138]. Moreover, the authors found that *FGFR2*-hypermethylated GC cell lines had decreased mRNA expression.

Overall, DNA methylation has been increasingly explored as an important regulator of expression of genes relevant for the malignancy of many types of cancer, including GC. However, very little is known about the regulation of these and other therapeutic targets by epigenetic mechanisms, in particular by DNA methylation. This type of epigenetic modification in the promoter of GC-associated actionable genes could constitute novel biomarkers for GC progression, prognosis and therapy response.

4.2. How to explore gene-specific and genome-wide DNA methylation

There are several methods to assess DNA methylation, which could be divided in two major types of analysis: genome-wide or region-specific [139-141]. Regardless of the type of methylation analysis, the most popular techniques are mainly based on the DNA chemical modification induced by sodium bisulfite. Sodium bisulfite is able to deaminate unmethylated cytosines, turning them into uracils, while methylated cytosines are resistant to this chemical reaction, remaining cytosines. After treatment, when a given sequence is amplified and analysed, detected cytosines correspond to originally methylated nucleotides while, detected thymines in CpG sites correspond to originally unmethylated cytosines [142]. Based on this method, two popular DNA methylation genome-wide techniques are: 1) Whole Genome Bisulfite Sequencing (WGBS) and; 2) Reduced Representation Bisulfite Sequencing (RRBS):

1) WGBS measures single-cytosine methylation levels genome-wide. DNA is fragmented, amplified and sequenced with adaptors, forming a library. This library

is then bisulfite-converted, and afterwards, the DNA is amplified again and sequenced using high-throughput sequencing technologies [143]. Despite the high coverage and specificity of this technique, WGBS is also associated with a high cost and a need for great amounts of starting material.

2) In the case of RRBS, there is an enrichment in CpG-rich regions by DNA enzyme digestion that recognizes a CpG rich region, such as the CCGG recognized by MspI [144]. Then, DNA is fragmented, amplified and a fragment library is formed. After that, DNA is bisulfite-converted and sequenced. In this case, smaller amounts of DNA are required than for WGBS, however only a portion of methylated DNA is assessed.

Region-specific methylation analysis aims at assessing the methylation levels of a regulatory region of interest, such as a promoter. For this, the most used methods continue to be based on bisulfite treatment. Such techniques are, for example: 1) bisulfite sequencing and; 2) methylation-specific PCR (MSP):

1) Bisulfite sequencing allows the direct observation of the methylation of CpG sites in a small region of DNA. Primers are designed overlapping the region of interest and afterwards DNA is bisulfite-converted and PCR-amplified. PCR-product is sequenced using Sanger sequencing and methylated/unmethylated cytosines can be identified [142].

2) MSP is a more straightforward method, also based on bisulfite conversion and PCR. Bisulfite-converted DNA is amplified by two different sets of primers, one recognizing a given methylated DNA region and another for the corresponding unmethylated DNA region [145]. Depending on which set of primers originates a gel band, the methylation status of the analysed DNA region can be inferred. However, only a couple of CpGs are usually assessed. For this reason, conclusions regarding the methylation status of a large region should be taken carefully.

Beyond these, there are several other methods being adopted for assessment of DNA methylation. The choice for the best method again depends on the question being asked. Nevertheless, DNA methylation of specific genes/regions could constitute good predictive biomarkers and the correlation between this modification and GC target gene expression has been explored over the years [146, 147].

With this introduction, we wanted to highlight that besides anti-HER2 therapy, other GC targeted therapies, both approved and unapproved, remain without proper predictive biomarkers to effectively anticipate patient response to treatment. This fact, which together with the late diagnosis underlie GC's high incidence and mortality, led us to focus this thesis on two relevant molecular targets in GC, VEGFR2 and FGFR2 and their epigenetic regulation, in the search for new predictive biomarkers that may improve the stratification and overall survival of GC patients.

II. Rationale and Aims

Gastric Cancer is one of the most incident and deadliest types of cancer in the world, with the majority of patients being diagnosed at late stages of the disease. The high intra- and inter-tumour heterogeneity associated with GC gives rise to a molecularly and phenotypically complex disease that further hinders the effective treatment of these patients. Currently, first- and second-line chemotherapy are the standard of care for most patients, with the exception of two targeted therapies approved for GC treatment: the *Trastuzumab* (anti-HER2) and *Ramucirumab* (anti-VEGFR2) antibodies, given in a first-line and second-line setting, respectively. Nevertheless, only *Trastuzumab* has an associated predictive biomarker, the genetic amplification and protein overexpression of HER2. For *Ramucirumab*, and the majority of other therapies tested/in test in clinical trials, an appropriate predictive biomarker for therapy response is still lacking.

VEGFR2 is known to be one of the main regulators of tumour angiogenesis, together with its main ligand VEGFA. Although an anti-VEGFR2 therapy is approved for GC treatment (*Ramucirumab*), VEGFR2 is often lowly expressed in tumour cells and a positive correlation between its expression and therapy efficacy has yet to be found. *FGFR2* is amplified in up to 10% of GC cases and FGFR2 overexpression has been found in varying frequencies. Furthermore, the deregulation of two tissue specific isoforms, FGFR2-IIIb and FGFR2-IIIc, have been detected in different types of cancer. FGFR2-IIIb has been found to be overexpressed in GC, however a correlation between *FGFR2* gene amplification or FGFR2/FGFR2-IIIb protein overexpression and response to anti-FGFR2 therapies is still missing, although many such therapies have been widely studied in several clinical trials. This is likely due to the fact that new biomarkers other than gene amplification or protein overexpression are yet to be uncovered to efficiently predict therapy response and improve the outcome of this disease.

Our group as performed a preliminary analysis in a cohort of 47 paired GC tumour and adjacent normal samples (Cohort 1) to evaluate the promoter methylation status and copy number of several genes, including *KDR* (codifying VEGFR2) and *FGFR2*. This analysis showed that the majority of GC cases presented *KDR* promoter hypermethylation and normal *VEGFA* copy number in tumour samples in comparison with their corresponding normal samples. An increased *VEGFA* copy number was observed in 8 out of 47 GC cases, half of which displayed also *KDR* promoter hypermethylation. In parallel, we observed that the majority of GC cases displayed *FGFR2* promoter hypomethylation in tumour samples in comparison with their normal counterparts. In addition, *ESRP1*, a known regulator of

FGFR2 alternative splicing, also showed promoter hypomethylation in most tumour samples.

Given the absence of predictive biomarkers for Ramucirumab, the lack of approved anti-FGFR2 therapies and our preliminary data, with this work we aimed at exploring the potential of promoter methylation status of FGFR2, ESRP1 and KDR, as well as VEGFA copy number status as novel predictive biomarkers of therapy response in GC. If valid, these molecular events, alone or in combination, may allow for a better selection of GC patients for anti-FGFR2 and/or anti-VEGFR2/anti-VEGFA therapies creating the rational to improve patient selection for these targeted therapies.

To accomplish this general aim, we have designed two specific aims:

- **Specific Aim 1:** To investigate the co-occurrence of variations in the *KDR* promoter methylation status, genetic events affecting its ligand *VEGFA* and alterations in the RNA expression of both genes;
- **Specific Aim 2:** To investigate the co-occurrence of variations in the *FGFR2* and its splicing regulator *ESRP1* promoter methylation status, together with alterations in the RNA expression of both genes and relevant isoforms.

To address **Specific Aim 1**, we have selected 16 GC samples from Cohort 1 and designed the following tasks:

Task 1.1) Validate the preliminary data regarding *KDR* promoter methylation status obtained with RRBS, using bisulfite Sanger sequencing;

Task 1.2) Correlate *KDR* promoter methylation status with its RNA expression;

Task 1.3) Validate the preliminary data regarding *VEGFA* copy number obtained with WGS, using qRT-PCR;

Task 1.4) Correlate *VEGFA* copy number with its RNA expression;

Task 1.5) Investigate the co-occurrence of *KDR* promoter methylation, *VEGFA* copy number, *KDR* and *VEGFA* RNA expression;

Task 1.6) Expand our observations in Task 1.5 to larger GC cohorts, using TCGA data available online and a bioinformatics approach (*in collaboration with bioinformaticians from Oliveira's group*).

To address **Specific Aim 2**, we have selected 13 GC samples from Cohort 1 and designed the following tasks:

Task 2.1) Validate the preliminary data regarding *FGFR2* promoter methylation status obtained with RRBS, using bisulfite Sanger sequencing;

Task 2.2) Correlate *FGFR2* promoter methylation status with total *FGFR2* RNA expression, as well as the RNA expression of the specific isoforms FGFR2-IIIb and FGFR2-IIIc;

Task 2.3) Correlate *ESRP1* promoter methylation status with the RNA expression of *ESRP1*;

Task 2.4) Investigate the co-occurrence of *ESRP1* and *FGFR2* promoter methylation status with the RNA expression of *ESRP1*, total *FGFR2* and *FGFR2* isoforms;

Task 2.5) Expand our observations in Task 2.4 to larger GC cohorts, using TCGA data available online and a bioinformatics approach (*in collaboration with bioinformaticians from Oliveira's group*).

Task 2.6) Explore the role of ESRP1 in the expression regulation of FGFR2-IIIb expression in GC using *in vitro* models

With the described specific aims and different tasks, this project may uncover novel predictive biomarkers for anti-VEGFR2 and/or anti-FGFR2 therapies, using an approach that could be applied to other relevant targets deregulated in GC, as well as in other cancer models.

III. Materials and Methods

1. Gastric cancer samples

We had access to a cohort of 47 GC tumour and paired adjacent normal samples (Cohort 1) that had been previously analysed by WGS and RRBS for genome-wide assessment of copy number variants (CNVs) and promoter methylation status.

2. Cell culture

Four GC cell lines, MKN45, MKN74, KATO-III and IPA-220 (obtained from Ipatimup's Cell Line Bank), were cultured in RPMI 1640 Medium (Gibco) supplemented with penicillin-streptomycin (1%, Invitrogen) and fetal bovine serum (10%, Biowest). MCF10A cells (obtained from Ipatimup's Cell Line Bank) were cultured in DMEM/F12 Glutamax™ medium (Gibco), complemented with horse serum (5%, Lonza), recombinant human insulin (5 ug/mL), penicillin-streptomycin (1%, Invitrogen), hydrocortisone (500 ng/mL, Sigma-Aldrich), cholera toxin (20ng/mL, Sigma-Aldrich), and recombinant human epidermal growth factor (20 ng/mL, Sigma-Aldrich). MCF10A mesenchymal cells (MCF10A M-cells) were obtained with the addition of the Transforming Growth Factor-1 (TGF-β1, 8 ng/ml) to the normal medium for 7 days. All cell lines were kept in culture flasks at approximately 37 ºC and 5% CO_2.

3. DNA extraction and methylation analysis

MKN45, MKN74, IPA-220, KATO-III, MCF10A and MCF10A M-cells were collected for DNA extraction using the NZY Tissue gDNA Isolation Kit (MB13502, NZYtech). Extracted DNA from cell lines, as well as DNA derived from tumour and normal samples from Cohort 1 (previously extracted), was bisulfite-converted using the EpiTect Bisulfite Kit (Qiagen) for DNA methylation analysis. This kit includes a mix containing Sodium Bisulfite that is able to convert unmethylated cytosines into uracils, while methylated cytosines are protected from this chemical modification and remain cytosines. For a correct conversion of all unmethylated cytosines, DNA incubation in a high temperature/low pH condition is required. Hence, a DNA protect buffer is also included in this kit and a thermal cycling program is setup for the reaction. In summary, a mix containing 200 ng of DNA and DNase/RNase-free water (up to 40 µL), 85 µL of Bisulfite Mix and 15 µL of DNA Protect Buffer in a final volume of 140 µL is prepared for each sample and the PCR protocol depicted in Table 1 is initiated. Afterwards, DNA is purified using EpiTect wash buffers in a series of column-based

purification steps, before converted DNA is eluted in 20 µL of DNase/RNase-free water and kept at -20ºC.

Temperatures	Time
95 ºC	5'
60 ºC	25'
95 ºC	5'
60 ºC	85'
95 ºC	5'
60 ºC	175'
20 ºC	indefinite

Table 1: Thermal cycling program for bisulfite conversion using the EpiTect Bisulfite Kit (Qiagen).

After bisulfite-conversion, a region chosen to assess the promoter methylation status of *KDR* and *FGFR2* was amplified and sequenced, for selected tumour and paired normal samples from cohort 1 and cultured cell lines. In addition, the promoter methylation status of *ESRP1* was also assessed exclusively in cell lines. Designed primers and the optimized PCR program for each primer pair is described in Table 2.

Table 2: Sequence of primers forward and reverse for DNA methylation analysis of *KDR*, *FGFR2* and *ESRP1*. TA corresponds to temperature of annealing and Ta1, Ta2 and Ta3 as the differential temperatures for the Touchdown PCR program used.

Promoter Region Assessed		**Gene**			
		KDR	***FGFR2***	***ESRP1 P2***	***ESRP1 P3***
		chr4: 55991720-55991851	chr10: 123358126-123358272	chr8: 95652477-95652664	chr8: 95652644-95652906
Forward Primer 5'- 3'		GGAGTTTATTTTGGAGGAGG	GGGAGGGTAGGGTTAGAG	GGAGTGATTAGGTGGTTGG	GTTGTGTTGGTTTAGGAGTTG
Reverse Primer 5'- 3'		CACCCCAACCTCCCAAAAAC	CCCTCTCTACCAATCAAC	CAACTCCTAAACCAACACAAC	CCACTAAAAATAACTAAAAAATAC
TA	Ta1	59,5ºC	57ºC	58ºC	55ºC
	Ta2	57,5ºC	55ºC	56ºC	53ºC
	Ta3	55,5ºC	53ºC	54ºC	51ºC
Positive Control		HCT-116	HeLa	HeLa	HeLa
Negative Control		MCF7	HCT-116	MCF7	MCF7

These primers were designed taking in consideration the area assessed by RRBS (2000 bp upstream of the TSS), referred as the mathematical promoter, and the closest CpG island, giving rise to a proxy region for promoter methylation analysis of each gene (Figure 1). Furthermore, positive and negative controls for this evaluation were selected according to Encode (Table 2).

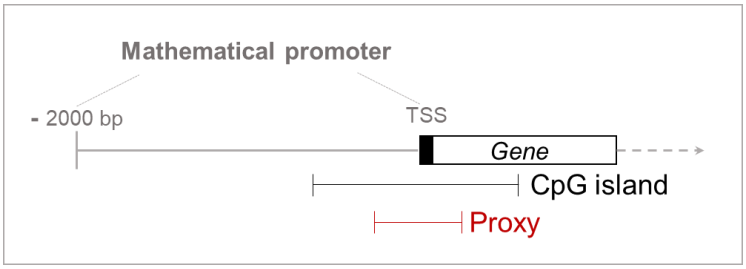


Figure 1: Representation of the region chosen in each gene to evaluate its promoter methylation status: the proxy region results from the overlapping area assessed by RRBS (mathematical promoter) and the nearest CpG island.

To amplify the chosen locus, a mix including Multiplex Mix (10 µL, Qiagen), primer forward (2 µL from a 10 µM solution), primer reverse (2 µL from a 2 µM stock solution) and Q-solution (2 µL Qiagen) was prepared for each pair of primers and multiplied by the number of samples being used. A final mix of converted DNA (1 µL) and 16 µL of the first mix was then subjected to a Touchdown PCR program (described in Table 3), optimized for DNA methylation assessment. In addition to the specific PCR program for each gene, the efficiency of each reaction results from the presence of the modified polymerase included in the Qiagen Multiplex Mix, the Hot Start Taq DNA polymerase, that reduces extension of non-specific DNA by being activated only at 95ºC. Moreover, the kit also includes the Q-solution, a signature PCR additive designed to modify the DNA melting properties and improve amplification. To confirm the correct amplification of expected DNA sequences, each PCR product was loaded in a 2% agarose gel.

Table 3: Touchdown PCR program for amplification of the chosen region for methylation assessment of each gene of interest. Ta1, Ta2 and Ta3 are annealing temperatures that change according to the pair of primers being used. The set of temperatures associated with Ta3 also varies in the number of cycles depending of the pair of primers being used.

Temperatures	Time	Cycles
95 (ºC)	15'	1x
94 (ºC) **Ta1** 72 (ºC)	30'' 1'30'' 1'30''	3x
94 (ºC) **Ta2** 72 (ºC)	1'30'' 1'30'' 1'30''	3x
94 (ºC) **Ta3** 72 (ºC)	1'30'' 1'30'' 1'30''	35-40x
72 (ºC)	1'30''	1x
12 (ºC)	1'30''	1x

After optimization of the PCR protocol and confirmation of the expected band by agarose gel, bisulfite Sanger sequencing was performed using the same pairs of primers and a different PCR protocol, described in Table 4. Before sequencing the PCR product, we purified the required amount of amplified DNA with ExoSAP-IT™ Express (Applied Biosystems), an enzyme-based reagent that removes extra primers and nucleotides that influence the quality of the sequencing reaction. To sequence, the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems™) was used in a mix prepared for each sample, multiplied by the 2 different primers. This mix contains 0.4 µL of either primer forward or reverse (from a 10 µM stock solution), 0.5 µL of BigDye, 1 µL of sequencing buffer and 2.1 µL of DNase/RNase-free water. To this mix, 1 µL of the purified post-PCR product was added.

Table 4: PCR program for Sanger sequencing of amplified products.

Temperatures	Time	Cycles
96 (ºC)	2'	1x
96 (ºC) 54 (ºC) 60 (ºC)	30'' 15'' 3'	35x
60 (ºC)	10'	1x
12 (ºC)	indefinite	1x

The resulting sequence was analysed, confirming the correct conversion of all previously unmethylated cytosines into thymines, as well as the methylation status of previously methylated cytosines.

4. Copy number variation quantification by qRT-PCR

The copy number of *VEGFA* was quantified in chosen tumour and normal samples from cohort 1 using Taqman Copy Number Assays (ThermoFisher Scientific). Two assays were run at the same time in a duplex quantitative real-time PCR (qRT-PCR) reaction: one for *VEGFA* detection (Hs03055271_cn, Applied Biosystems) and another to detect a reference gene (RNAseP, Applied Biosystems), known to present two copies in a diploid genome. A mix was prepared containing the two copy number assays, the Taqman Genotyping Master Mix (Applied Biosystems) and genomic DNA (1 ng/µL per reaction). This reaction was multiplied by three technical replicates. In addition, the same mix was prepared for a calibrator DNA sample, in which the region of interest presented a known copy number. This mix was multiplied by at least four reactions to generate a calibration curve. The *VEGFA* copy number was then calculated using the Copy Caller Software v2.0 (Applied Biosystems). This software determined the relative quantitation (RQ) using the $2^{-\Delta\Delta CT}$ method [148] by measuring the difference (ΔC_T) between the C_T values of *VEGFA* and the reference gene, as well as by comparing the ΔC_T of test samples to the ΔC_T of the calibrator sample.

5. RNA extraction and expression quantification

RNA was extracted from cultured cells using the TriPure isolation reagent (Roche) that, with the addition of chloroform, allows the simultaneous isolation of DNA, RNA and protein in a liquid-phase separation by centrifugation. The RNA-containing aqueous phase was collected and RNA was precipitated using isopropanol. After centrifugation, RNA was washed with ethanol, centrifuged again and air-dried. After elution with DNase/RNase-free water, RNA samples were kept at -80ºC.

To quantify RNA expression, a reverse transcription PCR (RT-PCR) was performed to originate cDNA in a two-step reaction. First, a mix containing 1 µL of random primers and 1 µg of RNA in a final volume of 12 µL was prepared and placed at 70ºC for 10 minutes and at 4ºC for 2 minutes, allowing the denaturation of RNA secondary structures followed by primer annealing. During the second step, an 8 µL mix for each sample was prepared, as described in Table 5. The 8 µL were added to the first mix and a final reaction was performed at 37ºC for 1h.

Table 5: RT-PCR reaction components.

Components	Volumes
Buffer 5x	4 uL
DTT (0,1 M)	2 uL
dNTPs (10 mM)	1 uL
RNAse inhibitor (40 U/uL)	0,2 uL
SSII (200 U/uL)	0,75 uL
H2O	0,75 uL

After converting RNA into cDNA, a quantitative real-time PCR (qRT-PCR) was performed to quantify RNA expression using a mix that included 5 µL of Kapa Probe Fast qPCR Master Mix (Roche), 1 µL of cDNA (~50ng/µL) and 3.5 µL of DNase/RNase-free water. RNA expression quantification was performed for *KDR*, *VEGFA*, *ESRP1*, total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* in chosen samples using PrimeTime qPCR assays specific for each gene of interest (Table 6). Furthermore, the same analysis was performed for the endogenous control *18s*, a gene expected to be expressed without any change across sample types and experimental conditions. The volumes described above were multiplied per three technical replicas and per the number of genes. To each mixture, 1.5 µL of the appropriate assay was added and 10µL of this final mix was transferred to individual wells in a 96-well plate. The qRT-PCR was then performed using an ABI Prism 7500 Sequence Detection System. The results were analysed using the comparative method $2^{(-\Delta\Delta C_T)}$ [148].

