Data in Brief 31 (2020) 105765

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data on the evaluation of *FGF2* gene expression in Colorectal Cancer

Helena Caiado[a,b,c], Natércia Conceição[b,c,d,*], Daniel Tiago[b,g], Ana Marreiros[c,d], Susana Vicente[e], Jose Luis Enriquez[e], Ana Margarida Vaz[f], Artur Antunes[f], Horácio Guerreiro[f], Paulo Caldeira[c,f], M. Leonor Cancela[b,c,d,h,**]

[a] *ProRegeM PhD Programme in Mechanisms of disease and regenerative medicine, University of Algarve, Faro, 8005-139, Portugal*
[b] *Center of Marine Sciences (CCMAR), University of Algarve, Faro, 8005-139, Portugal*
[c] *Department of Biomedical Sciences and Medicine, University of Algarve, Faro, 8005-139, Portugal*
[d] *Algarve Biomedical Center, University of Algarve, Faro, 8005-139, Portugal*
[e] *Pathology Department, University Hospital of Algarve, Faro, 8000-386*
[f] *Gastroenterology Department, University Hospital of Algarve, Faro, 8000-386*
[g] *Residency at University Hospital of Algarve, Faro 8000-386*
[h] *Centre for Biomedical Research, University of Algarve, Faro 8000-139, Portugal*

ARTICLE INFO

Article history:
Received 1 December 2019
Revised 14 May 2020
Accepted 19 May 2020
Available online 25 May 2020

Keywords:
Colorectal Cancer (CRC)
Matrix gla protein
Gene expression
Transcription factors
FGF2

ABSTRACT

The data presented in this article is related with the research paper entitled "Evaluation of MGP gene expression in colorectal cancer", available on Gene journal [1]. From all the transcription factors known to regulate *MGP*, FGF2 is the most described in colon adenocarcinoma and colon tumor cell lines, where it was shown to: i) contribute for the invasiveness potential; and ii) promote proliferation and survival of colorectal cancer cells. These *in vitro* studies pose the hypothesis that FGF2 associated signaling pathways could be

* Corresponding author.
** Corresponding author.
E-mail addresses: nconcei@ualg.pt (N. Conceição), lcancela@ualg.pt (M.L. Cancela).

https://doi.org/10.1016/j.dib.2020.105765

promoting the regulation of others genes, such as *MGP,* that may lead to tumor progression which ultimately could result in poor prognosis in colon adenocarcinoma.

Specifications Table

Subject	Molecular biology
Specific subject area	Colorectal cancer, Molecular biology
Type of data	Table Graph Figure
How data were acquired	qRT-PCR, SPSS
Data format	Raw Analysed
Parameters for data collection	FGF2 was shown to be both a regulator of *MGP* and an inhibitor of cellular differentiation in colorectal cancer organoids and FGF family proteins were proven to have an important role on the survival and growth of stem cells during embryogenesis, carcinogenesis and tissue regeneration
Description of data collection	*FGF2* gene expression analysis through qRT-PCR and assessment of the correlation with *MGP* gene expression and clinical and histopathological data analysis using SPSS software in colorectal patients
Data source location	University of Algarve Faro Portugal
Data accessibility	Data is available with this publication
Related research article	Caiado, H. et al. 2019 Evaluation of MGP gene expression in colorectal cancer Gene doi.org/10.1016/j.gene.2019.144120

Value of the Data

- The data presented here were obtained in order to evaluate *FGF2* gene expression in patients with colorectal cancer. This data may be of great relevance in trying to understand how *MGP* gene expression deregulation may affect patientś prognosis.
- Beneficiaries of these data are all those who seek knowledge about the molecular mechanisms that could be underlying *MGP* deregulation in tumorigenesis.
- These data report the upregulation of *FGF2* gene expression in tumor tissue and its positive correlation with *MGP* gene expression in CRC. These results could provide future insights for the search of new therapeutic targets associated with *MGP* gene expression and its deregulation in cancer.

1. Data Description

The fibroblast growth factor (FGF) signaling network has been implicated in several pathways, such as normal cell growth, differentiation, angiogenesis and tumor development [2]. The transcription factor FGF2 is one of the most studied in terms of its role in carcinogenesis including its role in tumor cell differentiation and proliferation [2]. Moreover, it is known that FGF2 induces transcription of the *MGP* gene [3].

