

## Extracellular matrix gene expression profiling using microfluidics for colorectal carcinoma stratification

Christopher J. Hayes,<sup>1,2</sup> Catriona M. Dowling,<sup>1,2,3</sup> Susan Dwane,<sup>1</sup>  
Mary E. McCumiskey,<sup>1,4</sup> Shona M. Tormey,<sup>4</sup> B. Anne Merrigan,<sup>4</sup>  
John C. Coffey,<sup>2,3,4,5</sup> Patrick A. Kiely,<sup>1,2,3,5</sup> and Tara M. Dalton<sup>1,a)</sup>

<sup>1</sup>Stokes Laboratories, Bernal Institute, University of Limerick, Limerick, Ireland

<sup>2</sup>Graduate Entry Medical School, University of Limerick, Limerick, Ireland

<sup>3</sup>Health Research Institute, University of Limerick, Limerick, Ireland

<sup>4</sup>Department of Surgery, University Hospital Limerick, Limerick, Ireland

<sup>5</sup>4i Centre for Interventions in Infection, Inflammation and Immunity, Graduate Entry Medical School, University of Limerick, Limerick, Ireland

(Received 16 August 2016; accepted 14 October 2016; published online 31 October 2016)

In cancer, biomarkers have many potential applications including generation of a differential diagnosis, prediction of response to treatment, and monitoring disease progression. Many molecular biomarkers have been put forward for different diseases but most of them do not possess the required specificity and sensitivity. A biomarker with a high sensitivity has a low specificity and vice versa. The inaccuracy of the biomarkers currently in use has led to a compelling need to identify more accurate markers with diagnostic and prognostic significance. The aim of the present study was to use a novel, droplet-based, microfluidic platform to evaluate the prognostic value of a panel of thirty-four genes that regulate the composition of extracellular matrices in colorectal carcinoma. Our method is a novel approach as it uses using continuous-flowing Polymerase Chain Reaction for the sensitive detection and accurate quantitation of gene expression. We identified a panel of relevant extracellular matrix genes whose expression levels were measured by real-time quantitative polymerase chain reaction using Taqman<sup>®</sup> reagents in twenty-four pairs of matched colorectal cancer tumour and associated normal tissue. Differential expression patterns occurred between the normal and malignant tissue and correlated with histopathological parameters and overall surgical staging. The findings demonstrate that a droplet-based microfluidic quantitative PCR system enables biomarker classification. It was further possible to sub-classify colorectal cancer based on extracellular matrix protein expressing groups which in turn correlated with prognosis. © 2016 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). [<http://dx.doi.org/10.1063/1.4966245>]

### INTRODUCTION

Colorectal cancer is a significant cause of mortality worldwide with more than 746 000 cases in men and 614 000 cases in women diagnosed annually.<sup>1</sup> The biomarkers used in colorectal cancer mainly aim to either (a) aid in determining suitability for chemotherapy or (b) accurately determine prognosis. Biomarkers are sub classified into clinical (emergency versus elective presentation), serologic (tumour antigens such as carcinoembryonic antigen), or histologic (including nodal status, lymph node ratio, lymphovascular invasion, and perineural invasion). Nevertheless, these biomarkers have limited utility as they are not accurate at the individual level. To overcome this, genetic-based biomarkers have been characterised. These

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: Tara.Dalton@ul.ie. Telephone: +00-353-61202961.

include microRNAs, single nucleotide polymorphisms, and RNA-based determination of gene expression. The convergence of these efforts with the development of array-based technologies has led to the development of *omics*, or high-throughput based, biomarkers. However, these are hampered by lack of consensus between laboratories and are not in routine clinical use given costs and other factors. Currently, the consensus is that biomarkers have failed to yield a valid basis for personalised medicine and so the approach to their development must be refreshed.