Table 6: PrimeTime qPCR predesigned assays or probe and primer sequences for custom assays used for qRT-PCR.

Gene	Predesigned Assays	Custom Assays		
		Probe	Primer 1	Primer 2
KDR	Hs.PT.58.3285240	-		
VEGFA	Hs.PT.58.1780230	-		
ESRP1	Hs.PT.58.24361486	-		
FGFR2	Hs.PT.58.1565679	-		
FGFR2-IIIb	-	AACAGCAAG/ZEN/CGC CTGGAAGAGAAA	CAATTATATAGGGCAGG CCAAC	CCCTATGCAGTAAATG GCTATC
FGFR2-IIIc	-	TCTGCATGG/ZEN/TTGA CAGTTCTGCCA	CTTGGCGGGTAATTCT ATTGG	CCCTATGCAGTAAATG GCTATC

6. 5-Aza-2′-deoxycytidine treatment

MCF10A, MKN74 and MKN45 cells were cultured in three 6-well plates each and, at ~60% of confluence, 5 µM of the demethylating agent 5-Aza-2′-deoxycytidine (5-AZA) or DMSO was added to each well. Cells were collected after 24h, 48h or 72h of incubation with

DMSO/5-AZA for DNA extraction to perform DNA methylation analysis and RNA extraction to quantify gene expression.

7. Bioinformatics Approach using TCGA Cohorts 2 and 3

TCGA-derived data for DNA methylation, CNVs and RNA expression was extracted from TCGA data portal for 2 independent GC cohorts: cohort 2, only with tumour samples and; cohort 3 with paired tumour and normal samples. In brief, DNA methylation data derived from the 450k Infinium chip, with probes located throughout the human genome; CNV data derived from Affymetrix SNP 6.0 array and; RNA expression data derived from RNA high throughput sequencing. All data extraction, analysis, production of boxplots and statistical calculations (Wilcoxon Rank-Sum Test) described in this Thesis were performed in collaboration with bioinformaticians from Oliveira's group.

IV. Results and Discussion

Current Gastric Cancer treatment is mostly based on standard chemotherapy regimens, offering limited survival improvement for GC patients. The two approved GC targeted therapies *Trastuzumab* (anti-HER2) and *Ramucirumab* (anti-VEGFR2) have improved patient survival by less than 6 months. Although *Trastuzumab* has a proven predictive biomarker, HER2 overexpression and *ERBB2*/HER2 gene amplification, the same has yet to be found for *Ramucirumab*. Other deregulated molecules in GC, such as FGFR2, have also been investigated as new potential therapeutic targets, however clinical trials have yet to show a significant improvement of survival for GC patients. One of the reasons underlying this inefficiency is the lack of predictive biomarkers that would improve the selection of patients that would benefit the most from one of these therapies being developed. Therefore, there is an urgent need to find such biomarkers in the setting of novel target therapies for GC management and treatment. Our general aim with this work was to explore the potential of promoter methylation status as a novel predictive biomarker for anti-*KDR*/VEGFR2 and anti-FGFR2 therapies, two known deregulated molecules in GC, for which there are currently no defined predictive biomarkers. If valid, these epigenetic events may allow for a better selection of GC patients for anti-FGFR2 and/or anti-VEGFR2 therapies, better therapeutic response and increased patient survival.

To address our general aim, we have designed two specific aims, tailored towards the state of the art regarding *KDR*/VEGFR2 (**Specific Aim 1, explored in Part 1**) and *FGFR2* (**Specific Aim 2, explored in Part 2**).

PART 1: KDR/VEGFR2 and VEGFA genetic, epigenetic and transcriptomic events in Gastric Cancer

1.1. Correlating *KDR* promoter methylation status with *KDR* RNA expression in gastric tumour *vs.* normal samples

A preliminary analysis assessing the promoter methylation status of several genes by RRBS, was performed for cohort 1, which encompassed 47 GC tumour and adjacent normal samples (47 GC pairs). RRBS bioinformatics analysis was previously performed focused on the upstream region of all genes, ranging from 2000 bp upstream of the TSS until the TSS. These regions were designated as “mathematical promoters”.

For each of these mathematical promoters, the number of methylated CpG sites was determined for each sample pair and a per-sample methylation ratio was calculated, *i.e.* the number of CpG sites in the mathematical promoter in the tumour sample divided by the number of CpG sites in the mathematical promoter in the paired normal sample. Obtained ratios were divided in three categories:

1) if the methylation ratio was equal to or above 1.5, this corresponded to tumours with increased methylation in the given mathematical promoter, *i.e.* hypermethylated tumours;

2) if the methylation ratio was equal to or below 0.66, this corresponded to tumours with decreased methylation in the given mathematical promoter, *i.e.* hypomethylated tumours;

3) if the methylation ratio was above 0.66 however below 1.5, this corresponded to tumours without variation in methylation in the given mathematical promoter.

To validate the obtained RRBS results, we have selected a small region for the *KDR* gene, to evaluate promoter methylation using Bisulfite Sanger sequencing in a selection of cases from cohort 1. This region, referred as proxy, was selected by overlapping the mathematical promoter of *KDR* with the nearest annotated CpG island (Figure 1). The proxy was amplified and sequenced using specific primers and the peak of each CpG dinucleotide was carefully analysed for DNA methylation. All previously unmethylated cytosines were confirmed to be converted to thymines, verifying the correct bisulfite conversion of the DNA. Next, all CpGs were annotated for their methylation status and classified as:

- Methylated, if only a cytosine peak appeared at the C site of the CpG dinucleotide;
- Unmethylated, if only a thymine peak appeared at the C site of the CpG dinucleotide;
- Hemimethylated, if both cytosine and thymine peaks appeared at the C site of the CpG dinucleotide (Figure 2).

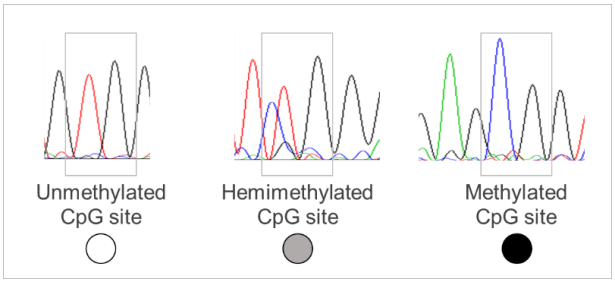


Figure 2: Example of unmethylated, hemimethylated and methylated CpG sites.

Moreover, the CpG methylation patterns of tumour and corresponding normal samples were compared and GC cases were classified as hypermethylated, hypomethylated or without variation for *KDR* promoter methylation (Figure 3 and Annex I).

The preliminary DNA methylation analysis performed by RRBS showed that the promoter of *KDR* was hypermethylated in 62% (*n*=29), hypomethylated in 13% (*n*=6) and without variation in methylation in 25% (*n*=12) of GC tumour samples in comparison with the corresponding normal samples. From these Cohort 1 cases, 16 pairs of tumour/normal samples were selected for promoter methylation status evaluation (task 1.1) using a different technique, in particular bisulfite Sanger sequencing (Table 7). We validated, according to RRBS, the following cases:

- 9 cases with *KDR* promoter hypermethylation;
- 5 with *KDR* promoter hypomethylation, and;
- 2 without methylation variation in the *KDR* promoter.

From the 9 hypermethylated cases, 8 were also hypermethylated by Sanger sequencing, while it was not possible to assess the methylation status of the remaining case. From the 5 hypomethylated cases, 4 also displayed *KDR* promoter hypomethylation, while 1 case presented no methylation variation in the assessed region. Lastly, the 2 cases without methylation variation by RRBS, also showed no variation in *KDR* promoter methylation.

Table 7: Comparison between the obtained *KDR* promoter methylation status by RRBS and Sanger sequencing in the 16 chosen tumour samples (T) *vs.* the corresponding normal samples (N) from Cohort 1.

T/N samples	*KDR* methylation status	
	RRBS	Sanger seq.
T1/N1	4,82	●
T2/N2	2,39	●
T3/N3	7,67	●
T4/N4	6,77	●
T5/N5	2,51	●
T6/N6	8,22	●
T7/N7	9,92	●
T8/N8	1,73	●
T9/N9	2,28	*NA*
T10/N10	0,59	**NV**
T11/N11	0,41	○
T12/N12	0,13	○
T13/N13	0,46	○
T14/N14	0,25	○
T15/N15	1,17	**NV**
T16/N16	1,19	**NV**

Footnote: 1) RRBS methylation ratio classification: ≥1.5, hypermethylation; ≤0.66, hypomethylation; >0.66 and <1.5, without variation. 2) Sanger sequencing classification: black circle, hypermethylation; white circle, hypomethylation; NV, no variation; NA, not available.

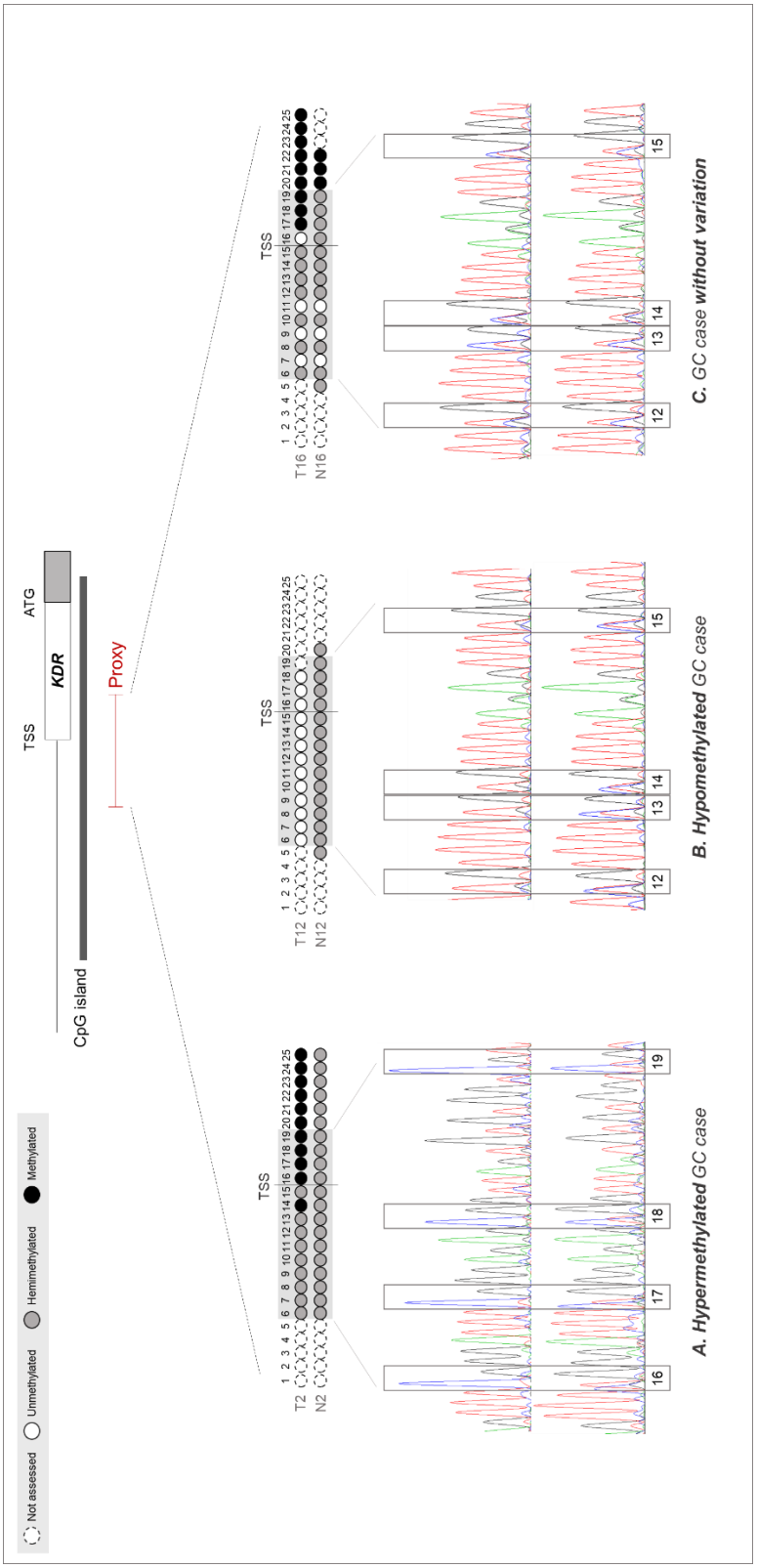


Figure 3: Example of the methylation patterns obtained at the *KDR* promoter proxy region by bisulfite Sanger sequencing in the analysed 16 GC cases from Cohort 1, with each CpG dinucleotide classified as unmethylated, methylated, hemimethylated or not assessed if the detected peak corresponded to a thymine, a cytosine, both or none was detected, respectively. **A)** Example of GC case in which the tumour sample (T2) was classified as hypermethylated in comparison with its paired normal sample (N2). **B)** Example of a GC case in which the tumour sample (T12) was classified as hypomethylated in comparison with its paired normal sample (N12). **C)** Example of GC case in which the tumour sample (T16) was classified as without variation in methylation in comparison with its paired normal sample (N16).

For a total of 14 out of 16 GC cases the results obtained with the RRBS technique were validated by Bisulfite Sanger sequencing for the KDR promoter methylation status. In addition, the proxy region chosen to validate the results from the RRBS was shown to be appropriate to infer the methylation status of the KDR promoter.

Next, *KDR* RNA expression was assessed by qRT-PCR in the same 16 GC cases and the expression level between the tumour and the normal counterpart was compared (Task 1.2, Table 8). Of notice, it was only possible to obtain RNA expression values for 10 GC cases, since the remaining 6 GC cases presented low RNA quality.

Of the 10 GC pairs with available RNA expression data:

- 4 revealed increased *KDR* expression in the tumour samples *vs.* the corresponding normal samples
- 5 showed similar *KDR* expression in the tumour samples *vs.* the corresponding normal samples
- 1 had decreased *KDR* expression in the tumour *vs.* the corresponding normal samples

From the 4 GC cases with increased RNA expression, all displayed *KDR* promoter hypermethylation (assessed by RRBS and Sanger sequencing, or just by RRBS), a scenario opposite to the "dogma" in which a methylated promoter often results in a transcriptionally repressed gene and a unmethylated promoter is transcriptionally active. These results may reflect the existence of a transcriptional repressor, *i.e.* the increase of methylation at the *KDR* promoter could be obstructing a binding site of a repressor that can be bound to a demethylated form of the *KDR* promoter. By impeding the binding of a repressor, the increase of *KDR* promoter methylation in tumours may underlie the observed increased RNA expression in comparison with the corresponding normal samples. To strengthen this hypothesis, the opposite scenario should also be detected in our cohort. Indeed, *KDR* promoter hypomethylation correlated with decreased *KDR* RNA expression, but only in 1 GC case.

Table 8: Concomitance of the *KDR* promoter methylation status found in each GC pair by RRBS or Bisulfite Sanger sequencing and the *KDR* RNA expression detected by qRT-PCR.

T/N samples	*KDR* methylation		*KDR* RNA expression
	RRBS	Sanger	
T1 / N1	●	●	2,75x
T2 / N2	●	●	5,35x
T3 / N3	●	●	1,04x
T4 / N4	●	●	*NA*
T5 / N5	●	●	*NA*
T6 / N6	●	●	*NA*
T7 / N7	●	●	*NA*
T8 / N8	●	●	2,49x
T9 / N9	●	*NA*	3,61x
T10 / N10	○	**NV**	*NA*
T11 / N11	○	○	0,49x
T12 / N12	○	○	1,1x
T13 / N13	○	○	1,25x
T14 / N14	○	○	1,01x
T15 / N15	**NV**	**NV**	0,94x
T16 / N16	**NV**	**NV**	*NA*

Footnote: 1) *KDR* methylation: black circle, hypermethylation; white circle, hypomethylation; NV, no variation; NA, not available. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

In summary, we observed that in 5/10 GC cases the methylation status of KDR was directly correlated with KDR RNA expression, i.e. KDR promoter increased methylation was accompanied by increased RNA expression in the tumour or KDR promoter decreased methylation was accompanied by decreased RNA expression in the tumour. These data support the hypothesis that KDR promoter hypermethylation is associated to KDR increased RNA expression.

1.2. Correlating *VEGFA* copy number with *VEGFA* RNA expression in gastric tumour *vs.* normal samples

A preliminary analysis was performed for the 47 GC tumour and adjacent normal samples mentioned above using WGS, in which the *VEGFA* copy number was assessed. Being the main VEGFR2 ligand, VEGFA is also an important factor to consider when targeting VEGFR2 signalling, as it could be adding complexity and variability to therapy response. It could be expected that targeting VEGFA in tumours with *KDR*/VEGFR2 downregulation would be a less efficient strategy than in tumours with *KDR*/VEGFR2 upregulation.

Hence, the copy number of *VEGFA* was determined for every sample by WGS, and a ratio between the values obtained for the tumour and normal samples of each GC pair was calculated.

> The obtained ratios were classified as follows:
>
> 1) If the copy number ratio was below or equal to 0.5, this corresponded to tumours with decreased *VEGFA* copy number in comparison with the corresponding normal samples, *i.e.* tumours with a *VEGFA* deletion;
>
> 2) If the copy number ratio was equal or above 1.5, this corresponded to tumours with increased *VEGFA* copy number in comparison with the corresponding normal samples *i.e.* tumours with a *VEGFA* amplification;
>
> 3) If the copy number ratio was above 0.5 and below 1.5, this corresponded to tumours with no change in *VEGFA* copy number in comparison with the corresponding normal samples.

WGS revealed that 8 cases from Cohort 1 (17%) displayed *VEGFA* amplification in the tumour sample, while the remaining cases (*n*=39) showed normal *VEGFA* copy number. To validate these results, qRT-PCR was performed for copy number quantification (Task 1.3) in the same 16 GC cases evaluated for *KDR* promoter methylation and RNA expression, which also included the previously mentioned 8 cases with increased *VEGFA* copy number. Using the Copy Caller software, the copy number for *VEGFA* was obtained from the qRT-PCR results and a ratio between the copy number found in tumour samples in comparison with its normal counterpart was calculated. The classification of the obtained ratios for each GC pair was performed using the same criteria mentioned above for the ratios obtained from WGS.

From the 8 cases that had shown increased *VEGFA* copy number by WGS, 5 also displayed further copies of the *VEGFA* gene in tumour samples when compared to their corresponding normal samples by qRT-PCR (Table 9). The remaining 3 GC cases presented normal copy number by qRT-PCR. Of the remaining GC cases that displayed normal copy number by WGS, 7 were concordant by qRT-PCR, while 1 revealed *VEGFA* amplification.

Table 9: Comparison between the *VEGFA* copy number obtained by WGS and that detected by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N). In blue are indicated the cases with *VEGFA* amplification.

T/N samples	*VEGFA* Copy Number	
	WGS	qRT-PCR
T1 / N1	**8,90**	**7,62**
T2 / N2	**1,54**	**2,79**
T3 / N3	0,48	0,79
T4 / N4	1,08	1,28
T5 / N5	**1,59**	0,85
T6 / N6	0,97	1,11
T7 / N7	1,00	0,90
T8 / N8	0,99	0,88
T9 / N9	**1,99**	**3,28**
T10 / N10	**1,59**	0,97
T11 / N11	1,45	**1,63**
T12 / N12	**2,02**	**1,82**
T13 / N13	1,00	1,02
T14 / N14	1,33	1,21
T15 / N15	**2,43**	**2,73**
T16 / N16	**2,97**	0,79

Footnote: *VEGFA* copy number: ≥1.5, amplification; ≤0.5, deletion; >0.5/<1.5, normal copy number.

Overall, 12/16 GC cases presented the same VEGFA copy number by WGS and qRT-PCR. From the remaining 4 GC cases, 3 were classified as VEGFA amplified by WGS however presented normal copy number by qRT-PCR, while 1 presented the opposite scenario. In the later non-validated cases, most values obtained either by WGS or qRT-PCR were borderline and therefore their classification was not consensual.