In this report, we describe data regarding the expression analysis performed by qRT-PCR for *FGF2,* for both normal and tumor tissues, of 23 out of 33 CRC patients [1] whose samples were

Table 1
Demographic features of colorectal patients

Characteristics	*MGP* (n=23)			*FGF2* (n=23)		
	Number (%)	Mean value of fold change	*p* value	Number (%)	Mean value of fold change	*p* value
Gender			0.033			0.439
Male	14 (61)	3.135		14 (61)	2.000	
Female	9 (39)	6.648		9 (39)	1.000	
Age (median: 71,70 years)			0.548			0.776
<72	8 (35)	2.898		8 (35)	4.437	
≥72	15 (65)	5.369		15 (65)	5.640	
Familial Cancer History			0.671			0.579
Yes	7 (30)	3.034		7 (30)	3.495	
No	16 (70)	5.155		16 (70)	5.977	
Previous Pathologies			0.691			1.000
Yes	18 (78)	3.732		18 (78)	5.460	
No	5 (22)	7.308		5 (22)	4.363	
Metastasis			0.177			0.812
Yes	6 (26)	8.082		6 (26)	5.445	
No	17 (74)	3.249		17 (74)	5.143	

Mann-Whitney U test

still available, and 9 samples from the control group (Fig. 1). The data showed that the expression of *FGF2* was significantly up-regulated in CRC tissues compared to matched normal tissues (p=0.002). Our data is in accordance with what was already described in the literature regarding the increase of *FGF2* expression in various tumor tissues, such as lung [4], colorectal [5], bladder [6] and prostate [7].

To evaluate if there is a correlation between *FGF2* expression and the clinical-pathological features of the patients, we analyzed all the variables shown in Tables 1 and 2. No statistically significant associations were found between *FGF2* expression and the clinical and pathological features of the patients.

We then evaluated the correlation between *FGF2* and *MGP* expression. *FGF2* mRNA expression determined by qRT-PCR was well correlated (r=0.572, p=0.004) with that determined for *MGP* [1] (Fig. 2).

In our previously published study, we found that the two step cluster analysis of the CRC samples allowed differentiating patients with a better or worse survival outcome [1]. Subsequently, we performed a multivariate classification of two step clusters [8] to determine possible patient profiles, taking into account the characteristics of categorical and numerical variables (Table 3). This type of analysis allows the exploitation of data taking into account each variable independently from each other's, to try to identify homogeneous groups depending on their characteristics. Since we did not find any correlation between the high expression of *FGF2* and the overall patient survival rate (Fig. 3), we then evaluated the prognostic value of different variables to differentiate patients in different groups according to the influence of these factors. The variables considered were: T classification, N classification, tumor staging, gender, deceased, fold change *MGP* categorized, fold change *FGF2* categorized, fold change *MGP*, fold change *FGF2*, tumor histology, KRAS mutations, tumor location, survival rate (months), polyposis and stroke. According to this analysis, patients were divided into clusters 1 and 2. Patients in cluster 1 presented a stage N0 of lymph node metastasis (50%), the tumor was either in stage II (33.3%) or stage III (44.4%), mostly male (72.2%), with low *MGP* (72.2%) and *FGF2* (55.6%) levels of expression, with a fold change for *MGP* of 3.09 (±3.03) and for *FGF2* of 4.89 (±6.81), with a tumor histology showing either a moderately (44.4%) or well differentiated tumor (44.4%), without mutation on KRAS (61.1%), with a T3 classification (72.2%), with a mean survival time of 49.61 (±18.6) months, with the tumor mostly located in rectum (38.8%) and without the presence of polyposis (88.9) and no stroke (88.9%). Patients in cluster 2 presented a stage N1 of lymph node