Since its introduction in 1992,<sup>2</sup> qPCR-based technologies have been widely exploited in the evaluation of biomarkers,<sup>3–5</sup> where potential markers are measured with greater accuracy to yield gene sets or signatures that can distinguish between a predefined defined subsets of tumours.<sup>6</sup> Micro total analysis systems, such as those utilising droplet microfluidics, provide clinical benefits in the setting of colorectal cancer by performing higher throughputs with enhanced data analysis capabilities. Furthermore, the quantity of available tissue has become a critical factor in *-omic* studies as we are now detecting cancer in polyps. Polyps can easily be removed but we cannot accurately determine the probability that the tumour will have spread to nodes within the associated mesentery. This spread is thought to be a critical step in metastasis development. As the risk of spread is unknown, subjecting patients to resection of the segment of colon or mesentery can expose patients to increased and unnecessary risks of complications or death. Therefore, by scaling of the reaction from the microliter down to the nanolitre range dramatically reduces the required volume of reagents and samples,<sup>7</sup> reducing the overall cost per reaction and conserving the clinical samples. Histologically, cancer can be diagnosed in a few hundred cells, so developing an approach to determine nodal status accurately with minimal tissue allows that these technologies can be preferred to obtain improved biological and translatable results.

To date, we have extensively characterised the role of several biomarkers in colorectal malignancy. Our research group has recently identified significant prognostic biomarkers in colorectal cancer<sup>8,9</sup> and established consensus profiles and associations with prognostic and predictive properties in colon and rectal adenocarcinomas.<sup>10,11</sup> Therefore, the hypothesis of this study is to use microfluidic droplets acting as distinct miniature reactors to identify differentially expressed extracellular-matrix (ECM) genes in colorectal cancer tissue and matched patient normal tissue to determine if differential expression could be correlated with an increased metastatic potential. The extracellular-matrix (ECM) has important roles in regulating normal cell and tissue development and function.<sup>12,13</sup> ECM components constantly interact with epithelia as ligands for receptors including integrins. In this manner, they instigate intracellular activities related to a vast array of biologic functions, including tissue development and homeostasis.<sup>14</sup> In fact, the ECM influences almost all cellular behaviour and is critical during major developmental processes.<sup>15,16</sup> However, in solid organ malignancy, ECM-based molecular and cellular interactions can be dramatically altered in a manner that potentiates metastasis development.<sup>17</sup> The metastatic process requires ECM remodelling, basement membrane degradation, and enhanced tumour cell motility.<sup>18</sup> Cell shape, function, proliferation, migration, and apoptosis are all influenced by cell-ECM interactions. These relationships are altered in solid organ malignancies. Therefore, the identification of prominent ECM tumour markers that derive biological insight into tumour development and progression would be of clinical value.

## MATERIALS AND METHODS

**Patients and Tissue Samples:** Colorectal cancer tissue and paired adjacent physiologically normal mucosa tissue samples were collected (ethical approval number 73/11) from a heterogeneous group of 24 patients at different stages of malignancy, undergoing colorectal resection at University Hospital Limerick. Primarily, the tumours are adenocarcinomas. After removal of the colon and associated mesentery, the specimen was opened. Mucosal samples were taken from the tumour and normal mucosa at a distance of 10 cm from the tumour. Samples were immediately placed in Allprotect<sup>®</sup> Tissue Reagent (Qiagen, Hilden, Germany) and stored in a  $-80^{\circ}\text{C}$  freezer.

**RNA extraction and cDNA Synthesis:** Frozen tissue was immersed in liquid nitrogen and ground into powder. Lysis buffer was added to tissue and the sample transferred to tubes. Total ribonucleic acid (RNA) was extracted using RNeasy Plus Mini Kit<sup>®</sup> (Qiagen, Hilden, Germany) according to manufacturer's instructions. RNA was evaluated for purity on the NanoDrop<sup>™</sup> 1000 Spectrophotometer 3.7 (Thermo Fisher, Massachusetts, United States). RNA has its absorption maximum at 260nm and the ratio of the absorbance at 260nm and 280nm is used to assess the RNA purity of an RNA preparation. RNA quality was evaluated through visualization of the 28S:18S ribosomal RNA ratio on a 1% agarose gel. cDNA synthesis was performed using Superscript<sup>®</sup> VILO cDNA Synthesis Kit (Thermo Fisher, Carlsbad, USA) according to the manufacturer's protocol. Reactions were incubated for 25 °C for 10 min followed by 60 °C for 45 min and 72 °C for final denaturation for 15 min. Samples were stored at −20 °C.