Next, *VEGFA* RNA expression was quantified in the same samples by qRT-PCR (Task 1.4, Table 10). Of notice, in 7 GC cases the *VEGFA* RNA expression was not possible to quantify due to low RNA quality. Of the remaining 9 GC cases, 7 presented increased *VEGFA* RNA expression (1 borderline increase), while 2 showed decreased *VEGFA* RNA expression. As we associated the copy number status of *VEGFA* with its RNA expression, we observed that:

- from the 5 GC cases with *VEGFA* amplification by both WGS and qRT-PCR, 4 also displayed increased *VEGFA* RNA expression in tumour samples in comparison with their normal counterparts; the remaining GC case presented borderline increased *VEGFA* RNA expression (1,45x);
- from the 3 GC cases with *VEGFA* amplification exclusively by WGS, RNA data was only available for 2 of them, with one displaying increased *VEGFA* RNA expression

and the other decreased *VEGFA* RNA expression in the tumour samples *vs.* the normal counterparts;

- the GC case with *VEGFA* amplification exclusively by qRT-PCR also displayed increased *VEGFA* RNA expression in the tumour samples *vs.* the normal counterparts;
- from the 7 GC cases with normal *VEGFA* copy number by both WGS and qRT-PCR, RNA data was only available for 1 of them, which presented decreased *VEGFA* RNA expression in the tumour sample *vs.* the normal counterpart.

Table 10: Concomitance of *VEGFA* copy number obtained by WGS and qRT-PCR and *VEGFA* RNA expression quantified by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N). In blue, are indicated the cases with *VEGFA* amplification and in white the cases with normal VEGFA copy number.

T/N samples	*VEGFA* Copy Number		*VEGFA* RNA expression
	WGS	qRT-PCR	
T1 / N1			5,24x
T2 / N2			1,45x
T3 / N3			*NA*
T4 / N4			*NA*
T5 / N5			*NA*
T6 / N6			0,14x
T7 / N7			*NA*
T8 / N8			*NA*
T9 / N9			1,71x
T10 / N10			0,16x
T11 / N11			3,07x
T12 / N12			3,12x
T13 / N13			*NA*
T14 / N14			*NA*
T15 / N15			8,32x
T16 / N16			2,15x

Footnote: 1) *VEGFA* copy number: ≥1.5, amplification; ≤0.5, deletion; >0.5/<1.5, normal copy number. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

In summary, in 7/16 GC cases with evidences of VEGFA amplification either by WGS or qRT-PCR a concomitant increased VEGFA RNA expression was observed in tumour samples in comparison with their normal counterparts. For the remaining 9 GC cases no apparent correlation between VEGFA copy number and RNA expression was observed. Overall, these data support the hypothesis that VEGFA amplification leads to VEGFA increased RNA expression.

1.3. Correlating *KDR* promoter methylation status and RNA expression with *VEGFA* copy number and RNA expression in gastric tumour *vs.* normal samples

VEGFA and VEGFR2 are known to actively cooperate during tumour angiogenesis [30], hence it is important to understand the expression status of both proteins when establishing therapeutic regimens, particularly for targeted therapies. Furthermore, as the expression of these proteins may be regulated by epigenetic events, we next explored our Cohort 1 cases for the status of the previously evaluated features in concomitance: *KDR*/VEGFR2 promoter methylation; *KDR*/VEGFR2 RNA expression; *VEGFA* copy number and; *VEGFA* RNA expression (Table 11).

Table 11: Concomitance of the *KDR* promoter methylation status evaluated either by RRBS or Bisulfite Sanger sequencing, *VEGFA* copy number assessed by WGS or qRT-PCR and RNA expression of *KDR* and *VEGFA* quantified by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N).

T/N samples	*KDR* methylation		*KDR* RNA expression	*VEGFA* Copy Number		*VEGFA* RNA expression
	RRBS	Sanger		WGS	qRT-PCR	
T1 / N1	●	●	↑			↑
T2 / N2	●	●	↑			=
T3 / N3	●	●	=			*NA*
T4 / N4	●	●	*NA*			*NA*
T5 / N5	●	●	*NA*			*NA*
T6 / N6	●	●	*NA*			↓
T7 / N7	●	●	*NA*			*NA*
T8 / N8	●	●	↑			*NA*
T9 / N9	●	*NA*	↑			↑
T10 / N10	○	**NV**	*NA*			↓
T11 / N11	○	○	↓			↑
T12 / N12	○	○	=			↑
T13 / N13	○	○	=			*NA*
T14 / N14	○	○	=			*NA*
T15 / N15	**NV**	**NV**	=			↑
T16 / N16	**NV**	**NV**	*NA*			↑

Footnote: 1) *KDR* methylation: black circle, hypermethylation; white circle, hypomethylation; NV, no variation; NA, not available. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available. 3) *VEGFA* copy number: ≥1.5, amplification; ≤0.5, deletion; >0.5/<1.5, normal copy number.

From the 16 GC cases analysed, 5 presented data for all evaluated features (T1/N1, T2/N2, T11/N11, T12/N12 and T15/N15):

- 2 cases (T1/N1 and T2/N2) presented *KDR* promoter hypermethylation, *VEGFA* amplification, increased *KDR* RNA expression and increased/equal *VEGFA* RNA expression in tumour samples *vs.* normal counterparts;

- 2 cases (T11/N11 and T12/N12) presented *KDR* promoter hypomethylation, *VEGFA* amplification (qRT-PCR), decreased/equal *KDR* RNA expression and increased *VEGFA* RNA expression in tumour samples *vs.* normal counterparts;
- 1 case (T15/N15) presented no variation in *KDR* promoter methylation, *VEGFA* amplification, no variation in *KDR* RNA expression and increased *VEGFA* RNA expression in the tumour sample *vs.* normal counterpart.

From the remaining 11 GC cases without available data for 1 or more of the evaluated features:

- 1 case (T9/N9) displayed *KDR* promoter hypermethylation (RRBS), *VEGFA* amplification and increased *KDR* and *VEGFA* RNA expression in tumour samples *vs.* normal counterparts;
- 2 cases (T13/N13 and T14/N14) displayed *KDR* promoter hypomethylation, *VEGFA* normal copy number and no variation in *KDR* RNA expression in tumour samples *vs*. normal counterparts;
- 2 cases (T3/N3 and T8/N8) displayed *KDR* promoter hypermethylation, *VEGFA* normal copy number and increased/equal *KDR* RNA expression in tumour samples *vs*. normal counterparts;
- 2 cases (T4/N4 and T7/N7) displayed *KDR* promoter hypermethylation, normal *VEGFA* copy number and no information regarding *KDR* or *VEGFA* RNA expression in tumour samples *vs.* normal counterparts;
- The remaining 4 cases (T5/N5, T6/N6, T10/N10 and T16/N16) displayed multiple scenarios for the evaluated features.

Overall, our results revealed 2 main scenarios that could be relevant in the context of GC patient treatment with anti-*KDR*/VEGFR2 and anti-VEGFA therapies:

- **Scenario 1**: GC cases with *KDR* promoter hypermethylation and increased *KDR* RNA expression, accompanied by *VEGFA* amplification and increased/equal *VEGFA* RNA expression;
- **Scenario 2**: GC cases with *KDR* promoter hypomethylation and equal or decreased *KDR* RNA expression, accompanied by *VEGFA* amplification or normal copy number and independently of *VEGFA* RNA expression levels;

Of all possible scenarios encompassing the 4 evaluated features, **scenarios 1 and 2** are the most represented in our small cohort and would be the most pertinent to select (scenario 1) or not (scenario 2) GC patients for anti-VEGFR2 and/or anti-VEGFA therapies. For **scenario 1**, we could hypothesize that patients would present the higher benefit in receiving anti-VEGFR2/anti-VEGFA therapies, given that both *KDR* and *VEGFA* were

overexpressed. For **scenario 2**, we could hypothesize that patients would not benefit from neither anti-VEGFR2 nor anti-VEGFA therapies. Given that tumours presented similar *KDR* RNA expression levels in comparison with their normal counterparts, targeting this receptor in tumour cells would not constitute any benefit. Moreover, targeting VEGFA would also be inefficient, given that its main receptor *KDR*/VEGFR2 is expressed at similar levels in tumour cells as their normal counterparts. To expand these observations, we evaluated the same features in larger TCGA cohorts.

1.4. Correlating *KDR* promoter methylation with *VEGFA* copy number and *VEGFA* and *KDR* RNA expression in other GC cohorts

To validate the observations made for the 16 GC pairs from Cohort 1, we evaluated the same four features, *i.e. KDR* promoter methylation status, *VEGFA* copy number and the RNA expression of *KDR* and *VEGFA*, in two TCGA datasets, referred as Cohort 2 and Cohort 3 (task 1.6). Cohort 2 encompassed 310 gastric adenocarcinomas with data for all 4 evaluated features. Cohort 3 encompassed 25 paired gastric tumour and normal samples, however, no data was available for normal samples for the methylation status of *KDR*. Therefore, gastric tumours from Cohort 3 were analysed together with Cohort 2, totalling 335 gastric tumours.

Cohorts 2 and 3 methylation data derived from the 450k Infinium chip, with probes located throughout the human genome. To address *KDR* promoter methylation, first the normalized intensity of probes located at the *KDR* mathematical promoter, at the nearest CpG island and at the proxy region assessed by Sanger sequencing was collected. The normalized intensity was made available by the TCGA project in the form of beta-values, which range from 0 for demethylated DNA to 1 for methylated DNA. Next, the obtained beta-values for each assessed region was compared. We observed that:

1) the region corresponding to the *KDR* mathematical promoter displayed significantly different methylation levels when compared to the CpG island and the proxy region (data not shown);

2) the CpG island and proxy regions presented similar levels of methylation (data not shown).

Therefore, for further analyses we would consider the methylation of the proxy region to validate the observations made for Cohort 1. However, given that Cohort 2 did not encompass normal samples and that no methylation data was available for normal samples

from Cohort 3, *KDR* methylation status was only evaluated in tumours arising from both cohorts, being classified in three categories:

- If the beta-value was lower or equal to 0.33, the tumour was classified as demethylated;
- If the beta-value was equal or above 0.66, the tumour was classified as methylated;
- If the beta-value was between 0.33 and 0.66, the tumour was classified as hemimethylated.

Considering the proxy region, we observed that:

- 43% of tumour samples presented *KDR* promoter demethylation (*n*=144/335, Table 12);
- 49% of tumour samples presented *KDR* promoter hemimethylation (*n*=164/335, Table 12) and;
- 8% of tumour samples presented *KDR* promoter methylation (*n*=27/335, Table 12).

Table 12: *KDR* promoter methylation status detected at its mathematical promoter, CpG island and proxy region in the tumour samples from Cohorts 2 and 3.

Number of Tumour Samples *(Cohorts 2/3, n=335)*		
Analysed Region	***KDR* Promoter Methylation Status**	
Mathematical Promoter	Demethylated	182
	Hemimethylated	131
	Methylated	22
CpG Island	Demethylated	151
	Hemimethylated	163
	Methylated	21
Proxy	Demethylated	144
	Hemimethylated	164
	Methylated	27

Footnote: Tumours with *KDR* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Of notice, almost half of all tumour samples from Cohorts 2/3 presented intermediate levels of *KDR* promoter methylation (49% of hemimethylation), which we cannot safely consider as being demethylated or methylated, given the lack of normal tissue information. Therefore, we cannot confidently compare the previously obtained *KDR* promoter

hypermethylation in GC cases from Cohort 1 with the promoter methylation/hemimethylation found in tumour samples from Cohorts 2/3.

Nevertheless, we were able to assess *KDR* RNA expression in tumour samples from Cohorts 2/3 and normal samples from Cohort 3. Both *KDR* and *VEGFA* RNA expression data in Cohorts 2 and 3 was generated via RNA sequencing experiments and retrieved from the TCGA website in FPKM/RPKM units [149-151]. RNA expression data from Cohorts 2 and 3 tumours was analysed all together, as well as categorized as upregulated or downregulated using as a cut-off the median RNA expression obtained for the 25 normal samples within Cohort 3. We observed that *KDR* RNA expression was significantly increased in tumour samples from Cohorts 2/3 in comparison with the 25 normal samples from Cohort 3 (Figure 4).

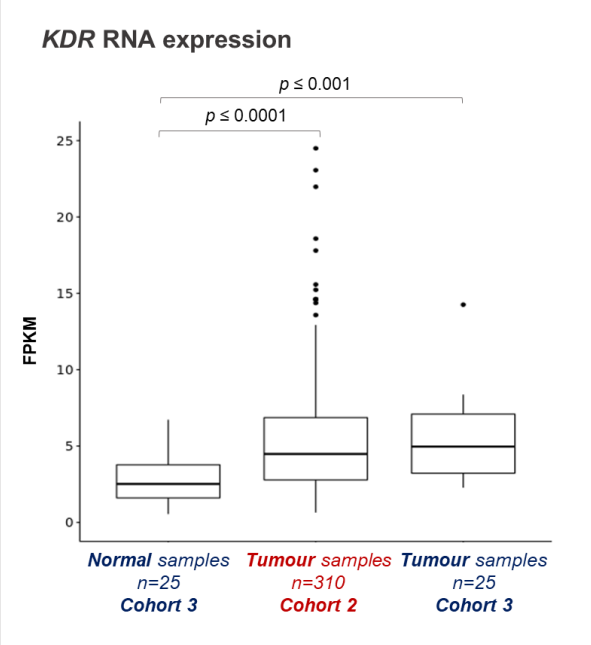


Figure 4: Comparison between the *KDR* RNA expression levels detected in tumour samples from Cohort 2 or Cohort 3 *vs.* that obtained in the normal samples from Cohort 3.

This is in accordance with the increased *KDR* RNA expression observed in the analysed 16 GC cases from Cohort 1 (Table 8). However, when correlating *KDR* promoter methylation status with *KDR* RNA expression, no significant correlation was found between *KDR* methylation levels in all three analysed regions and its level of RNA expression for both Cohorts (data not shown). In fact, as we categorized *KDR* RNA expression levels into up and downregulated, we observed that regardless of the methylation level found at its promoter, *KDR* was upregulated in most tumour samples (Table 13). For the methylation status found at the *KDR* proxy region, we observed that:

- 86% of tumour samples with *KDR* promoter demethylation (*n*=124/144, Table 13) presented also upregulation of *KDR* RNA expression;
- 81% of tumour samples with *KDR* promoter hemimethylation (*n*=133/164, Table 13) presented also upregulation of *KDR* RNA expression and;
- 48% of tumour samples with *KDR* promoter methylation (*n*=13/27, Table 13) presented also *KDR* RNA expression upregulation.

Table 13: *KDR* promoter methylation status detected at its mathematical promoter, CpG island and proxy region in the tumour samples from Cohorts 2 and 3. *KDR* RNA expression in the same tumour samples is also represented regardless of the RNA expression categories or classified as upregulated or downregulated, depending if the expression value found in the tumour samples is above or below, respectively, of the median RNA expression value detected at normal samples from Cohort 3.

Analysed Region	**Promoter Methylation Status**	**Number of Tumour Samples** *(Cohorts 2/3, n=335)*		
		Regardless of RNA Expression Categories	***Taking into account RNA Expression Categories***	
			Upregulated	**Downregulated**
Mathematical Promoter	Demethylated	182	152	30
	Hemimethylated	131	108	23
	Methylated	22	10	12
CpG Island	Demethylated	151	131	20
	Hemimethylated	163	129	34
	Methylated	21	10	11
Proxy	Demethylated	144	124	20
	Hemimethylated	164	133	31
	Methylated	27	13	14

Footnote: Tumours with *KDR* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Therefore, the results from Cohorts 2/3 do not support those obtained with Cohort 1 regarding *KDR* promoter methylation status and *KDR* RNA expression, a fact that could be due to the differences between the methods used for methylation and RNA expression assessment in these cohorts. In particular:

- In Cohort 1, constituted by tumour and adjacent normal samples, *KDR* promoter methylation was assessed by RRBS and validated by bisulfite Sanger sequencing, while RNA expression was evaluated by qRT-PCR;
- In Cohorts 2 and 3, the *KDR* promoter methylation status was obtained using the 450k Infinium chip and only for tumour samples, leading to the sample categorization in 3 methylation levels; moreover, the RNA expression was solely classified as upregulated or downregulated using the median of the values obtained for the small subset of 25 normal samples of Cohort 3.

Hence, the different technological approaches used for data generation may play a role in the lack of reproducibility of Cohort 1 results, with different levels of *KDR* promoter methylation found in gastric tumour samples associated with varying levels of *KDR* RNA expression**. Nevertheless, 81% of tumour samples from Cohorts 2/3 revealed upregulation of *KDR* RNA expression, regardless of *KDR* promoter methylation status, unlike Cohort 1** (Table 14).

Table 14: *KDR* RNA expression in the 335 tumour samples from Cohorts 2/3, classified as upregulated or downregulated, depending if the expression value found in the tumour samples is above or below, respectively, of the median RNA expression value detected at normal samples from Cohort 3.

***KDR* RNA expression** *(Cohorts 2/3, n=335)*		
RNA Expression Categories	**Cohort 2/3 Tumour Samples**	**Cohort 3 Normal Samples**
Upregulated	270	13
Downregulated	65	12

Most studies thus far have been focused on VEGFR2 protein expression in GC, rather than *KDR* RNA expression and have failed to establish a correlation with anti-VEGFR2 therapy response. *Given our results, KDR RNA expression could be in fact the missing predictive biomarker for an efficient response to Ramucirumab in GC patients.*

Next, Cohorts 2 and 3 *VEGFA* copy number data was generated via Affymetrix SNP 6.0 array and retrieved from the TCGA website as previously computed segment mean values overlapping the *VEGFA locus* and categorized similarly to described in [152]:

- If the segment mean value was below or equal to -0.1, the tumour was classified as *VEGFA* deleted;
- If the segment mean value was equal or above +0.1, the tumour was classified as *VEGFA* amplified;
- If the segment mean value was between -0.1 and +0.1, the tumour was classified as with normal *VEGFA* copy number.

This analysis showed that 33% (*n*=111/335) of the tumour samples from Cohorts 2/3 presented *VEGFA* amplification, while 9% (*n*=31/335) displayed *VEGFA* deletion and 58% (*n*=193/335) showed a normal *VEGFA* copy number. To correlate this *VEGFA* copy number status with its level of RNA expression, *VEGFA* RNA expression data was retrieved from the TCGA website as previously mentioned. We observed that *VEGFA* RNA expression

was significantly increased in tumour samples from Cohorts 2/3 in comparison with the 25 normal samples from Cohort 3 (Figure 5).

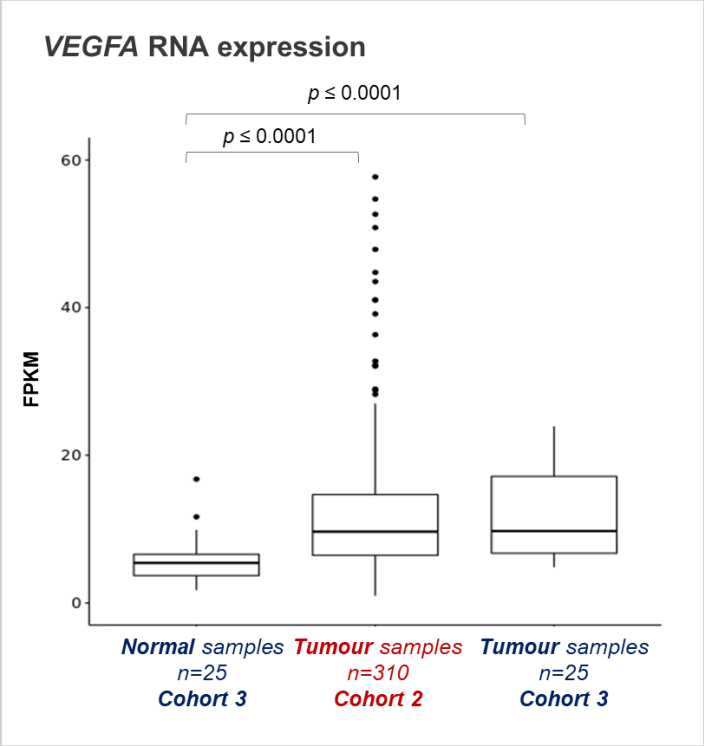


Figure 5: Comparison between the *VEGFA* RNA expression levels detected in tumour samples from Cohort 2 or Cohort 3 *vs.* that obtained in the normal samples from Cohort 3.

In particular, 82% of the tumour samples (*n*=276) were upregulated in comparison with the median expression of Cohort 3 normal samples (Table 15). As we correlated *VEGFA* copy number status and RNA expression, overall no correlation was found (data not shown).

Table 15: *VEGFA* RNA expression in the 335 tumour samples from Cohorts 2/3, classified as upregulated or downregulated, depending if the expression value found in the tumour samples is above or below, respectively, of the median RNA expression value detected at normal samples from Cohort 3.