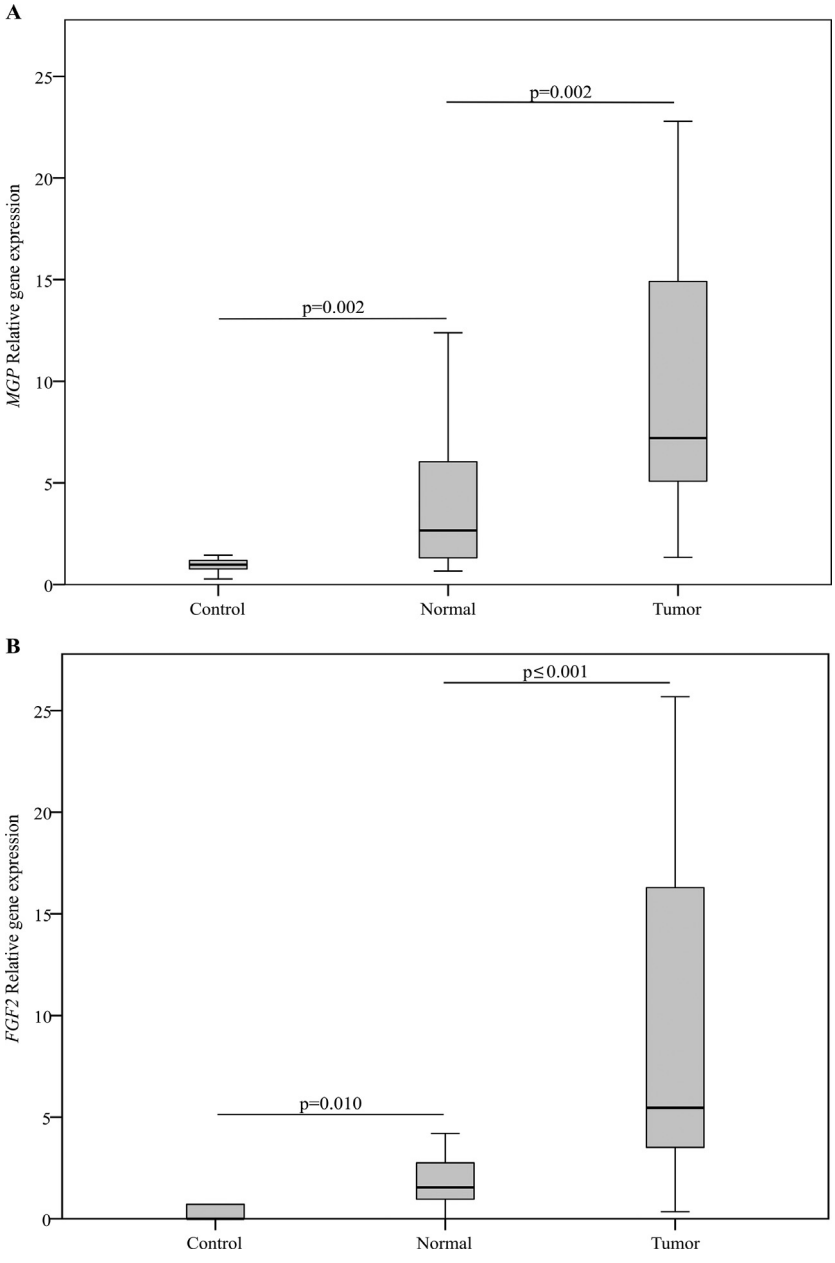

Fig. 1. Relative *MGP* and *FGF2* gene expression in samples from patients with colon adenocarcinoma. Relative *MGP* (A) and *FGF2* (B) gene expression levels were analyzed by *q*RT-PCR in a total of 9 samples from control group and 23 samples from colorectal cancer tissue (normal and tumor mucosa). The latter showed higher mRNA levels of *MGP* and *FGF2* than non-tumor tissues (*MGP* p=0.002; *FGF2* p≤0.001). Values are presented as mean ± SD. The Mann-Whitney and Kruskal Wallis non parametric tests were performed for the statistical analysis.