***VEGFA* RNA expression** *(Cohorts 2/3, n=335)*		
RNA Expression Categories	**Cohort 2/3 Tumour Samples**	**Cohort 3 Normal Samples**
Upregulated	276	13
Downregulated	59	12

In fact, as we categorized *VEGFA* RNA expression levels into up and downregulated, we observed that 94% (*n*=104/111), 76% (*n*=146/193) and 84% (*n*=26/31) of *VEGFA* amplified,

normal copy number and deleted tumours presented upregulated *VEGFA* RNA expression, respectively (Table 16).

Table 16: *VEGFA* RNA expression observed in tumours from Cohorts 2/3 with *VEGFA* amplification, normal copy number or deletion. RNA expression is either depicted regardless of the RNA expression categories or divided in the upregulated or downregulated categories, *i.e.* tumours with a *VEGFA* RNA expression level above or below the median level of the RNA expression found in the normal samples from Cohort 3 are considered as upregulated or downregulated, respectively.

Gene	Copy Number Status	Number of Tumour Samples *(Cohorts 2/3, n=335)*		
		Regardless of RNA Expression Categories	***Taking into account RNA Expression Categories***	
			Upregulated	**Downregulated**
VEGFA	Amplified	111	104	7
	Normal	193	146	47
	Deleted	31	26	5

Therefore, these results challenge *VEGFA* genetic amplification as an activation mechanism, which may imply that different molecular mechanisms regulate *VEGFA* expression in GC. Nevertheless, the widespread observation that *VEGFA* RNA was upregulated in tumour samples, mimics that of *KDR*, **again suggesting that *VEGFA* RNA expression could be a valid predictive biomarker for anti-VEGFA therapies**, rather than its protein expression or genetic amplification. Moreover, this observation was also valid in Cohort 1, given that from the 9 GC pairs with *VEGFA* RNA information, 6 presented increased expression in the tumour samples *vs.* paired normal samples (Table 10). This is even more striking given the approach differences between Cohort 1 and Cohorts 2/3:

1) Small number of normal gastric tissue samples used to define the up/downregulation of *VEGFA* in tumours from Cohorts 2/3;

2) Lack of paired normal samples for Cohort 2;

3) Different technologies for RNA evaluation, with RNA-sequencing for Cohorts 2/3 and qRT-PCR for Cohort 1.

Only with access to larger paired GC cohorts could this observation be further and more confidently investigated.

Recalling Scenarios 1 and 2 observed in Cohort 1 cases, we searched for the same co-ocurrences between *KDR* RNA expression levels, and *VEGFA* RNA expression in Cohorts 2/3, and we excluded *KDR* promoter methylation status and *VEGFA* copy number status as reproducible mechanisms between cohorts. According to scenario 1, *KDR* and *VEGFA*

were both overexpressed, a co-ocurrence observed in 67% (*n*=225/335) of tumour samples from cohorts 2/3 (Table 17). Therefore, and according to our hypothesis, this feature could be valuable for the stratification of patients eligible for anti-VEGFR2/VEGFA therapies. According to scenario 2, *KDR* RNA expression was downregulated while *VEGFA* RNA expression was either up or downregulated, a co-ocurrence observed in 19% of tumour samples from Cohorts 2/3 (*n*=51/335 plus *n*=14/335, as depicted in Table 17). Hence, and according to our hypothesis, this feature could be valuable as a negative predictor for the application of anti -VEGFR2/VEGFA therapies.

Table 17: Concomitance of *KDR* and *VEGFA* RNA expression in the 335 tumour samples from Cohorts 2/3. RNA expression is classified as upregulated or downregulated, depending if the expression value found in the tumour samples is above or below, respectively, of the median RNA expression value detected at normal samples from Cohort 3.

Number of Tumour Samples *(Cohorts 2/3, n=335)*			
KDR RNA expression		***VEGFA RNA Expression***	
		Upregulated	**Downregulated**
Upregulated	270	**225**	45
Downregulated	65	**51**	**14**

These observations support the relevance of further studies on the importance of *KDR* and *VEGFA* RNA expression as putative predictive biomarkers. In fact, several examples already exist strengthening the relevance of RNA expression of certain genes as predictive biomarkers. For example, the *MammaPrint* is a prognostic and predictive test that relies on the analysis of the RNA expression of 70 genes in the context of breast cancer [153]. Therefore, one may hypothesize that *KDR* and *VEGFA* RNA expression could represent the missing predictive biomarkers for anti-*KDR*/VEGFR2 and anti-VEGFA therapy application. For this, we believe that our results support further studies using different GC cohorts with paired normal samples and with available clinical data.

PART 2: FGFR2 and ESRP1 epigenetic and transcriptomic events in Gastric Cancer

2.1. Correlating *FGFR2* promoter methylation status with total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression in tumour *vs.* normal samples

The same preliminary analysis assessing *KDR* promoter methylation levels in Cohort 1 by RRBS was performed for the promoters of *FGFR2 and ESRP1*. The RRBS results were only confirmed for a small region of the *FGFR2* promoter, using bisulfite Sanger sequencing to evaluate the promoter methylation levels in 13 selected cases from Cohort 1. The "proxy" region to evaluate was selected using the same criteria as previously mentioned: an area overlapping both the mathematical promoter of *FGFR2* and the nearest annotated CpG island (Figure 1). After amplifying and sequencing this "proxy" region of the *FGFR2* promoter in the selected 13 GC pairs, tumours were classified as being hypermethylated, hypomethylated or without variation in methylation in comparison with their normal counterparts (Figure 6 and Annex II).

The preliminary analysis by RRBS showed that 83% (*n*=39/47) of GC cases displayed *FGFR2* promoter hypomethylation in tumours in comparison with their corresponding normal samples, while the remaining 17% (*n*=8/47) presented no variation in methylation. In addition, 79% (*n*=37/47) of cases showed *ESRP1* promoter hypomethylation, while 21% (*n*=10/47) showed no methylation variation at the *FGFR2* promoter. Moreover, we observed that 97% of GC cases that displayed *ESRP1* promoter hypomethylation (*n*=36/37) was also hypomethylated at the *FGFR2* promoter. Of notice, this preliminary analysis revealed two main *FGFR2* mathematical promoters, as two main transcription start sites were annotated to *FGFR2* locus. However, only the *FGFR2* mathematical promoter overlapping a predicted CpG island was considered for further analyses.

In the 13 GC pairs selected for bisulfite Sanger sequencing validation (Task 2.1), 10 that had been classified as hypomethylated by RRBS also showed *FGFR2* promoter hypomethylation for the proxy region in tumour samples in comparison with their corresponding normal samples (Table 18). From the 3 GC cases that showed no variation in *FGFR2* promoter methylation by RRBS, 2 also presented no change in methylation in tumour samples by Sanger sequencing, while 1 GC pair was hypermethylated in the proxy region.

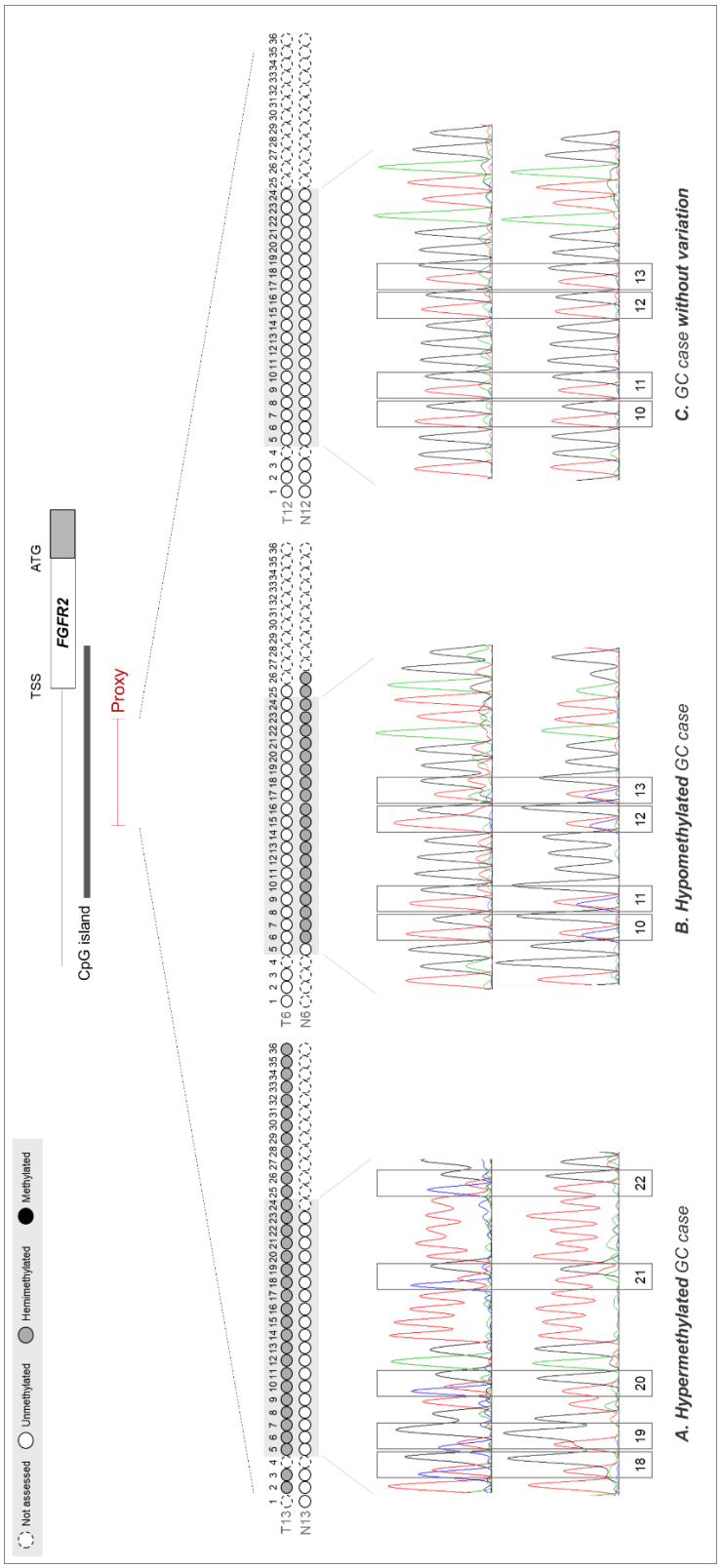


Figure 6: Example of the methylation patterns obtained at the *FGFR2* promoter proxy region by Bisulfite Sanger sequencing in the analysed 13 GC cases from Cohort 1, with each CpG dinucleotide classified as unmethylated, methylated, hemimethylated or not assessed if the detected peak corresponded to a thymine, a cytosine, both or none was detected, respectively. **A)** Example of GC case in which the tumour sample (T13) was classified as hypermethylated in comparison with its paired normal sample (N13). **B)** Example of a GC case in which the tumour sample (T6) was classified as hypomethylated in comparison with its paired normal sample (N6). **C)** Example of GC case in which the tumour sample (T12) was classified as without variation in methylation in comparison with its paired normal sample (N12).

Table 18: Comparison between the obtained *FGFR2* promoter methylation status by RRBS and Sanger sequencing in 13 chosen tumour samples (T) *vs.* the corresponding normal (N) samples from Cohort 1.

T/N samples	*FGFR2* methylation status	
	RRBS	Sanger seq.
T1 / N1	0,21	○
T2 / N2	0,18	○
T3 / N3	0,08	○
T4 / N4	0,18	○
T5 / N5	0,44	○
T6 / N6	0,12	○
T7 / N7	0,38	○
T8 / N8	0,08	○
T9 / N9	0,24	○
T10 / N10	0,16	○
T11 / N11	1,16	**NV**
T12 / N12	0,91	**NV**
T13 / N13	1,09	●

Footnote: 1) RRBS methylation ratio classification: ≥1.5, hypermethylation; ≤0.66, hypomethylation; >0.66 and <1.5, without variation. 2) Sanger sequencing classification: black circle, hypermethylation; white circle, hypomethylation; NV, no variation.

For a total of 12 out of 13 GC cases, the results from RRBS regarding the FGFR2 promoter methylation status were confirmed by Bisulfite Sanger sequencing. These results support the validity of both the RRBS results and the proxy region chosen to evaluate FGFR2 promoter methylation analysis.

As mentioned before, *FGFR2* undergoes a major splicing event at the third immunoglobulin domain, originating three different isoforms: FGFR2-IIIa, FGFR2-IIIb and FGFR2-IIIc. The expression of FGFR2-IIIb and FGFR2-IIIc is known to be tissue-specific and to be differently deregulated in various tumour types [154-156]. FGFR2-IIIb has been found overexpressed in GC and the interest in targeting this specific isoform in GC patients has increased [75, 157, 158]. We hypothesize that *FGFR2* promoter hypomethylation could increase the total expression of the gene and consequently impact the expression of *FGFR2-IIIb* in GC. Therefore, it is conceivable that targeting the later isoform, in detriment of total FGFR2, would be more therapeutically effective. For this reason, RNA expression of total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* was quantified in the 13 GC pairs which was integrated with the *FGFR2* promoter methylation status (Task 2.2, Table 19). RNA expression data of total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* was not available in 7, 3 and 9 out of the 13 GC pairs, respectively, due to low RNA quality. Two out of the 6 GC cases with total *FGFR2* RNA expression data presented an increase of total *FGFR2* RNA expression in the tumour

samples, while the remaining 4 GC cases showed equal levels of RNA expression between the tumour and normal samples. *FGFR2-IIIb* isoform was overexpressed in 7 out of 10 GC cases with RNA expression data, while in 2 cases presented equal RNA expression levels and in 1 case showed a decrease of expression. For *FGFR2-IIIc*, all 4 GC cases with RNA expression data available presented decreased expression in tumour *vs.* normal samples.

Table 19: Concomitance of the *FGFR2* promoter methylation status found in each GC pair by RRBS or Bisulfite Sanger sequencing and the RNA expression level of total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* detected by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N).

T/N samples	*FGFR2* methylation		*FGFR2* RNA expression	*FGFR2-IIIb* RNA expression	*FGFR2-IIIc* RNA expression
	RRBS	Sanger			
T1 / N1	○	○	2,28x	7,34x	0,11x
T2 / N2	○	○	*NA*	1,50x	*NA*
T3 / N3	○	○	*NA*	2,37x	*NA*
T4 / N4	○	○	*NA*	1,67x	0,08x
T5 / N5	○	○	1,12x	2,91x	*NA*
T6 / N6	○	○	0,54x	0,21x	0,39x
T7 / N7	○	○	*NA*	1,12x	0,07x
T8 / N8	○	○	2,23x	*NA*	*NA*
T9 / N9	○	○	*NA*	*NA*	*NA*
T10 / N10	○	○	*NA*	*NA*	*NA*
T11 / N11	NV	NV	1,40x	32,19x	*NA*
T12 / N12	NV	NV	1,04x	1,48x	*NA*
T13 / N13	NV	●	*NA*	5,11x	*NA*

Footnote: 1) *FGFR2* methylation: black circle, hypermethylation; white circle, hypomethylation; NV, no variation. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

As we correlated *FGFR2* promoter methylation status with total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression (Table 19), we observed that:

- From the 10 hypomethylated GC cases, 2 had an increase of total *FGFR2* RNA expression, 5 presented *FGFR2-IIIb* overexpression and 4 showed decreased *FGFR2-IIIc* expression;
- The 2 GC cases without variation in *FGFR2* promoter methylation showed equal levels of total *FGFR2* RNA expression between tumour and normal samples, one case presented overexpression of *FGFR2-IIIb* and the other equal levels between tumour and normal (borderline increase), and no information was available concerning *FGFR2-IIIc* expression for both cases;

There were only 2 GC cases with information available for all 4 features: *FGFR2* promoter methylation status and *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression. Nevertheless, despite both GC cases presenting *FGFR2* promoter hypomethylation, one case showed overexpression of total *FGFR2* and *FGFR2-IIIb* and decreased *FGFR2-IIIc* expression while

the other case showed equal levels of total *FGFR2* RNA expression and decreased *FGFR2-IIIb* and *FGFR2-IIIc* expression.

In summary, 5/10 GC cases with FGFR2 promoter hypomethylation displayed increased FGFR2-IIIb expression, while total FGFR2 overexpression was only detected in 2 out of the 10 hypomethylated GC cases. These data support the hypothesis that *FGFR2* promoter hypomethylation is associated with *FGFR2*/*FGFR2-IIIb* increased RNA expression.

2.2. Correlating *ESRP1* promoter methylation status with *ESRP1* RNA expression in tumour *vs*. normal samples

The role of the splicing regulator ESRP1 in *FGFR2* alternative splicing is known in the EMT context, during which it contributes for the tissue-specific expression of the isoforms *FGFR2-IIIb* and *FGFR2-IIIc* [93, 96]. However, despite *FGFR2-IIIb* overexpression has been reported in GC, the contribution of ESRP1 in this deregulation remains unexplored in this type of cancer. Therefore, we aimed at exploring the promoter methylation status and RNA expression of *ESRP1* and its association with the expression of *FGFR2-IIIb* in GC.

We observed that from the 13 GC pairs previously chosen for *FGFR2* promoter methylation validation and RNA expression analysis, 11 also displayed *ESRP1* promoter hypomethylation while 2 showed no variation in promoter methylation between tumour and normal samples (Table 20). For *ESRP1*, no bisulfite Sanger sequencing validation of the promoter methylation status was performed and RRBS results were taken as valid, considering the previous results for *FGFR2* and *KDR*. We next quantified *ESRP1* RNA expression in the same 13 GC pairs by qRT-PCR, however RNA expression data was only obtained for 9 out of the 13 GC cases (Table 20). From these, 8 showed *ESRP1* overexpression, while 1 presented no variation in RNA expression between tumour and normal samples. Correlating *ESRP1* promoter methylation status and its level of RNA expression, it was possible to observe that from the 8 GC cases with RNA expression data available and *ESRP1* overexpression, 6 were also hypomethylated while 2 showed no variation in methylation.

Table 20: Concomitance of the *ESRP1* promoter methylation status found in each GC pair by RRBS at the *ESRP1* mathematical promoter and the RNA expression level of *ESRP1* detected by qRT-PCR in the 13 chosen tumour samples (T) in comparison with the corresponding normal samples (N) from Cohort 3.

T/N samples	*ESRP1* methylation RRBS		*ESRP1* RNA expression
T1 / N1	0,44	◯	**20,12x**
T2 / N2	0,40	◯	**3,84x**
T3 / N3	0,15	◯	**17,86x**
T4 / N4	0,28	◯	*NA*
T5 / N5	0,60	◯	**6,85x**
T6 / N6	0,17	◯	*NA*
T7 / N7	0,30	◯	**8,52x**
T8 / N8	0,17	◯	**1,32x**
T9 / N9	0,33	◯	*NA*
T10 / N10	0,28	◯	**1,62x**
T11 / N11	1,43	**NV**	**4,58x**
T12 / N12	0,66	◯	*NA*
T13 / N13	1,03	**NV**	**10,39x**

Footnote: 1) RRBS methylation ratio classification: ≥1.5, hypermethylation; ≤0.66, hypomethylation; >0.66 and <1.5, without variation. White circle, hypomethylation; NV, no variation in methylation. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

In summary, ESRP1 was overexpressed and hypomethylated in 6 out of the 8 GC cases with available RNA data. These data support the hypothesis that ESRP1 promoter hypomethylation is associated with increased ESRP1 RNA expression.

2.3. Correlating *ESPR1* promoter methylation status and RNA expression with *FGFR2* promoter methylation and *FGFR2-IIIb* RNA expression in tumour *vs.* normal samples

After observing overexpression of both *ESRP1* and *FGFR2-IIIb* in the majority of analysed cases and given the known role of ESRP1 in *FGFR2* splicing, we hypothesized that *ESRP1* promoter hypomethylation could be contributing for ESRP1 overexpression and consequent upregulation of the FGFR2-IIIb isoform in GC cases. From the 6 GC pairs with *ESRP1* promoter hypomethylation and increased *ESRP1* RNA expression 4 showed also *FGFR2-IIIb* overexpression (Table 21).

Table 21: Concomitance of the *ESRP1* promoter methylation status found in each GC pair by RRBS at the *ESRP1* mathematical promoter and the RNA expression level of *ESRP1* and *FGFR2-IIIb* detected by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N).

T/N samples	*ESRP1* methylation RRBS	*ESRP1* RNA expression	*FGFR2-IIIb* RNA expression
T1 / N1	◯	**20,12x**	**7,34x**
T2 / N2	◯	**3,84x**	**1,50x**
T3 / N3	◯	**17,86x**	**2,37x**
T4 / N4	◯	*NA*	**1,67x**
T5 / N5	◯	**6,85x**	**2,91x**
T6 / N6	◯	*NA*	**0,21x**
T7 / N7	◯	**8,52x**	**1,12x**
T8 / N8	◯	**1,32x**	*NA*
T9 / N9	◯	*NA*	*NA*
T10 / N10	◯	**1,62x**	*NA*
T11 / N11	**NV**	**4,58x**	**32,19x**
T12 / N12	◯	*NA*	**1,48x**
T13 / N13	**NV**	**10,39x**	**5,11x**

Footnote: 1) *ESRP1* methylation: White circle, hypomethylation; NV, no variation in methylation. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

Therefore, both ESRP1 promoter hypomethylation and increased ESRP1 RNA expression could be contributing for the overexpression of FGFR2-IIIb in GC cases.

Despite the possible regulatory role of *ESRP1* promoter hypomethylation in the observed related increase of expression of *ESRP1* and *FGFR2-IIIb* in gastric tumours, this deregulation could also be due to *FGFR2* promoter demethylation. In the 13 analysed GC pairs, 10 displayed promoter hypomethylation of both *ESRP1* and *FGFR2* (Table 22). Despite half of these cases (*n*=5) showing no information regarding expression of either *ESRP1* or *FGFR2-IIIb*, 4 of the 5 remaining GC cases showed both *ESRP1* and *FGFR2-IIIb* overexpression. In parallel, from the 2 GC cases without *ESRP1* promoter methylation variation, one presented *FGFR2* hypomethylation and the other showed *FGFR2* hypermethylation in tumour *vs.* normal samples. Nevertheless, both GC cases presented increased *ESRP1* and *FGFR2-IIIb* RNA expression in tumour samples.

Table 22: Concomitance of the *ESRP1* and *FGFR2* promoter methylation status found in each GC pair by RRBS at the *ESRP1* and *FGFR2* mathematical promoter, respectively, and the RNA expression level of *ESRP1* and *FGFR2-IIIb* detected by qRT-PCR in tumour samples (T) in comparison with the corresponding normal samples (N).

T/N samples	*ESRP1* methylation RRBS	*ESRP1* RNA expression	*FGFR2* methylation Sanger	*FGFR2-IIIb* RNA expression
T1 / N1	○	20,12x	○	7,34x
T2 / N2	○	3,84x	○	1,50x
T3 / N3	○	17,86x	○	2,37x
T4 / N4	○	*NA*	○	1,67x
T5 / N5	○	6,85x	○	2,91x
T6 / N6	○	*NA*	○	0,21x
T7 / N7	○	8,52x	○	1,12x
T8 / N8	○	1,32x	○	*NA*
T9 / N9	○	*NA*	○	*NA*
T10 / N10	○	1,62x	○	*NA*
T11 / N11	NV	4,58x	○	32,19x
T12 / N12	○	*NA*	○	1,48x
T13 / N13	NV	10,39x	●	5,11x

Footnote: 1) Methylation: Black circle, hypermethylation; White circle, hypomethylation; NV, no variation in methylation. 2) RNA expression: ≥1.5, increased expression; ≤0.5, decreased expression; >0.5/<1.5, no variation; NA, not available.

In summary, from the 5 GC cases with ESRP1 and FGFR2 promoter, 4 GC showed also increased ESRP1 and FGFR2-IIIb RNA expression.

These data support the hypothesis that the transcriptionally-permissive state of the promoter of both FGFR2 and ESRP1 allows the overexpression of both genes. Additionally, the increased availability of FGFR2 pre-mRNA and increased level of the ESRP1 splicing factor likely *favours the higher abundance of the FGFR2-IIIb isoform. Therefore, we could hypothesize that GC patients harbouring tumours with FGFR2/ESRP1 promoter hypomethylation and increased RNA expression of ESRP1 and FGFR2-IIIb would benefit the most from anti-FGFR2-IIIb therapies.*

2.4. Correlating *FGFR2* and *ESRP1* promoter methylation with their RNA expression in other GC cohorts

To validate these observations from Cohort 1, similarly to the previous analysis for *KDR*/VEGFR2 and *VEGFA*, the correlations found using 13 GC cases from Cohort 1 regarding *FGFR2* and *ESRP1* promoter methylation and RNA expression levels were re-evaluated using the previously described TCGA-derived Cohorts 2 and 3 (task 2.5). Cohort 2 encompassed 300 gastric adenocarcinomas with data for all 4 evaluated features (*FGFR2*

and *ESRP1* promoter methylation and *FGFR2* and *ESRP1* RNA expression). Cohort 3 encompassed 24 paired gastric tumour and normal samples, however, no data was available for normal samples for the methylation status of *FGFR2* and *ESRP1*. Hence, gastric tumours from Cohort 3 were analysed together with Cohort 2, totalling 324 gastric tumours.

To address *FGFR2* and *ESRP1* promoter methylation status, three regions were again defined: the mathematical promoters, the nearest CpG islands and the proxy regions. Of notice, although no bisulfite Sanger sequencing analysis was performed for the *ESRP1* proxy region in GC cases from Cohort 1, we opted for its definition and usage in Cohorts 2/3 analyses.

Regarding *FGFR2*, the promoter methylation found in the three analysed regions did not present any statistical difference (data not shown). Therefore, we used the level of methylation found at the *FGFR2* proxy region for further analyses. In this region, we observed that 319 out of 324 tumours (98.5%) from Cohorts 2/3 displayed *FGFR2* promoter demethylation, while only 5 out of 324 tumours (1.5%) presented *FGFR2* promoter hemimethylation. *These data support the observed FGFR2 promoter hypomethylation found in tumour samples from Cohort 1, in comparison with their normal counterparts.*

Next, the expression of total *FGFR2* was assessed, revealing that tumour samples from Cohorts 2/3 did not display statistically significant differences in the RNA expression levels when compared with the normal samples from Cohort 3 (Figure 7).

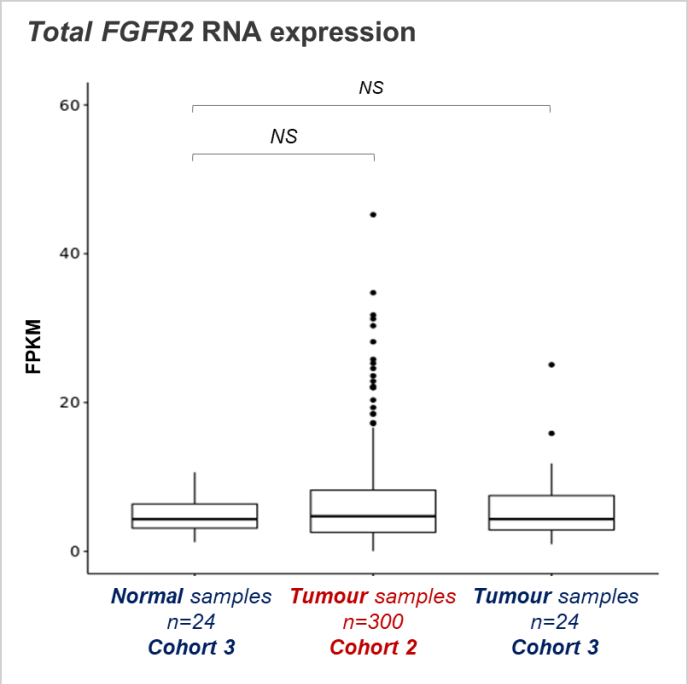


Figure 7: Comparison between the RNA expression levels of total *FGFR2* in tumour samples from Cohorts 2 and 3 *vs.* those obtained in the normal samples from Cohort 3.

When assessing the expression of the specific exons that are differently expressed in *FGFR2-IIIb* and *FGFR2-IIIc*, referred to as exons IIIb and IIIc, we observed an increased expression of exon IIIb in comparison with exon IIIc in tumour samples from Cohorts 2/3 (Figure 8). Moreover, for the normal samples from Cohort 3, this difference between the expression of exons IIIb and IIIc was not observed. Furthermore, the RNA expression of exon IIIb was significantly increased in tumour samples from Cohorts 2/3 in comparison with normal samples from Cohort 3.

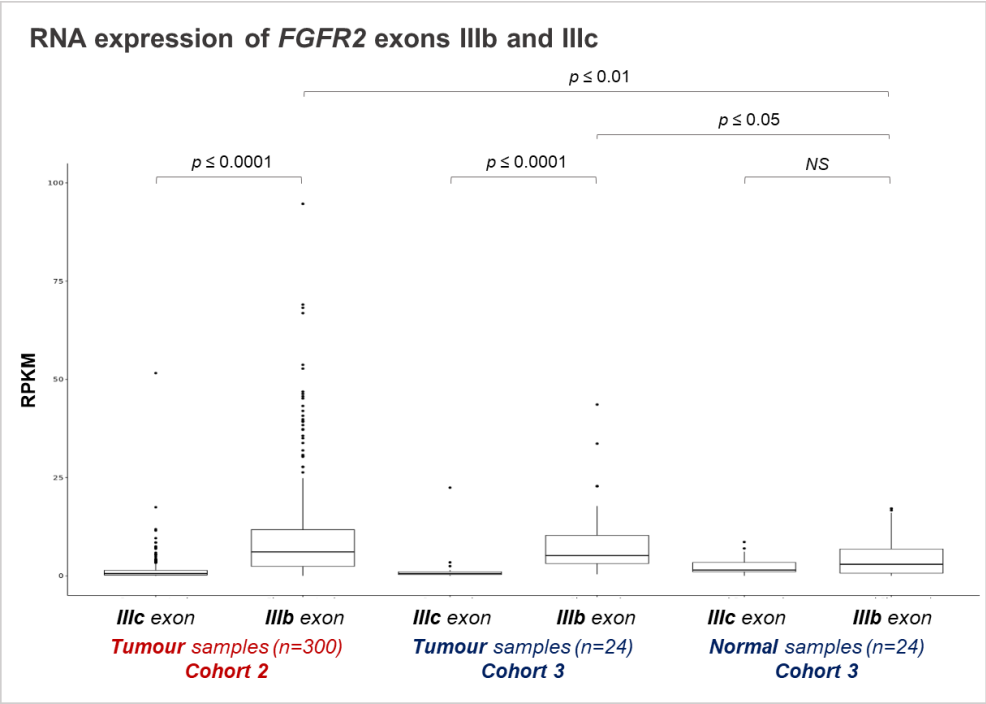


Figure 8: Comparison between the RNA expression levels of *exon*s IIIb and IIIc in tumour samples from Cohorts 2 and 3 *vs.* those obtained in the normal samples from Cohort 3.

Despite the fact that total *FGFR2* expression did not significantly change between tumour and normal samples (Figure 7), 55% of the tumour samples (*n*=177/324) actually presented total *FGFR2* upregulation in comparison with the median expression of Cohort 3 normal sample. Following this trend, 71% of tumour samples (*n*=230/324) presented upregulation of *FGFR2-IIIb*. Moreover, and following cohort 1 data, 78% of tumour samples (*n*=254/324) presented downregulation of exon IIIc (Figure 8, Table 23).

Table 23: *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression level found in the 324 tumour samples from Cohorts 2/3. RNA expression observed in tumours is represented in relation to that of normal samples from Cohort 3, *i.e.* tumours are categorized as upregulated or downregulated if the expression level found in tumour samples of total *FGFR2*, *FGFR2-IIIb* or *FGFR2-IIIc* is above or below, respectively, to that found in the normal samples from Cohort 3.

Number of Tumour Samples *(Cohorts 2/3, n=324)*			
Gene	**RNA Expression Categories**	**Cohort 2/3 Tumour Samples**	**Cohort 3 Normal Samples**
Total FGFR2	Upregulated	177	12
	Downregulated	147	12
FGFR2-IIIb	Upregulated	230	12
	Downregulated	94	12
FGFR2-IIIc	Upregulated	70	12
	Downregulated	254	12

We further calculated a ratio between the expression level of exons IIIb and IIIc to evaluate the balance between both exons across samples. We observed that tumour samples from both cohorts presented an increased expression ratio IIIb/IIIc in comparison with the normal samples from Cohort 3 (Figure 9). In addition, most tumour samples presented expression ratios above 1, which further demonstrated the increased expression of exon IIIb over exon IIIc in gastric tumour samples.

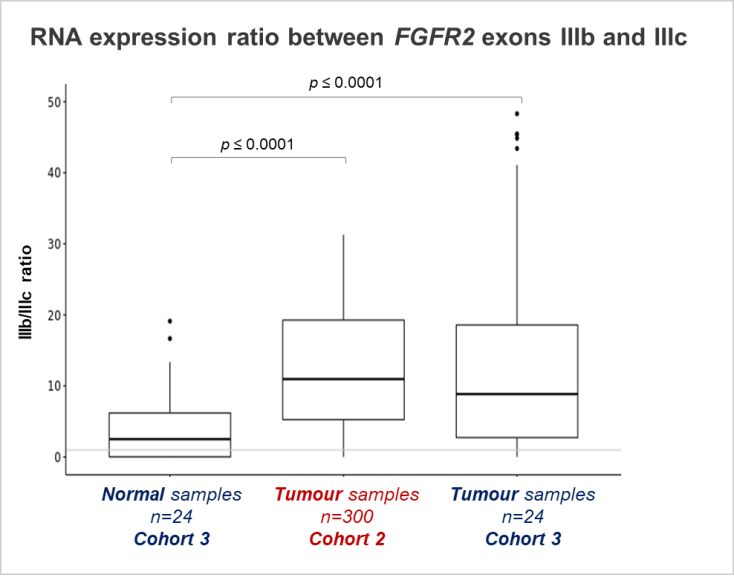


Figure 9: Comparison between the RNA expression ratios exon IIIb/exon IIIc of *FGFR2* in the 324 tumour samples from Cohorts 2 and 3 *vs.* those obtained in the normal samples from Cohort 3.

These data support the preferential overexpression of FGFR2-IIIb in GC tumours, in detriment of the FGFR2-IIIc isoform. Moreover, these results reinforce the importance of addressing the expression of the specific FGFR2-IIIb isoform, instead of detecting all

FGFR2 isoforms, which may be diluting the effect of one dominant isoform. Despite the fact that some studies have already assessed FGFR2-IIIb protein expression in GC, less than 5% of cases displayed overexpression of this isoform [74, 75]. Therefore, and considering availability of RNA as an important factor, it is likely that *FGFR2-IIIb* RNA expression, rather than its protein expression levels, could constitute a valuable predictive biomarker for anti-FGFR2/FGFR2-IIIb therapies.

Next, we correlated the promoter methylation status of *FGFR2* with the expression of total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc*, however, no statistically significant correlation was found for any of the analysed regions (mathematical promoter, CpG island and proxy regions, data not shown). This result is expected as 98.5% of GC tumours were demethylated in cohorts 2/3.

As we categorized total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression levels into up and downregulated, and integrated in the promoter methylation levels found at the *FGFR2* promoter proxy region, we observed that 55% of demethylated tumours (*n*=177/319) displayed increased total *FGFR2* and 72% (*n*=229/319) presented *FGFR2-IIIb* increased RNA expression (Table 24). However, the same was not observed from *FGFR2-IIIc* RNA expression: most demethylated tumours presented downregulation of this isoform (79%, *n*=251/319, Table 24). Confirming the relationship between *FGFR2* promoter demethylation and increased expression, is the opposite correlation observed in the only 5 cases presenting *FGFR2* promoter hemimethylation: all 5 displayed total *FGFR2* RNA expression downregulation and 4/5 showed also *FGFR2-IIIb* downregulation.

Table 24: Promoter methylation status of *FGFR2* at proxy region in the 324 tumour samples from Cohorts 2/3 and the total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression found in the same samples in comparison with the normal samples from Cohort 3. RNA expression is represented regardless of the expression categories as well according to the categories upregulated or downregulated *i.e.* tumours with an RNA expression level above or below the median expression value of a given gene in normal samples from Cohort 3, are classified as upregulated or downregulated, respectively.

Analysed Region	Promoter Methylation Status	Number of Tumour Samples (Cohorts 2/3, n=324)						
		Regardless of RNA Expression Categories	*Taking into account RNA Expression Categories*					
			Total FGFR2		*FGFR2-IIIb*		*FGFR2-IIIc*	
			Upregulated	Downregulated	Upregulated	Downregulated	Upregulated	Downregulated
FGFR2 Proxy	Demethylated	**319**	177	142	229	90	68	251
	Hemimethylated	**5**	0	5	1	4	2	3
	Methylated	**0**	0	0	0	0	0	0

Footnote: Tumours with *FGFR2* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Therefore, similarly to Cohort 1 results, these data support the hypothesis that the transcriptionally-permissive state of the FGFR2 promoter allows for the increased

expression of total FGFR2 and of the specific FGFR2-IIIb isoform. Therefore, FGFR2-IIIb RNA expression could be an appropriate biomarker for anti-FGFR2/FGFR2-IIIb therapy response. In this case, FGFR2 promoter methylation status would be an inadequate biomarker, given that the FGFR2 promoter is consistently demethylated in gastric tumour samples. In contrast, its overexpression and particularly of the FGFR2-IIIb isoform may arise as an interesting predictive marker of therapy response.

Next, we assessed the promoter methylation of *ESRP1* in Cohorts 2/3 and observed that more than 99% of tumour samples displayed *ESRP1* promoter demethylation in all three analysed regions (Table 25). Therefore, we were able to consider the methylation level found in either region for further analyses. Of notice, given that the *ESRP1* promoter proxy region was not yet validated in the 13 GC cases from Cohort 1, we considered the *ESRP1* mathematical promoter for further correlations.

Table 25: Promoter methylation status of *ESRP1* found at its mathematical promoter, CpG island and proxy region in the 324 tumour samples from Cohorts 2/3.

Number of Tumour Samples *(Cohorts 2/3, n=324)*		
Analysed Region	***ESRP1* Promoter Methylation Status**	
Mathematical Promoter	Demethylated Hemimethylated Methylated	323 0 1
CpG Island	Demethylated Hemimethylated Methylated	322 1 1
Proxy	Demethylated Hemimethylated Methylated	322 1 1

Footnote: Tumours with *ESRP1* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Next, the RNA expression of *ESRP1* was assessed, revealing that tumour samples from Cohorts 2/3 displayed a significant increase of *ESRP1* RNA expression in comparison with Cohort 3 normal samples (Figure 10).

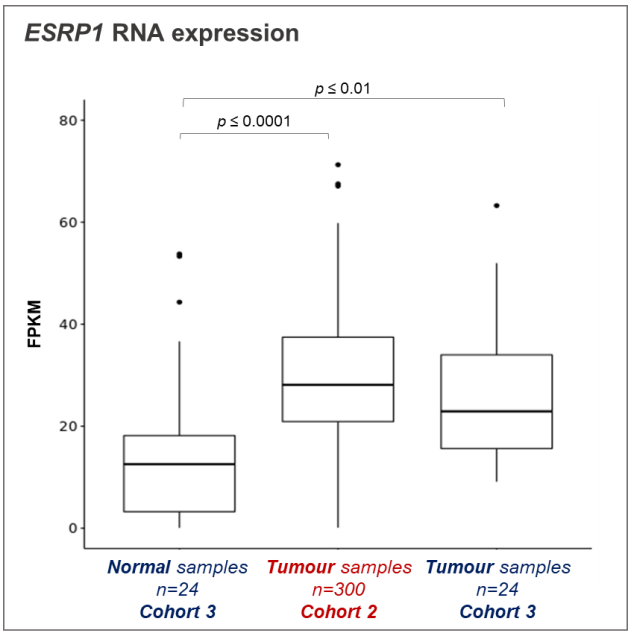


Figure 10: Comparison between the *ESRP1* RNA expression levels found in the tumour samples from Cohorts 2 or 3, in comparison with that obtained in the normal samples from Cohort 3.

Similarly to *FGFR2*, and as expected, no statistically significant correlation was found between *ESRP1* methylation levels in any of the assessed regions and *ESRP1* RNA expression (data not shown). This is clearly due to the fact that over 99% of tumour samples from Cohorts 2/3 displayed *ESRP1* promoter demethylation (Table 25). Considering the *ESRP1* mathematical promoter, we observed that 92% of the tumour samples with *ESRP1* promoter demethylation (*n*=297/323) were upregulated in comparison with the median expression of Cohort 3 normal samples (Table 26).

Therefore, these results support the hypothesis that ESRP1 promoter demethylation allows for the increased ESRP1 RNA expression in GC.

Figure 26: Methylation status of *ESRP1* at the mathematical promoter in the 324 tumour samples from Cohorts 2/3 and *ESRP1* RNA expression found in the same samples in comparison with the normal samples from Cohort 3. RNA expression is represented regardless of the expression categories as well according to the categories upregulated or downregulated *i.e.* tumours with an RNA expression level above or below the median expression value of a given gene in normal samples from Cohort 3, are classified as upregulated or downregulated, respectively.