Table 2
Histopathological features of patients

Characteristics	*MGP* (n=23)			*FGF2* (n=23)		
	Number (%)	Mean value of fold change	*p* value	Number (%)	Mean value of fold change	*p* value
Tumor Location			0.618			0.493
Rectum	12 (52)	4.672		12 (52)	3.967	
Rectosigmoid Junction	3 (13)	6.217		3 (13)	2.479	
Ascending Colon	2 (9)	2.633		2 (9)	10.730	
Sigmoid	1 (4)	8.004		1 (4)	3.653	
Cecum	2 (9)	2.793		2 (9)	14.938	
Hepatic Angle	3 (13)			3 (13)		
Tumor Histology			0.196			0.655
Well Differentiated	10 (44)	4.014		10 (44)	3.400	
Moderately Differentiated	9 (39)	2.164		9 (39)	7.867	
Poorly Differentiated	1 (4)	24.042		1 (4)	5.530	
Mucinous	1 (4)	8.004		1 (4)	3.653	
Mucinous Well Differentiated	2 (9)	6.028		2 (9)	3.054	
Tumor Stage			0.201			0.336
I - II	9 (39)	3.155		9 (39)	3.017	
III - IV	14 (61)	5.380		14 (61)	6.639	
T classification			0.815			0.447
pT2	4 (18)	3.983		4 (18)	1.866	
pT3	18 (78)	4.763		18 (78)	5.918	
pT4	1 (4)	2.055		1 (4)	6.109	
N classification			0.372			0.592
N0	9 (39)	3.155		9 (39)	3.017	
N1	8 (35)	5.626		8 (35)	6.717	
N2	6 (26)	5.053		6 (26)	6.536	
M classification			0.227			0.745
M0	18 (78)	3.294		18 (78)	5.505	
M1	5 (22)	8.884		5 (22)	4.201	
Hepatic Metastasis			0.227			0.745
Yes	5 (22)	8.884		5 (22)	4.201	
No	18 (78)	3.294		18 (78)	5.505	
Pulmonary Metastasis			0.158			0.198
Yes	2 (9)	14.057		2 (9)	8.597	
No	21 (91)	3.600		21 (91)	4.900	
KRAS mutations			0.728			0.265
Yes	8 (35)	4.022		8 (35)	7.826	
No	15 (65)	4.770		15 (65)	3.833	

Mann-Whitney U test

metastasis (60%), the tumor was either in stage III (20%) or stage IV (80%), mostly female (80%), with high *MGP* (100%) and *FGF2* (80%) levels of expression, with a fold change for *MGP* of 9.61 (±8.4) and for *FGF2* of 6.38 (±5.0), with a well differentiated tumor histology (40%), without mutation on KRAS (80%), with a T3 classification (100%), with a mean survival time of 18.00 (±8.2) months, with the tumor located in rectum (100%) and without the presence of polyposis (100%) and no stroke (100%).

Moreover, we performed a Kaplan-Meier survival analysis to assess if *MGP* and *FGF2* could be in fact good prognostic factors in terms of overall survival rate for the two groups of patients found in the two-step cluster analysis. Patients in cluster 2, which presented a worst prognosis, had a higher mortality rate when compared with patients in cluster 1 (log-rank test $p \leq 0.001$) (Fig. 4).

From the analysis it was perceived that patients in cluster 2 had a worst prognosis, in the way that all of these patients presented a small survival rate, and higher tumor stages when compared with patients in cluster 1. It's also worthy of note, that the variables that significantly contributed to the division of the patients were the tumor staging, the presence of high level