Analysed Region	***ESRP1* Promoter Methylation Status**	**Number of Tumour Samples** *(Cohorts 2/3, n=324)*		
		Regardless of RNA Expression Categories	***Taking into account RNA Expression Categories***	
			Upregulated	**Downregulated**
Mathematical Promoter	Demethylated	323	297	26
	Hemimethylated	0	0	0
	Methylated	1	0	1

Footnote: Tumours with *ESRP1* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Next, recalling our **specific aim 2** and the results from Cohort 1, we assessed the co-ocurrence of *FGFR2*/*ESRP1* promoter demethylation and overexpression of *ESRP1* and *FGFR2-IIIb* in the tumour samples from Cohorts 2/3. Moreover, the expression of *FGFR2-IIIc* was also considered, to confirm the preferential expression of the *FGFR2-IIIc* isoform. We observed that 319 out of 324 tumour samples (98%) presented both *ESRP1* and *FGFR2* promoter demethylation. From these 319 samples, 214 (67%) samples presented *ESRP1* and *FGFR2-IIIb* upregulation (Table 27). Interestingly, from these 214 samples, 176 (82%) samples displayed also *FGFR2-IIIc* downregulation. Regarding the entirety of Cohorts 2/3 tumour samples, we observed a co-occurrence of *ESRP1* and *FGFR2* promoter hypomethylation, *ESRP1* and *FGFR2-IIIb* expression upregulation and *FGFR2-IIIc* downregulation in 54% of all tumours in Cohorts 2/3 (*n*=176 out of 324).

Table 27: Number of tumour samples from Cohorts 2/3 distributed according to *FGFR2* and *ESRP1* promoter methylation status at the proxy region and mathematical promoter, respectively, and *ESPR1*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression.

Number of Tumour Samples *(Cohorts 2/3, n=324)*					
***FGFR2* Promoter Methylation Status**	***ESRP1* Promoter Methylation Status**	**RNA Expression**			
		ESRP1	***FGFR2-IIIb***	***FGFR2-IIIc***	**Total**
Demethylated ***n*=319**	Demethylated ***n*=319**	Upregulated ***n*=293**	Upregulated ***n*=214**	Upregulated	38
				Downregulated	**176**
			Downregulated	Upregulated	22
				Downregulated	57
		Downregulated	Upregulated	Upregulated	3
				Downregulated	12
			Downregulated	Upregulated	5
				Downregulated	6
	Hemimethylated	*NA*			
	Methylated	*NA*			

Footnote: Tumours with *KDR* methylation beta-values: ≤0.33, demethylated; ≥0.66, methylated; >0.33/<0.66, hemimethylated.

Altogether, the observed concomitance of ESRP1 and FGFR2 promoter demethylation, ESRP1/FGFR2-IIIb overexpression and FGFR2-IIIc downregulation suggests that the promoter of both genes is in a transcriptionally-permissive state, allowing ESRP1 to regulate the alternative splicing of FGFR2 favouring the expression of the FGFR2-IIIb isoform.

To better understand this interplay, we next attempted to use a gastric cancer line model to modulate the methylation status of *ESRP1*/*FGFR2* promoters and understand how it affects the RNA expression of both molecules.

PART 3. Exploring the role of ESRP1 in the expression regulation of FGFR2-IIIb expression in GC using *in vitro* models

3.1. Characterizing *ESRP1* and *FGFR2* promoter methylation status and RNA expression in MCF10A and GC cell lines

To further investigate the interplay between *ESRP1*/*FGFR2* promoter methylation status and the RNA expression levels of ESRP1 and FGFR2-IIIb, we search for the ideal model in which a cell line displayed *ESRP1* and *FGFR2* promoter methylation and basal levels of RNA expression of both molecules (task 2.6). For this, we characterized 4 GC cell lines regarding *ESRP1* and *FGFR2* promoter methylation status by Bisulfite Sanger sequencing, and *ESRP1*, total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression, using qRT-PCR: MKN74, MKN45, KATO-III and IPA-220 cells. In addition, given the known role of ESRP1 during EMT [96], in particular, in the control of tissue-specific expression of FGFR2 isoforms during EMT, we also assessed *ESRP1* promoter methylation status and the RNA expression levels of *ESRP1* and *FGFR2-IIIb* and *FGFR2-IIIc* isoforms in epithelial and mesenchymal cells. For this purpose, we used MCF10A epithelial (E-cells) and TGF-β-induced mesenchymal cells (M-cells) from an MCF10A EMT/MET model previously generated [159].

To assess *FGFR2* and *ESRP1* promoter methylation, the same proxy regions evaluated in Cohorts 1, 2 and 3 were sequenced in these cell lines. From the 4 analysed GC cell lines, all displayed *ESRP1* promoter demethylation, while the near-normal cell line MCF10A was hemimethylated. Indeed, both MCF10A E- and M-cells presented *ESRP1* promoter hemimethylation (Figure 11A). Regarding *FGFR2* promoter methylation, all analysed cell lines displayed promoter demethylation (Figure 11B), as observed for most GC tumours analysed.

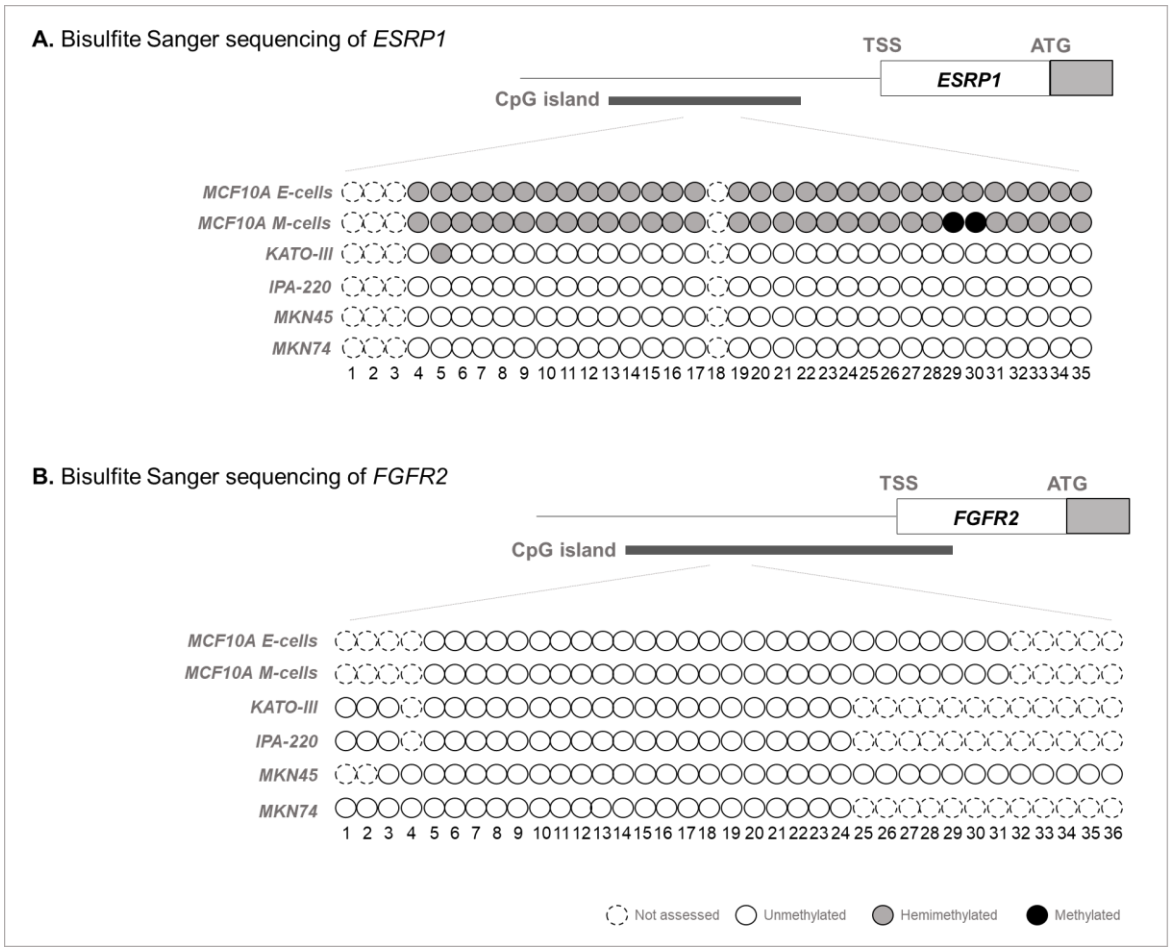


Figure 11: Representation of the methylation status found at the *ESRP1* and *FGFR2* promoter proxy regions in several cell lines. **A)** Methylation status of each CpG dinucleotide found at the *ESRP1* promoter proxy region in MCF10A E- and M-cells, as well as in the KATO-III, IPA-220, MKN45 and MKN74 cell lines. **B)** Methylation status of each CpG dinucleotide found at the *FGFR2* promoter proxy region in MCF10A E- and M-cells, as well in the KATO-III, IPA-220, MKN45 and MKN74 cells lines. CpG dinucleotides were classified as not assessed, unmethylated, methylated or hemimethylated if, by Bisulfite Sanger Sequencing, it was not possible to evaluate the CpG site, or a peak corresponding to a thymine, a cytosine or both, respectively, was obtained at each site.

Regarding RNA expression, we observed that MCF10A E- and M-cells presented similar levels of *ESRP1*, *FGFR2-IIIb* and *FGFR2-IIIc* RNA expression (Figure 12). When comparing with the RNA expression levels found in GC cell lines, we observed that all GC cell lines presented increased RNA expression of *ESRP1* and *FGFR2-IIIb* and similar *FGFR2-IIIc* RNA expression both in comparison with MCF10A E- and M-cells. Of notice, KATO-III GC cell line exhibited the highest RNA expression of *FGFR2-IIIb*, likely due to *FGFR2* genetic amplification [160].

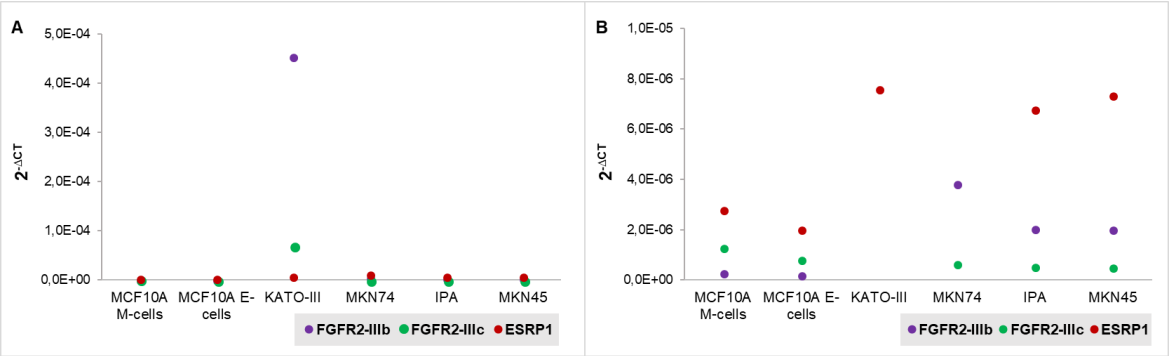


Figure 12: RNA expression levels of *FGFR2-IIIb*, *FGFR2-IIIIc* and *ESRP1* across several cell lines. **A.** Expression in $2^{-\Delta CT}$ values of *FGFR2-IIIb* (purple dots), *FGFR2-IIIIc* (green dots) and *ESRP1* (red) in MCF10A E- and M-cells, as well as in the GC cell lines KATO-III, MKN74, IPA-220 and MKN45, assessed by qRT-PCR. **B.** Zoom in representation of panel A data.

Moreover, *ESRP1* showed the highest RNA expression in GC cell lines, which in turn also presented *ESPR1* promoter demethylation (Figure 11A), unlike MCF10A E- and M-cells. We also observed that *ESRP1* RNA expression was accompanied by an increased *FGFR2-IIIb* RNA expression in all GC cell lines analysed.

In summary, these results confirm the promoter demethylation of ESRP1 and FGFR2 and increased ESRP1 and FGFR2-IIIb RNA expression in GC cells lines, as opposed to the near-normal cell line. However, none of these constituted the ideal model to explore this correlation. For that reason, we focused on MCF10A E-cells that presented ESRP1 promoter hemimethylation, FGFR2 promoter demethylation and lower levels of ESRP1 and FGFR2-IIIb RNA expression, in comparison with GC cell lines. We used MCF10A cells to further investigate promoter methylation as an expression regulation mechanisms of ESRP1 and FGFR2.

3.2. Determining the effect of global DNA demethylation of MCF10A E-cells in *ESRP1* and *FGFR2* RNA expression

MCF10A E-cells were treated with the global demethylating agent 5-Aza-2′-deoxycytidine (5-Aza). This cell line was selected as it displayed *ESRP1* promoter methylation (Figure 11A). MKN45 and MKN74 cell lines were also treated with 5-Aza and used as controls for this experiment. MCF10A E-cells, MKN74 and MKN45 cells were cultured in the presence of 5-Aza or DMSO (vehicle) for 24, 48 or 72 hours and 35 CpG sites in the proxy region were evaluated.

After 24 hours of treatment with 5-Aza, CpG site number 35 became demethylated, in comparison with MCF10 E-cells treated with DMSO, in which the same CpG site was

hemimethylated (Figure 13A). After 48 hours, MCF10A E-cells treated with DMSO presented full hemimethylation of the *ESPR1* promoter, unlike observed at 24 hours. Moreover, MCF10A E-cells treated with 5-AZA at 48 hours presented 3 CpG sites (#31, #34, #35) *de novo* demethylated. After 72 hours of treatment, MCF10A E-cells in the presence of 5-Aza presented demethylation in 6 CpG sites (#7, #9, #17, #28, #31) unlike the DMSO control. CpG sites #34 and #35 could not be confidently assessed. *FGFR2* promoter demethylation was maintained across all experimental conditions (Figure 13B).

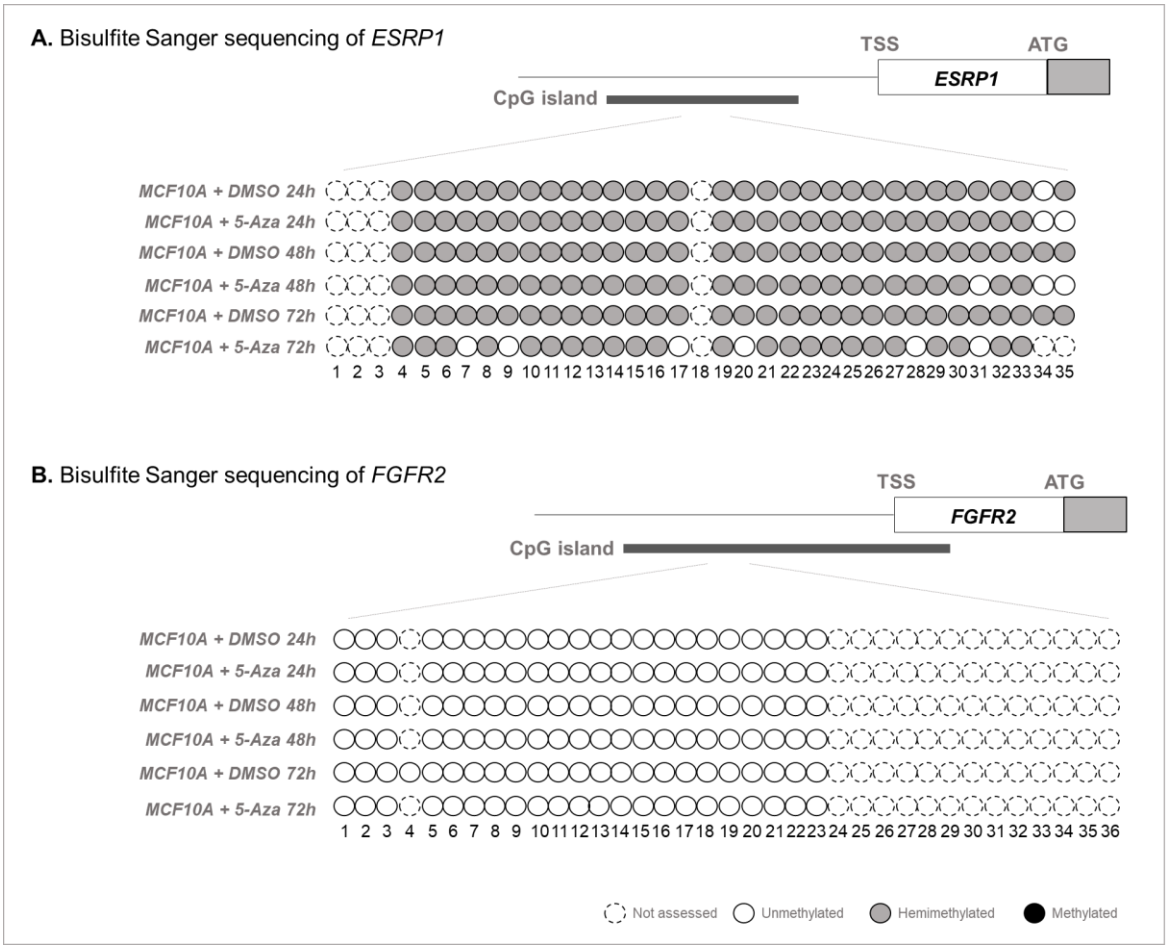


Figure 13: Representation of the methylation status found at the *ESRP1* and *FGFR2* promoter proxy regions in MCF10A E-cells treated with DMSO or 5-Aza across three time-points. **A)** Methylation status of each CpG dinucleotide found at the *ESRP1* promoter proxy region in MCF10A E-cells treated with DMSO or 5-Aza for 24, 48 or 72 hours. **B)** Methylation status of each CpG dinucleotide found at the *FGFR2* promoter proxy region in MCF10A E-cells treated with DMSO or 5-Aza for 24, 48 or 72 hours. CpG dinucleotides were classified as not assessed, unmethylated, methylated or hemimethylated if, by Bisulfite Sanger Sequencing, it was not possible to evaluate the CpG site, or a peak corresponding to a thymine, a cytosine or both, respectively, was obtained at each site.

Afterwards, RNA expression of *ESRP1*, total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc* was quantified in MCF10A E-cells, MKN74 and MKN45 cells grown in the presence of 5-Aza or DMSO for 24, 48 and 72 hours. MKN74 cells treated with 5-Aza showed no variation in the RNA expression of any of the analysed genes across timepoints, in comparison with MKN74 cells treated with DMSO (Figure 14A). MKN45 cells presented an increase of *FGFR2-IIIb*

RNA expression after 48 hours of treatment with 5-Aza (1.6x), in comparison with MKN45 cells treated with DMSO (Figure 14B). RNA expression of the remaining genes in MKN45 cells treated with either DMSO or 5-Aza did not differ after 24 or 72 hours.

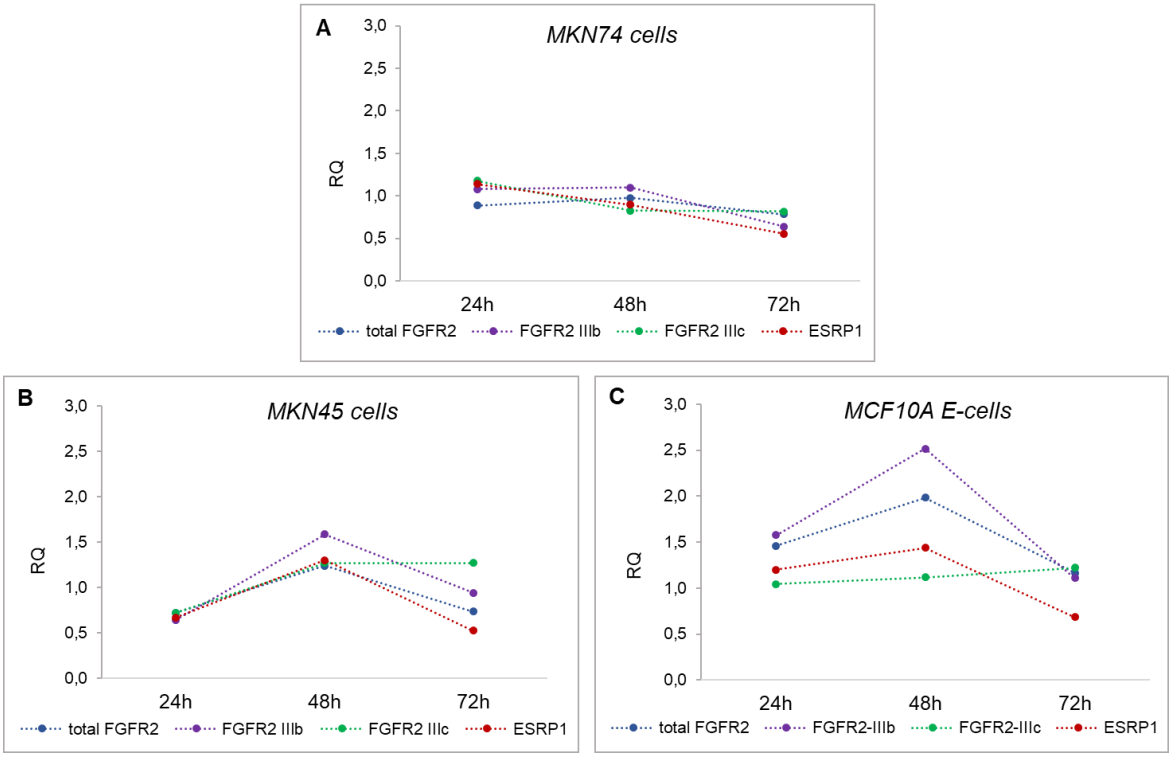


Figure 14: Relative Quantitation (RQ) of the RNA expression of total *FGFR2* (blue line), *FGFR2-IIIb* (purple line), *FGFR2-IIIc* (green line) and *ESRP1* (red line) across several cell lines and timepoints. **A)** RQ expression values in MKN74 cells treated with 5-Aza for 24, 48 and 72 hours, in relation to MKN74 E-cells treated with DMSO. **B)** RQ expression values in MKN45 cells treated with 5-Aza for 24, 48 and 72 hours, in relation to MKN45 cells treated with DMSO. **C)** MCF10A E-cells treated with 5-Aza for 24, 48 and 72 hours, in relation to MCF10A E-cells treated with DMSO.

With an already transcriptionally-permissive state at the ESRP1 and FGFR2 promoter, we confirmed that the demethylation of MKN74 and MKN45 showed no major impact in the expression of the assessed genes.

Regarding MCF10A E-cells RNA expression (Figure 14C), we observed that:

- after 24 hours of treatment with 5-Aza, there was an increase of total *FGFR2* (1.5x) and *FGFR2-IIIb* (1.6x) RNA expression and no change in *FGFR2-IIIc* and *ESRP1* RNA expression, in comparison with MCF10A E-cells treated with DMSO;
- after 48 hours of treatment with 5-Aza, MCF10A E-cells presented an increase of total *FGFR2* (2.0x) and *FGFR2-IIIb* (2.5x), a borderline increase of *ESRP1* (1.4x)

RNA expression and no change in *FGFR2-IIIc* RNA expression, in comparison with MCF10A E-cells treated with DMSO;

- After 72 hours of treatment with 5-Aza, MCF10A E-cells presented no change in RNA expression of any of the analysed genes, in comparison with MCF10A E-cells treated with DMSO.

Of notice, the RNA expression of total *FGFR2* and *FGFR2-IIIb* observed in MCF10A E-cells treated with 5-Aza showed a time-dependant increase from 24 to 48 hours of treatment. Moreover, this increase of total *FGFR2* and *FGFR2-IIIb* was also accompanied, to a lesser degree, by an increase of *ESRP1* RNA expression after 24 and 48 hours of treatment (1.2x and 1.4x, respectively). At 72 hours, total *FGFR2*, *FGFR2-IIIb* and *ESRP1* RNA expression concurrently decreased in MCF10A E-cells treated with 5-Aza, in comparison with MCF10A E-cells treated with 5-Aza for 24 and 48 hours (Figure 14C).

Considering *ESRP1* and *FGFR2* promoter methylation analysis and the RNA expression of total *FGFR2*, *FGFR2-IIIb* and *FGFR2-IIIc*, we observed that:

- After 24 hours of treatment with 5-Aza, MCF10A E-cells showed *de novo* demethylation at the *ESRP1* promoter in 1 CpG site accompanied by increased total *FGFR2* and *FGFR2-IIIb* RNA expression, in comparison with MCF10A E-cells treated with DMSO;
- After 48 hours of treatment with 5-Aza, MCF10A E-cells displayed *de novo* demethylation at the *ESRP1* promoter in 3 CpG sites accompanied by increased total *FGFR2* and *FGFR2-IIIb* RNA expression, in comparison with MCF10A E-cells treated with DMSO;
- After 72 hours of treatment with 5-Aza, MCF10A E-cells presented *de novo* demethylation at the *ESRP1* promoter in 6 CpG sites however no change in RNA expression of any gene was observed, in comparison with MCF10A E-cells treated with DMSO.

In summary, the treatment of MCF10A E-cells with 5-Aza led to: (1) demethylation of some CpG sites at the ESRP1 promoter, leading to an increase of its RNA expression; (2) increase of total FGFR2 and FGFR2-IIIb RNA expression in a time-dependant manner, independently of FGFR2 promoter methylation. Although this experiment was performed only once, and 5-Aza treatment affects many other genes which were not accounted for, our observations clearly support a cause-effect relation between loss of ESRP1 promoter methylation, increased ESRP1 RNA expression that likely triggers a splicing bias towards FGFR2-IIIb isoform. These observations strongly support further studies to better understand the interplay between ESRP1 and FGFR2-IIIb in GC.

V. Conclusion

Our results highlight the potential of evaluating alterations found in GC tumours, other than gene amplification and protein overexpression, as predictive biomarkers for therapy response. Moreover, considering the deregulation of not only one molecule but of key signalling partners, could also be relevant in the stratification of highly heterogeneous tumours as those observed in GC.

Considering as example VEGFR2/VEGFA, our data indicates that RNA overexpression of both molecules likely pinpoints a target group (67%) to therapy intervention with anti-VEGFR2 with or without anti-VEGFA therapies. Moreover, our results further indicated that downregulation of *KDR*/VEGFR2, concomitant or not with *VEGFA* overexpression, selects a group of tumours (19%) that will likely not benefit from any of these therapies. Considering the FGFR2/ESRP1 example, our data revealed the importance of considering the concomitance of *FGFR2* promoter demethylation*, FGFR2-IIIb* overexpression, *FGFR2-IIIc* downregulation, as well as *ESRP1* promoter demethylation and overexpression. With a combination of biomarkers including *FGFR2-IIIb* overexpression, *FGFR2-IIIc* downregulation, and *ESRP1* overexpression, we identified 55% of GC cases that will likely benefit the most from anti-FGFR2/anti-FGFR2-IIIb therapies. This is a current need for GC treatment, as neither genetic amplification nor protein overexpression of VEGFR2 and FGFR2 have shown potential as predictive biomarkers of treatment response.

As a follow-up work, we intent to further explore the role of ESRP1 in FGFR2-IIIb deregulation in GC by inhibiting *ESRP1* RNA expression in an *in vitro* GC model and evaluate its effect on *FGFR2-IIIb* RNA expression. Furthermore, we will investigate whether the observed *FGFR2-IIIb* RNA overexpression is directly correlated with the corresponding protein expression in GC cohorts. We believe that independent validation studies using different GC cohorts with paired normal samples and available clinical data are also required, to truly understand the potential use of RNA expression of *KDR*, *VEGFA*, *FGFR2-IIIb* and *ESRP1* and the promoter methylation status of *FGFR2* and *ESRP1* as novel biomarkers for patient stratification and prediction of therapy efficacy.

VI. Bibliographic References

1. Torre, L.A. et al. (2016) Global Cancer Incidence and Mortality Rates and Trends--An Update. Cancer Epidemiol Biomarkers Prev 25 (1), 16-27.

2. Global Burden of Disease Cancer, C. et al. (2017) Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 3 (4), 524-548.

3. Maximiano, S. et al. (2016) Trastuzumab in the Treatment of Breast Cancer. BioDrugs 30 (2), 75-86.

4. Salgia, R. (2016) Mutation testing for directing upfront targeted therapy and post-progression combination therapy strategies in lung adenocarcinoma. Expert Rev Mol Diagn 16 (7), 737-49.

5. Ferlay, J. et al. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 136 (5), E359-86.

6. Gullo, I. et al. (2018) Heterogeneity in Gastric Cancer: From Pure Morphology to Molecular Classifications. Pathobiology 85 (1-2), 50-63.

7. Karimi, P. et al. (2014) Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev 23 (5), 700-13.

8. Ajani, J.A. et al. (2017) Gastric adenocarcinoma. Nat Rev Dis Primers 3, 17036.

9. Oliveira, C. et al. (2015) Familial gastric cancer: genetic susceptibility, pathology, and implications for management. Lancet Oncol 16 (2), e60-70.

10. Lauren, P. (1965) The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification. Acta Pathol Microbiol Scand 64, 31-49.

11. Bosman, F.T. et al. (2010) WHO Classification of Tumours of the Digestive System., WHO.

12. Qiu, M.Z. et al. (2013) Clinicopathological characteristics and prognostic analysis of Lauren classification in gastric adenocarcinoma in China. J Transl Med 11, 58.

13. Liu, L. et al. (2013) A cohort study and meta-analysis between histopathological classification and prognosis of gastric carcinoma. Anticancer Agents Med Chem 13 (2), 227-34.

14. Fenoglio-Preiser, C.M. et al. (2003) TP53 and gastric carcinoma: a review. Hum Mutat 21 (3), 258-70.

15. Iqbal, N. and Iqbal, N. (2014) Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. Mol Biol Int 2014, 852748.

16. Carneiro, P. et al. (2012) E-cadherin dysfunction in gastric cancer - Cellular consequences, clinical applications and open questions. FEBS Letters 586 (18), 2981-2989.

17. Cancer Genome Atlas Research, N. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. Nature 513 (7517), 202-9.

18. Cristescu, R. et al. (2015) Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. Nat Med 21 (5), 449-56.

19. Smyth, E.C. et al. (2016) Gastric cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 27 (suppl 5), v38-v49.

20. Kim, R. et al. (2013) Geographic differences in approach to advanced gastric cancer: Is there a standard approach? Crit Rev Oncol Hematol 88 (2), 416-26.

21. Bang, Y.J. et al. (2010) Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. Lancet 376 (9742), 687-97.

22. Fuchs, C.S. et al. (2014) Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. Lancet 383 (9911), 31-39.

23. Wilke, H. et al. (2014) Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. Lancet Oncol 15 (11), 1224-35.

24. Hudis, C.A. (2007) Trastuzumab--mechanism of action and use in clinical practice. N Engl J Med 357 (1), 39-51.

25. Slamon, D.J. et al. (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 344 (11), 783-92.

26. Gunturu, K.S. et al. (2013) Gastric cancer and trastuzumab: first biologic therapy in gastric cancer. Ther Adv Med Oncol 5 (2), 143-51.

27. Hofmann, M. et al. (2008) Assessment of a HER2 scoring system for gastric cancer: results from a validation study. Histopathology 52 (7), 797-805.

28. Aprile, G. et al. (2013) Critical appraisal of ramucirumab (IMC-1121B) for cancer treatment: from benchside to clinical use. Drugs 73 (18), 2003-15.

29. Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. Cell 144 (5), 646-74.

30. Simons, M. et al. (2016) Mechanisms and regulation of endothelial VEGF receptor signalling. Nat Rev Mol Cell Biol 17 (10), 611-25.

31. Spratlin, J.L. et al. (2010) Phase I pharmacologic and biologic study of ramucirumab (IMC-1121B), a fully human immunoglobulin G1 monoclonal antibody targeting the vascular endothelial growth factor receptor-2. J Clin Oncol 28 (5), 780-7.

32. Chan, M.M. et al. (2015) Clinical utility of ramucirumab in advanced gastric cancer. Biologics 9, 93-105.

33. Fuchs, C.S. et al. (2016) Biomarker analyses in REGARD gastric/GEJ carcinoma patients treated with VEGFR2-targeted antibody ramucirumab. Br J Cancer 115 (8), 974-982.

34. Van Cutsem, E. et al., Biomarker Analyses of Second-Line Ramucirumab in Patients With Advanced

Gastric Cancer From RAINBOW, a Global, Randomized, Double-Blind, Phase 3 Study, ESMO 18th World Congress on Gastrointestinal Cancer, Barcelona, Spain, 2016.

35. Jayson, G.C. et al. (2016) Antiangiogenic therapy in oncology: current status and future directions. Lancet 388 (10043), 518-29.

36. Claesson-Welsh, L. (2003) Signal transduction by vascular endothelial growth factor receptors. Biochem Soc Trans 31 (Pt 1), 20-4.

37. Ang, Y.L. et al. (2016) Translating gastric cancer genomics into targeted therapies. Crit Rev Oncol Hematol 100, 141-6.

38. Liu, L. et al. (2012) Prognostic value of vascular endothelial growth factor expression in resected gastric cancer. Asian Pac J Cancer Prev 13 (7), 3089-97.

39. Hurwitz, H. et al. (2004) Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med 350 (23), 2335-42.

40. Escudier, B. et al. (2007) Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. Lancet 370 (9605), 2103-11.

41. Friedman, H.S. et al. (2009) Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma. J Clin Oncol 27 (28), 4733-40.

42. Miller, K. et al. (2007) Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. N Engl J Med 357 (26), 2666-76.

43. Reck, M. et al. (2009) Phase III trial of cisplatin plus gemcitabine with either placebo or bevacizumab as first-line therapy for nonsquamous non-small-cell lung cancer: AVAil. J Clin Oncol 27 (8), 1227-34.

44. Roviello, G. et al. (2016) Monoclonal antibodies-based treatment in gastric cancer: current status and future perspectives. Tumour Biol 37 (1), 127-40.

45. Ohtsu, A. et al. (2011) Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a randomized, double-blind, placebo-controlled phase III study. J Clin Oncol 29 (30), 3968-76.

46. Van Cutsem, E. et al. (2012) Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a biomarker evaluation from the AVAGAST randomized phase III trial. J Clin Oncol 30 (17), 2119-27.

47. Shen, L. et al. (2015) Bevacizumab plus capecitabine and cisplatin in Chinese patients with inoperable locally advanced or metastatic gastric or gastroesophageal junction cancer: randomized, double-blind, phase III study (AVATAR study). Gastric Cancer 18 (1), 168-76.

48. Kim, M.A. et al. (2008) EGFR in gastric carcinomas: prognostic significance of protein overexpression and high gene copy number. Histopathology 52 (6), 738-46.

49. Sibertin-Blanc, C. et al. (2016) Monoclonal antibodies for treating gastric cancer: promises and pitfalls. Expert Opin Biol Ther 16 (6), 759-69.

50. Waddell, T. et al. (2013) Epirubicin, oxaliplatin, and capecitabine with or without panitumumab for patients with previously untreated advanced oesophagogastric cancer (REAL3): a randomised, open-label phase 3 trial. Lancet Oncol 14 (6), 481-9.

51. Zhao, B. et al. (2017) Mechanisms of resistance to anti-EGFR therapy in colorectal cancer. Oncotarget 8 (3), 3980-4000.

52. Appleman, L.J. (2011) MET signaling pathway: a rational target for cancer therapy. J Clin Oncol 29 (36), 4837-8.

53. Catenacci, D.V.T. et al. (2017) Rilotumumab plus epirubicin, cisplatin, and capecitabine as first-line therapy in advanced MET-positive gastric or gastro-oesophageal junction cancer (RILOMET-1): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 18 (11), 1467-1482.

54. Doi, T. et al. (2015) A phase 3, multicenter, randomized, double-blind, placebo-controlled study of rilotumumab in combination with cisplatin and capecitabine (CX) as first-line therapy for Asian patients (pts) with advanced MET-positive gastric or gastroesophageal junction (G/GEJ) adenocarcinoma: The RILOMET-2 trial. Journal of Clinical Oncology 33 (3_suppl), TPS226-TPS226.

55. Shah, M.A. et al. (2016) A Randomized Phase II Study of FOLFOX With or Without the MET Inhibitor Onartuzumab in Advanced Adenocarcinoma of the Stomach and Gastroesophageal Junction. Oncologist 21 (9), 1085-90.

56. Shah, M.A. et al. (2017) Effect of Fluorouracil, Leucovorin, and Oxaliplatin With or Without Onartuzumab in HER2-Negative, MET-Positive Gastroesophageal Adenocarcinoma: The METGastric Randomized Clinical Trial. JAMA Oncol 3 (5), 620-627.

57. Kang, Y.-K. et al. (2015) Phase I study of ABT-700, an anti-c-Met antibody, in patients (pts) with advanced gastric or esophageal cancer (GEC). Journal of Clinical Oncology 33 (3_suppl), 167-167.

58. Wang, J. et al. (2016) Anti-c-Met monoclonal antibody ABT-700 breaks oncogene addiction in tumors with MET amplification. BMC Cancer 16, 105.

59. NIH (2018) www.clinicaltrials.gov. www.clinicaltrials.gov, (accessed).

60. Turner, N. and Grose, R. (2010) Fibroblast growth factor signalling: from development to cancer. Nat Rev Cancer 10 (2), 116-29.

61. Touat, M. et al. (2015) Targeting FGFR Signaling in Cancer. Clin Cancer Res 21 (12), 2684-94.

62. Hattori, Y. et al. (1990) K-sam, an amplified gene in stomach cancer, is a member of the heparin-binding growth factor receptor genes. Proc Natl Acad Sci U S A 87 (15), 5983-7.

63. Johnson, D.E. et al. (1991) The human fibroblast growth factor receptor genes: a common structural arrangement underlies the mechanisms for generating receptor forms that differ in their third immunoglobulin domain. Mol Cell Biol 11 (9), 4627-34.

64. Yayon, A. et al. (1992) A confined variable region confers ligand specificity on fibroblast growth factor receptors: implications for the origin of the immunoglobulin fold. EMBO J 11 (5), 1885-90.

65. Ishii, H. et al. (1994) Preferential expression of the third immunoglobulin-like domain of K-sam product provides keratinocyte growth factor-dependent growth in carcinoma cell lines. Cancer Res 54 (2), 518-22.

66. Hattori, Y. et al. (1996) Immunohistochemical detection of K-sam protein in stomach cancer. Clin Cancer Res 2 (8), 1373-81.

67. Liu, Y.J. et al. (2014) HER2, MET and FGFR2 oncogenic driver alterations define distinct molecular segments for targeted therapies in gastric carcinoma. Br J Cancer 110 (5), 1169-78.

68. Nagatsuma, A.K. et al. (2015) Expression profiles of HER2, EGFR, MET and FGFR2 in a large cohort of patients with gastric adenocarcinoma. Gastric Cancer 18 (2), 227-38.

69. Murase, H. et al. (2014) Prognostic significance of the co-overexpression of fibroblast growth factor receptors 1, 2 and 4 in gastric cancer. Mol Clin Oncol 2 (4), 509-517.

70. Su, X. et al. (2014) FGFR2 amplification has prognostic significance in gastric cancer: results from a large international multicentre study. Br J Cancer 110 (4), 967-75.

71. Das, K. et al. (2014) Mutually exclusive FGFR2, HER2, and KRAS gene amplifications in gastric cancer revealed by multicolour FISH. Cancer Lett 353 (2), 167-75.

72. Park, Y.S. et al. (2015) FGFR2 Assessment in Gastric Cancer Using Quantitative Real-Time Polymerase Chain Reaction, Fluorescent In Situ Hybridization, and Immunohistochemistry. Am J Clin Pathol 143 (6), 865-72.

73. Matsumoto, K. et al. (2012) FGFR2 gene amplification and clinicopathological features in gastric cancer. Br J Cancer 106 (4), 727-32.

74. Han, N. et al. (2015) Evaluation of Fibroblast Growth Factor Receptor 2 Expression, Heterogeneity and Clinical Significance in Gastric Cancer. Pathobiology 82 (6), 269-79.

75. Ahn, S. et al. (2016) FGFR2 in gastric cancer: protein overexpression predicts gene amplification and high H-index predicts poor survival. Mod Pathol 29 (9), 1095-103.

76. Inokuchi, M. et al. (2015) Therapeutic targeting of fibroblast growth factor receptors in gastric cancer. Gastroenterol Res Pract 2015, 796380.

77. Van Cutsem, E. et al. (2017) A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification. Ann Oncol 28 (6), 1316-1324.

78. Hierro, C. et al. (2017) Targeting the fibroblast growth factor receptor 2 in gastric cancer: promise or pitfall? Ann Oncol 28 (6), 1207-1216.

79. Ornitz, D.M. and Itoh, N. (2015) The Fibroblast Growth Factor signaling pathway. Wiley Interdiscip Rev Dev Biol 4 (3), 215-66.

80. Beenken, A. and Mohammadi, M. (2009) The FGF family: biology, pathophysiology and therapy. Nat Rev Drug Discov 8 (3), 235-53.

81. Babina, I.S. and Turner, N.C. (2017) Advances and challenges in targeting FGFR signalling in cancer. Nat Rev Cancer 17 (5), 318-332.