Table 3
Multivariate analysis of predictor factors

Characteristics	Cluster 1 (n=18, %)	Cluster 2 (n=5, %)	*p* value
N Classification			p=0.126[1]
N0	9 (50)	0 (0)	
N1	5 (27.8)	3 (60)	
N2	4 (22.2)	2 (40)	
Tumor Staging			**p=0.05**[1]
Stage I	3 (16.7)	0 (0)	
Stage II	6 (33.3)	0 (0)	
Stage III	8 (44.4)	1 (20)	
Stage IV	1 (5.6)	4 (80)	
Gender			**p=0.05**[1]
Male	13 (72.2)	1 (20)	
Female	5 (27.8)	4 (80)	
Deceased			**p=0.05**[1]
No	18 (100)	0 (0)	
Yes	0 (0)	5 (100)	
Fold change *MGP* categorized			**p=0.05**[1]
High *MGP*	5 (27.8)	5 (100)	
Fold change *FGF2* categorized	8 (44.4)	4 (80)	p=0.159[1]
High *FGF2*			
Fold Change *MGP*, mean (SD[2])	3.09(±3.03)	9.61(±8.4)	**p=0.05**[3]
Fold change *FGF2*, mean (SD[2])	4.89(±6.81)	6.38(±5.00)	p=0.403[3]
MGP vs *FGF2*[5]	r=0.373;		
p=0.128	r=-0.200; p=0.747		
Tumor Histology			p=0.246[1]
Well differentiated	8 (44.4)	2 (40)	
Moderately differentiated	8 (44.4)	1 (20)	
Poorly Differentiated	0 (0)	1 (20)	
Mucinous	1 (5.6)	0 (0)	
Mucinous well differentiated	1 (5.6)	1 (20)	
KRAS mutations			p=0.433[1]
No	11 (61.1)	4 (80)	
T classification			p=0.412[1]
T1	0 (0)	0 (0)	
T2	4 (22.2)	0 (0)	
T3	13 (72.2)	5 (100)	
T4	1 (5.6)	0 (0)	
Survival Rate (Months), mean (SD[2])	49.61(±18.6)	18.00(±8.2)	**p=0.05**[4]
Tumor Location			p=0.320[1]
Rectum	7 (38.8)	5 (100)	
Rectosigmoid junction	3 (16.7)	0 (0)	
Ascending colon	2 (11.1)	0 (0)	
Sigmoid	1 (5.6)	0 (0)	
Cecum	2 (11.1)	0 (0)	
Hepatic angle	3 (16.7)	0 (0)	
Polyposis			p=0.435[1]
No	16 (88.9)	5 (100)	
Stroke			p=0.435[1]
No	16 (88.9)	5 (100)	

Boldfaced values - Variables with $p \leq 0.05$
[1] Chi Square test
[2] Standard Deviation
[3] Mann-Whitney test
[4] Log Rank test
[5] Spearman coefficient correlation test