82. Weiss, J. et al. (2010) Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. Sci Transl Med 2 (62), 62ra93.

83. Reis-Filho, J.S. et al. (2006) FGFR1 emerges as a potential therapeutic target for lobular breast carcinomas. Clin Cancer Res 12 (22), 6652-62.

84. Turner, N. et al. (2010) Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. Oncogene 29 (14), 2013-23.

85. Gong, S.G. (2014) Isoforms of receptors of fibroblast growth factors. J Cell Physiol 229 (12), 1887-95.

86. Sarabipour, S. and Hristova, K. (2016) Mechanism of FGF receptor dimerization and activation. Nat Commun 7, 10262.

87. Del Piccolo, N. et al. (2017) A New Method to Study Heterodimerization of Membrane Proteins and Its Application to Fibroblast Growth Factor Receptors. J Biol Chem 292 (4), 1288-1301.

88. Sarabipour, S. (2017) Parallels and Distinctions in FGFR, VEGFR, and EGFR Mechanisms of Transmembrane Signaling. Biochemistry 56 (25), 3159-3173.

89. Givol, D. and Yayon, A. (1992) Complexity of FGF receptors: genetic basis for structural diversity and functional specificity. FASEB J 6 (15), 3362-9.

90. Holzmann, K. et al. (2012) Alternative Splicing of Fibroblast Growth Factor Receptor IgIII Loops in Cancer. J Nucleic Acids 2012, 950508.

91. Zhang, X. et al. (2006) Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. J Biol Chem 281 (23), 15694-700.

92. Huang, T. et al. (2017) FGF7/FGFR2 signal promotes invasion and migration in human gastric cancer through upregulation of thrombospondin-1. Int J Oncol 50 (5), 1501-1512.

93. Warzecha, C.C. et al. (2009) ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. Mol Cell 33 (5), 591-601.

94. Warzecha, C.C. et al. (2009) The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events. RNA Biol 6 (5), 546-62.

95. Roche, J. (2018) The Epithelial-to-Mesenchymal Transition in Cancer. Cancers (Basel) 10 (2).

96. Warzecha, C.C. et al. (2010) An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. EMBO J 29 (19), 3286-300.

97. Warzecha, C.C. and Carstens, R.P. (2012) Complex changes in alternative pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT). Semin Cancer Biol 22 (5-6), 417-27.

98. Yang, Y. et al. (2016) Determination of a Comprehensive Alternative Splicing Regulatory Network and Combinatorial Regulation by Key Factors during the Epithelial-to-Mesenchymal Transition. Mol Cell Biol 36 (11), 1704-19.

99. Ishii, H. et al. (2014) Epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) suppress cancer cell motility via different mechanisms. J Biol Chem 289 (40), 27386-99.

100. Ueda, J. et al. (2014) Epithelial splicing regulatory protein 1 is a favorable prognostic factor in pancreatic cancer that attenuates pancreatic metastases. Oncogene 33 (36), 4485-95.

101. Yae, T. et al. (2012) Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. Nat Commun 3, 883.

102. Mizutani, A. et al. (2016) The Arkadia-ESRP2 axis suppresses tumor progression: analyses in clear-cell renal cell carcinoma. Oncogene 35 (27), 3514-3523.

103. Hayakawa, A. et al. (2017) Dual Roles for Epithelial Splicing Regulatory Proteins 1 (ESRP1) and 2 (ESRP2) in Cancer Progression. Adv Exp Med Biol 925, 33-40.

104. Fagoonee, S. et al. (2017) The RNA-binding protein ESRP1 promotes human colorectal cancer progression. Oncotarget 8 (6), 10007-10024.

105. Wang, X. et al. (2016) Recurrent amplification of MYC and TNFRSF11B in 8q24 is associated with poor survival in patients with gastric cancer. Gastric Cancer 19 (1), 116-27.

106. Koch, S. and Claesson-Welsh, L. (2012) Signal transduction by vascular endothelial growth factor receptors. Cold Spring Harb Perspect Med 2 (7), a006502.

107. Plein, A. et al. (2014) Neuropilin regulation of angiogenesis, arteriogenesis, and vascular permeability. Microcirculation 21 (4), 315-23.

108. Koch, S. et al. (2014) NRP1 presented in trans to the endothelium arrests VEGFR2 endocytosis, preventing angiogenic signaling and tumor initiation. Dev Cell 28 (6), 633-46.

109. Jin, Z.G. et al. (2003) Ligand-independent activation of vascular endothelial growth factor receptor 2 by fluid shear stress regulates activation of endothelial nitric oxide synthase. Circ Res 93 (4), 354-63.

110. Mitola, S. et al. (2010) Gremlin is a novel agonist of the major proangiogenic receptor VEGFR2. Blood 116 (18), 3677-80.

111. Kim, J.Y. et al. (2013) The prognostic significance of growth factors and growth factor receptors in gastric adenocarcinoma. APMIS 121 (2), 95-104.

112. Hirashima, Y. et al. (2009) Impact of vascular endothelial growth factor receptor 1, 2, and 3 expression on the outcome of patients with gastric cancer. Cancer Sci 100 (2), 310-5.

113. Brien, G.L. et al. (2016) Exploiting the Epigenome to Control Cancer-Promoting Gene-Expression Programs. Cancer Cell 29 (4), 464-476.

114. Yang, X. et al. (2010) Targeting DNA methylation for epigenetic therapy. Trends Pharmacol Sci 31 (11), 536-46.

115. Mann, B.S. et al. (2007) FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. Oncologist 12 (10), 1247-52.

116. Padmanabhan, N. et al. (2017) How to stomach an epigenetic insult: the gastric cancer epigenome. Nat Rev Gastroenterol Hepatol 14 (8), 467-478.

117. Tsai, M.M. et al. (2016) Potential Diagnostic, Prognostic and Therapeutic Targets of MicroRNAs in Human Gastric Cancer. Int J Mol Sci 17 (6).

118. Ueda, T. et al. (2010) Relation between microRNA expression and progression and prognosis of gastric cancer: a microRNA expression analysis. Lancet Oncol 11 (2), 136-46.

119. Zhang, L. et al. (2009) Genome-wide analysis of histone H3 lysine 27 trimethylation by ChIP-chip in gastric cancer patients. J Gastroenterol 44 (4), 305-12.

120. Choi, I.-S. and Wu, T.-T. (2005) Epigenetic alterations in gastric carcinogenesis. Cell Research 15, 247.

121. Park, J.H. et al. (2011) Identification of DNA methylation changes associated with human gastric cancer. BMC Med Genomics 4, 82.

122. Toyota, M. et al. (1999) CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 96 (15), 8681-6.

123. Toyota, M. et al. (1999) Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. Cancer Res 59 (21), 5438-42.

124. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 13 (7), 484-92.

125. Saxonov, S. et al. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proceedings of the National Academy of Sciences 103 (5), 1412.

126. Hackett, J.A. and Surani, M.A. (2013) DNA methylation dynamics during the mammalian life cycle. Philos Trans R Soc Lond B Biol Sci 368 (1609), 20110328.

127. Ooi, S.K. et al. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448 (7154), 714-7.

128. Feinberg, A.P. and Vogelstein, B. (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301, 89.

129. Esteller, M. (2008) Epigenetics in cancer. N Engl J Med 358 (11), 1148-59.

130. Moarii, M. et al. (2015) Changes in correlation between promoter methylation and gene expression in cancer. BMC Genomics 16, 873.

131. Guillaumet-Adkins, A. et al. (2014) Hypermethylation of the alternative AWT1 promoter in hematological malignancies is a highly specific marker for acute myeloid leukemias despite high expression levels. J Hematol Oncol 7, 4.

132. Upchurch, G.M. et al. (2016) Aberrant Promoter Hypomethylation in CLL: Does It Matter for Disease Development? Front Oncol 6, 182.

133. Soes, S. et al. (2014) Hypomethylation and increased expression of the putative oncogene ELMO3 are associated with lung cancer development and metastases formation. Oncoscience 1 (5), 367-74.

134. Yuille, M.R. et al. (2001) TCL1 is activated by chromosomal rearrangement or by hypomethylation. Genes Chromosomes Cancer 30 (4), 336-41.

135. Maekita, T. et al. (2006) High levels of aberrant DNA methylation in Helicobacter pylori-infected gastric mucosae and its possible association with gastric cancer risk. Clin Cancer Res 12 (3 Pt 1), 989-95.

136. Doorakkers, E. et al. (2018) Helicobacter pylori eradication treatment and the risk of gastric adenocarcinoma in a Western population. Gut.

137. Kim, J. et al. (2012) Promoter methylation status of VEGF receptor genes: a possible epigenetic biomarker to anticipate the efficacy of intracellular-acting VEGF-targeted drugs in cancer cells. Epigenetics 7 (2), 191-200.

138. Park, S. et al. (2007) Aberrant hypermethylation of the FGFR2 gene in human gastric cancer cell lines. Biochemical and Biophysical Research Communications 357 (4), 1011-1015.

139. Tost, J. (2016) Current and Emerging Technologies for the Analysis of the Genome-Wide and Locus-Specific DNA Methylation Patterns. Adv Exp Med Biol 945, 343-430.

140. Bujko, M. et al. (2014) Repetitive genomic elements and overall DNA methylation changes in acute myeloid and childhood B-cell lymphoblastic leukemia patients. Int J Hematol 100 (1), 79-87.

141. Yan, P. et al. (2012) Genome-wide methylation profiling in decitabine-treated patients with acute myeloid leukemia. Blood 120 (12), 2466-74.

142. Li, Y. and Tollefsbol, T.O. (2011) DNA methylation detection: Bisulfite genomic sequencing analysis. Methods in molecular biology (Clifton, N.J.) 791, 11-21.

143. Wreczycka, K. et al. (2017) Strategies for analyzing bisulfite sequencing data. J Biotechnol 261, 105-115.

144. Meissner, A. et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33 (18), 5868-77.

145. Herman, J.G. et al. (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 93 (18), 9821-6.

146. Wen, J. et al. (2017) Promoter methylation of tumor-related genes as a potential biomarker using blood samples for gastric cancer detection. Oncotarget 8 (44), 77783-77793.

147. Tahara, T. and Arisawa, T. (2015) DNA methylation as a molecular biomarker in gastric cancer. Epigenomics 7 (3), 475-86.

148. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25 (4), 402-8.

149. Mortazavi, A. et al. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5 (7), 621-8.

150. Trapnell, C. et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28 (5), 511-5.

151. TCGA (2018) Genomic Data Commons Data Portal. https://portal.gdc.cancer.gov/, (accessed).

152. Laddha, S.V. et al. (2014) Mutational landscape of the essential autophagy gene BECN1 in human cancers. Mol Cancer Res 12 (4), 485-90.

153. Duffy, M.J. et al. (2017) Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). Eur J Cancer 75, 284-298.

154. Chaffer, C.L. et al. (2006) Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2. Cancer Res 66 (23), 11271-8.

155. Guo, M. et al. (2012) FGFR2 isoforms support epithelial-stromal interactions in thyroid cancer progression. Cancer Res 72 (8), 2017-27.

156. Yoshino, M. et al. (2005) Keratinocyte growth factor receptor expression in normal colorectal epithelial cells and differentiated type of colorectal cancer. Oncol Rep 13 (2), 247-52.

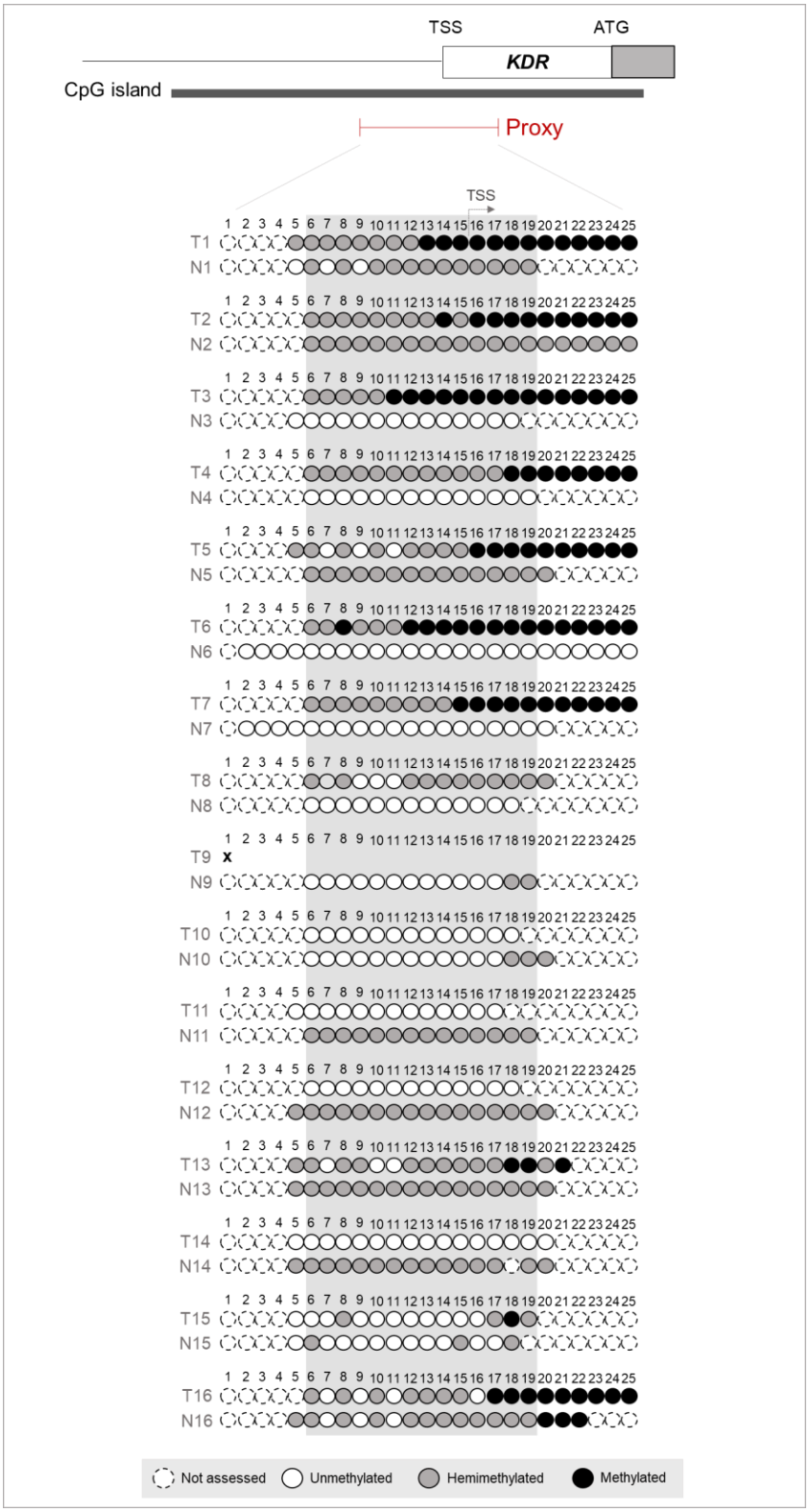
157. Bai, A. et al. (2010) GP369, an FGFR2-IIIb-specific antibody, exhibits potent antitumor activity against human cancers driven by activated FGFR2 signaling. Cancer Res 70 (19), 7630-9.

158. Catenacci, D.V.T. et al. (2017) Updated antitumor activity and safety of FPA144, an ADCC-enhanced, FGFR2b isoform-specific monoclonal antibody, in patients with FGFR2b+ gastric cancer. Journal of Clinical Oncology 35 (15_suppl), 4067-4067.

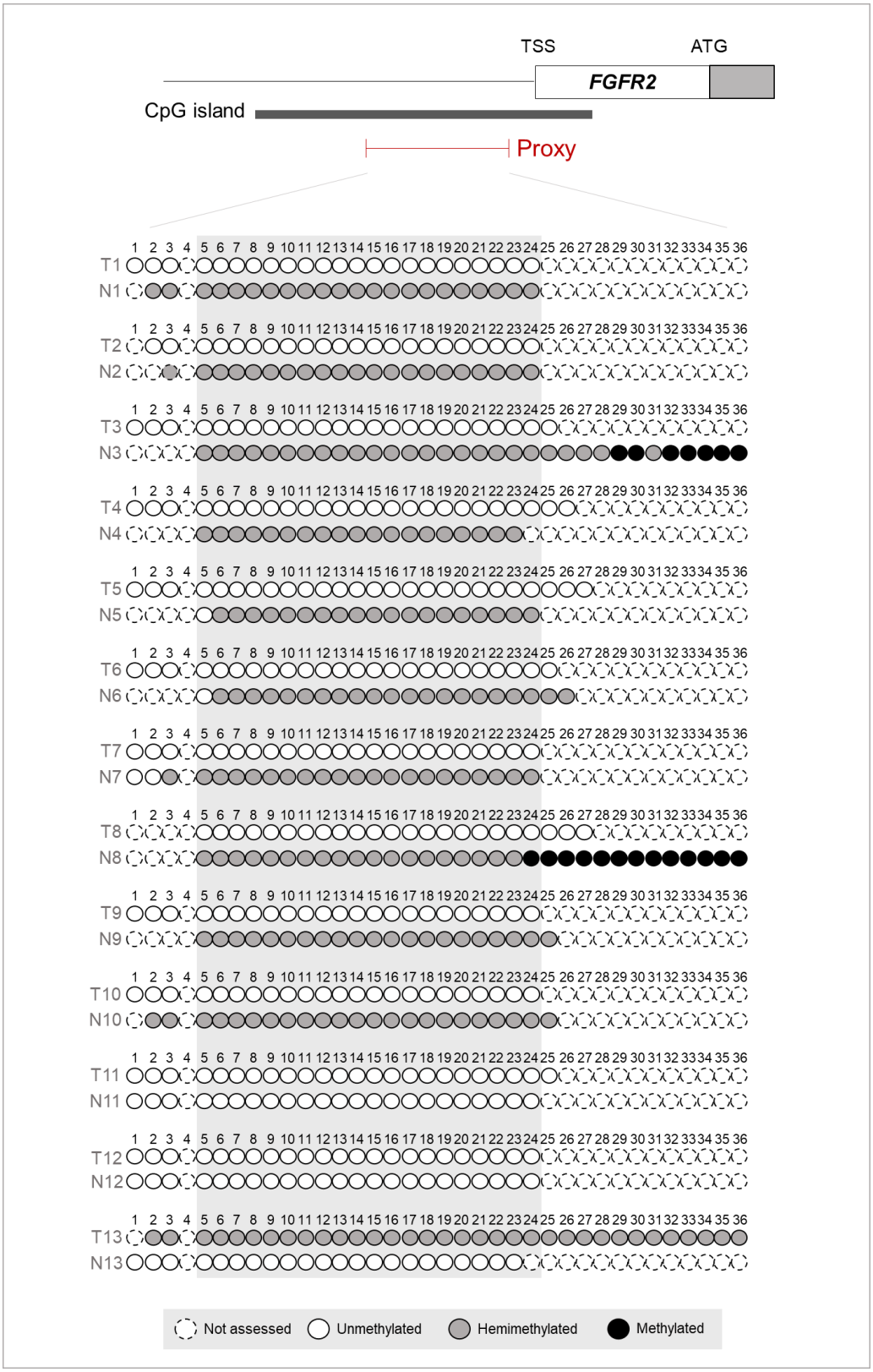
159. Pinho, S.S. et al. (2012) Loss and recovery of Mgat3 and GnT-III Mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal-epithelial transitions. PLoS One 7 (3), e33191.

160. Kunii, K. et al. (2008) FGFR2-amplified gastric cancer cell lines require FGFR2 and Erbb3 signaling for growth and survival. Cancer Res 68 (7), 2340-8.

VII. Annexes



Annex 1: Representation of the methylation status found at the *KDR* promoter proxy region in the 16 tumour (T) and adjacent normal (N) samples selected from Cohort 1. For the T9 sample it was not possible to obtain the methylation status of the *KDR* promoter due to low DNA quality. CpG dinucleotides were classified as not assessed, unmethylated, methylated or hemimethylated if, by bisulfite Sanger Sequencing, it was not possible to evaluate the CpG site, or a peak corresponding to a thymine, a cytosine or both, respectively, was obtained at each site. In grey is highlighted the area compared between tumour and normal samples.



Annex II: Representation of the methylation status found at the *FGFR2* promoter proxy region in the 13 tumour (T) and adjacent normal (N) samples selected from Cohort 1. CpG dinucleotides were classified as not assessed, unmethylated, methylated or hemimethylated if, by Bisulfite Sanger Sequencing, it was not possible to evaluate the CpG site, or a peak corresponding to a thymine, a cytosine or both, respectively, was obtained at each site. In grey is highlighted the area compared between tumour and normal samples.