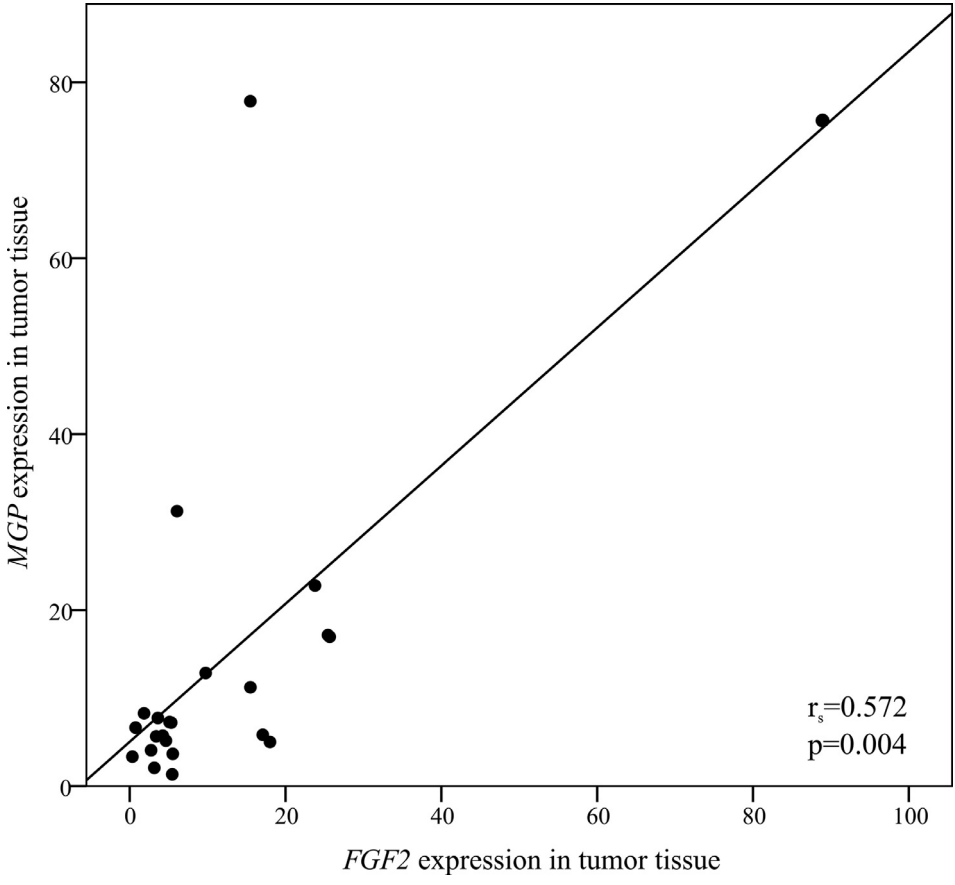

Fig. 2. Correlation between *FGF2* and *MGP* gene expression in tumor tissue. As described in experimental design in materials and methods, the correlation between *MGP* and *FGF2* gene expression was evaluated through the SPSS software, applying the Spearman coefficient correlation test in the tumor tissue and establishing a positive and significant correlation between expression of both genes (r=0.572; p=0.004).

of *MGP*, gender and the survival rate. This means that, per se, the high levels of *FGF2* alone are not sufficient for the clustering of patients, but in combination with other multiple variables can profile the patients into groups with a better or worst prognosis.

Despite the presence of some patients in cluster 1 presenting a T staging of 3 or even 4, this does not mean that these patients will actually have an associated worst prognosis. In fact, it was already shown in the literature that patients who presented a tumor stage III could have a better prognosis than those with a tumor stage II. For example, according to the American Joint Committee (AJCC) staging manual [9], when TNM staging is being evaluated, the clinicians have to take into account the tumor size (T), the number of lymph node metastasis, and the presence of metastasis. The stage is then categorized according to the combination of those three major factors, but the prognosis of the disease is reflected by its combination with other external variables that may also contribute to a worst and better prognosis. The conclusion from this analysis shows that it is the combination of the multiple variables analyzed, together with the high expression of *FGF2* in tumor tissue that can differentiate patients in two groups associated to a better or worst prognosis.

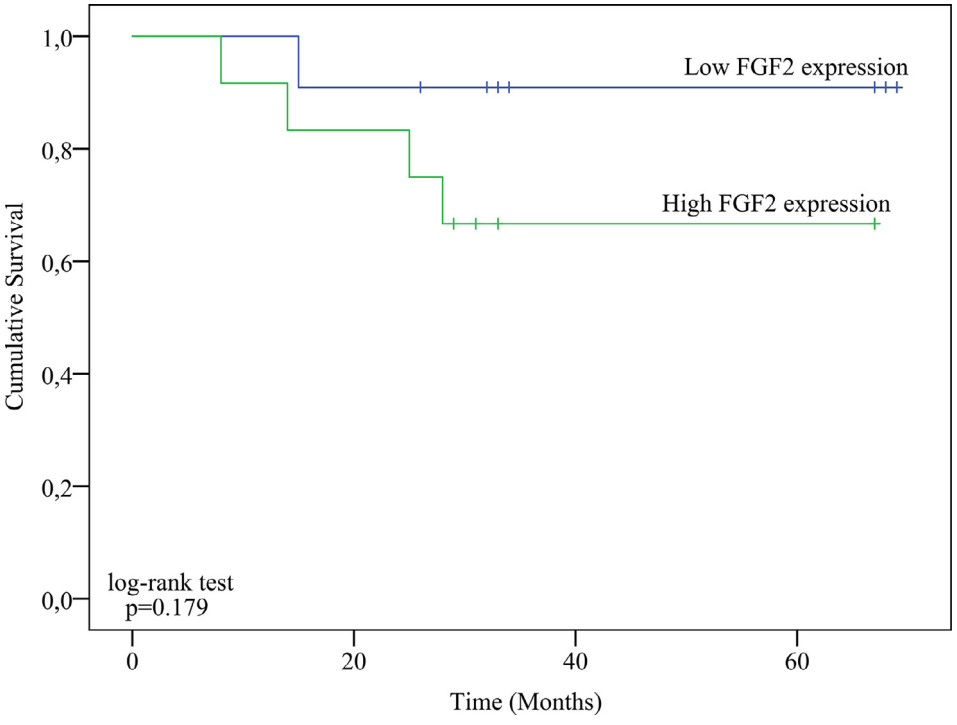

Fig. 3. Overall survival curve of patients with overexpression of *FGF2*. Patients with high *FGF2* gene expression appear to have a lower survival rate although this was not statistically significant (p=0.179). Small vertical lines indicate the censored cases referring to the number of patients that have not reached the terminal event during the data collection. *p*-value was calculated by log-rank test.

2. Experimental Design, Materials, and Methods

In this report we present briefly the materials and methods used to obtain the data here described. To see a more detailed material and methods, please refer to [1].

2.1. Clinical, demographic and pathological characteristics of patients

Tissue samples, as well as clinical and pathological information, were obtained as described in the research article "Evaluation of MGP Gene Expression in Colorectal Cancer".

Clinical, demographic and histopathological information regarding patients is depicted in tables 1 and 2.

2.2. qRT-PCR

Total RNA was extracted from fresh biopsies stored in RNALater (CRC (n=23) including normal adjacent tissue and healthy colonic tissue (n=9)). After quality and quantity measurements, cDNA synthesis was performed using 1 μg of the extracted RNA treated with RQ1 DNase (1U per μg of RNA; Promega) and M-MLV reverse transcriptase (ThermoFisher Scientific) according to manufacturer's instructions.

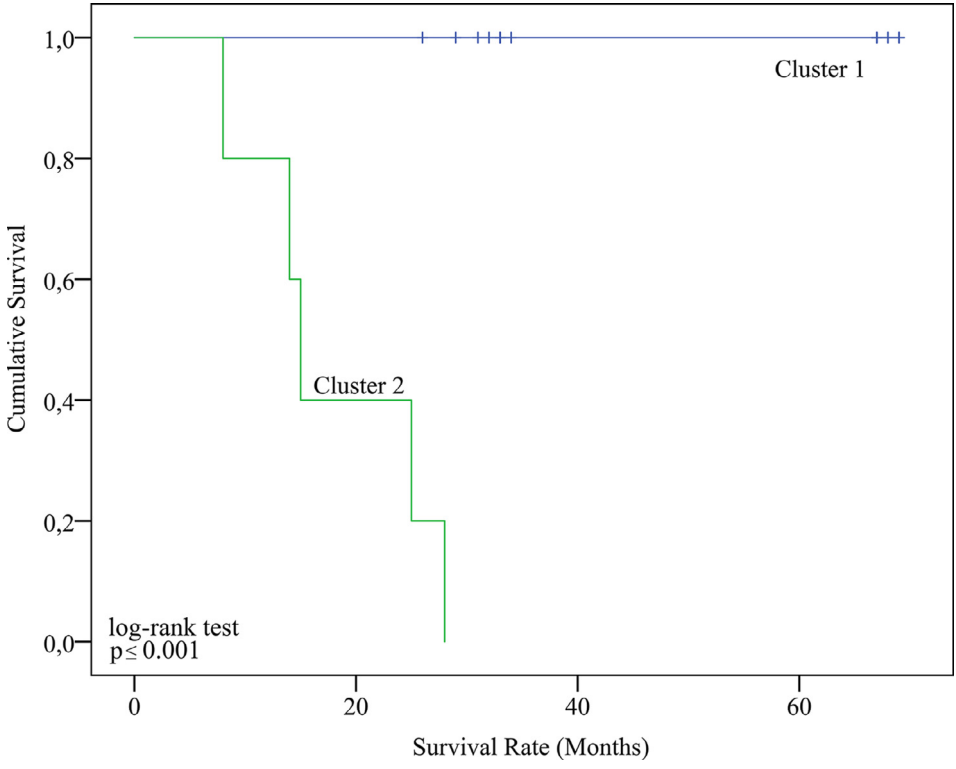

Fig. 4. Overall survival curve for patients categorized by clusters 1 and 2. Patients in Cluster 1 present a better survival rate, when compared with patients in cluster 2, who have a lower survival rate and a worse prognosis. Small vertical lines indicate the censored cases referring to the number of patients that have not reached the terminal event during data collection. *p*-value was calculated by log-rank test.

The expression of mRNA for *FGF2* was analyzed by $2^{-\Delta\Delta Ct}$ method and normalized with the expression of glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as reference gene. Primer sequences for *GAPDH* and *FGF2* were as follows: *GAPGH*: forward: 5'-TCAACGGATTTGGTCGTATTGGGCG-3' and reverse: 5'-CTCGCTCCTGGAAGATGGTGATGGG-3'; *FGF2*: forward: 5'-CAAAAACGGGGGCTTCTTCCTG-3' and reverse: 5'-CCATCTTCCTTCATAGCCAGGTAACG-3'.

Data were presented as the relative quantity of target mRNA normalized with *GAPDH* and relative to the mean expression of the control group. Please refer to the research article "Evaluation of MGP Gene Expression in Colorectal Cancer" for the analyses of expression of mRNA for *MGP* [1].

2.3. Statistical analysis

Statistical analysis was performed using SPSS software program version 25. Values for gene expression are presented as mean and standard deviation (SD) and two-sided *P* value less than 0.05 was defined as statistically significant. Fold changes presented correspond to the ratio of the values from tumor mucosa *versus* normal mucosa. Comparisons between group variables and gene expression were estimated using non parametric statistical tests: Mann–Whitney U and Kruskal–Wallis.

The cutoff value to distinguish the patients with low and high *MGP* and *FGF2* levels were estimated taking into account the median value of the fold change for both *MGP* and *FGF2*.

A multivariate classification of two step clusters [8] was performed to determine possible patient profiles, taking into account the characteristics of categorical and numerical variables (Table 3). This allowed the formation of cluster 1 (n=18) and cluster 2 (n=5). Spearman coefficients were considered to analyze the correlation between *MGP* and *FGF2* fold change values by the interest groups, namely, clusters and tissue samples. Overall survival probability for two groups of patients (clusters 1 and 2) was calculated using the Kaplan–Meier method; intergroup differences were determined using a log-rank test. Logistic regression analysis and χ^2 analysis were used to evaluate the independent influence of factors on the final prognosis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

This research was supported in part by national funds from the Portuguese Science and Technology Foundation (FCT), under the project UID/Multi/04326/2019 (CCMAR) and by Sociedade Portuguesa de Gastrenterologia through the awarded projected entitled “Insights into Matrix Gla protein (MGP) regulation in colorectal cancer” (Grant SPG 1/2015). During the data collection NC and DT where supported by a FCT fellowship (grant numbers: SFRH/BPD/111898/2015 and SFRH/BPD/111289/2015 respectively). HC is supported by a doctoral fellowship (Grant number: PD/BD/128341/2017) from FCT.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.105765.

References

[1] H. Caiado, et al., Evaluation of MGP gene expression in Colorectal Cancer, Gene 723 (October 2019) 144120, 2019.
[2] M.R. Akl, et al., Molecular and clinical significance of fibroblast growth factor 2 (FGF2/bFGF) in malignancies of solid and hematological cancers for personalized therapies, Oncotarget 7 (28) (2016) 44735–44762.
[3] C. Stheneur, et al., Basic fibroblast growth factor as a selective inducer of matrix Gla protein gene expression in proliferative chondrocytes, Biochem. J. 369 (1) (Jan. 2003) 63–70.
[4] L. Li, et al., FGF2 and FGFR2 in patients with idiopathic pulmonary fibrosis and lung cancer, Oncol. Lett. 16 (2) (2018) 2490–2494.
[5] M.L. George, et al., Plasma basic fibroblast growth factor levels in colorectal cancer: A clinically useful assay? Clin. Exp. Metastasis 19 (8) (2002) 735–738.
[6] P. Gazzaniga, et al., Detection of basic Fibroblast Growth Factor mRNA in urinary bladder cancer: Correlation with local relapses, Int. J. Oncol. 14 (6) (1999) 1123–1127.
[7] N. Soulitzis, et al., Expression analysis of peptide growth factors VEGF, FGF2, TGFB1, EGF and IGF1 in prostate cancer and benign prostatic hyperplasia, Int. J. Oncol. 29 (2) (2006) 305–314.
[8] T. Chiu, et al., A robust and scalable clustering algorithm for mixed type attributes in large database environment, Proc. seventh ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '01 (2001) 263–268.
[9] M.R. Weiser, AJCC 8th Edition: Colorectal Cancer, Ann. Surg. Oncol. 25 (6) (2018) 1454–1455.