**Endogenous reference controls:** Relative quantification is based on the expression levels of a target gene versus one or more reference gene(s), and in many experiments it is adequate for investigating physiological changes in gene expression levels.<sup>19</sup> Peptidylprolyl Isomerase A (PPIA) and Hypoxanthine guanine phosphoribosyl transferase 1 (HPRT1) have been showed previously to be stably expressed in a large cohort of colorectal tissues,<sup>20,21</sup> and so were chosen to normalise expression values in the present study. The use of multiple reference genes increases the accuracy of quantitation compared to the use of a single reference gene, especially when the aim is to show relatively small fold-changes in RNA levels.<sup>22</sup> Further to this, our research group have characterised the expression level of nine candidate reference genes and determined their stability in malignant and normal colorectal tissue.<sup>23</sup> From this, Phosphoglycerate Kinase 1 (PGK1), Glucuronidase Beta (GUSB), and PPIA were ranked as the most stable genes between normal and cancerous tissue. Therefore, PGK1 and GUSB were used as additional reference genes. A list of selected reference genes can be seen in Table I.

**qPCR of mRNA:** The instrument that has been developed on which the PCR will be performed<sup>24</sup> is a droplet based microfluidic platform that performs qPCR in a continuous flowing process. The reaction droplets (~300nl) are surrounded by an immiscible oil to prevent contamination and carryover. Micro-scale polytetrafluoroethylene (PTFE) capillary tubing is integrated into dipping heads that are attached to modular robotic stages. The stages programmed sequence allows dipping into the solutions and aspirates predefined quantities of patient cDNA and nuclease free water (35nl:100nl), gene expression assay primer-probe mix (15nl) (TaqMan<sup>®</sup> Gene Expression Assays, Thermo Fisher, Massachusetts, United States), and Mastermix (150nl) (TaqMan<sup>®</sup> Gene Expression Master Mix<sup>®</sup>-Thermo Fisher, Massachusetts, United States). The individual droplets are then combined at the liquid bridge which consists of capillaries where the droplets are suspended before merging in a free-balanced system.<sup>25</sup> The single PCR reactions are then conveyed through the system to the thermal zone. Thermal cycling is performed in a flowing serpentine cycler, which gives superior thermal performance because of the relative size of the droplet resulting in more sensitive experiments. The temperature blocks of 95 °C (denaturation) and 65 °C (annealing) are isothermal to prevent a non-uniform temperature field which may lead to low amplification efficiency of nucleic acids and even non-specific PCR products due to insufficient annealing temperature of the PCR process.<sup>26</sup> Each cycle through a serpentine performs one cycle of PCR and there are 40 cycles in total. To detect the droplet signal a bio-fluorescence detection platform, comprising of reflectance probes, fibre optical cores, and high resolution cameras, analyse each microfluidic reactor droplet. From this data,

TABLE I. Reference genes used for normalisation.

Gene name	Gene symbol	Location	Assay ID	bp	References
Phosphoglycerate kinase 1	PGK1	Xq13.3	Hs00943178_g1	73	111, 21, and 23
Peptidylprolyl isomerase A	PPIA	7p13	Hs04194521_s1	97	
Glucuronidase beta	GUSB	7q21.11	Hs00939627_m1	96	
Hypoxanthine guanine phosphoribosyl transferase 1	HPRT1	Xq26.1	Hs02800695_m1	82	

the amplification *s*-curves for each sample are generated. A baseline is then set and the *s*-curve is post processed to output a threshold cycle (Ct) value or, according to the MIQE guidelines,<sup>27</sup> a quantification cycle (Cq) value, which is indicative of the number of RNA targets in the initial samples. The Cq is determined from a log-linear plot of PCR signal versus the cycle number. The background levels are normally determined from the first triplicate no-template-control (NTC) droplets in a run. This is subtracted from the subsequent ROX reference dye normalisation value. The  $\Delta\Delta Cq$  method is currently the most commonly used method for studies that report changes in the expression of genes of interest relative to a reference gene.<sup>28</sup> For studies performed on the platform, the relative expression levels of the target genes are calculated using the Livak and Schmittgen method<sup>29</sup> using the formula  $2^{-(\Delta\Delta Cq)}$  where

$$\Delta\Delta Cq = [(Cq \text{ target gene}_{\text{normal sample}}) - (Cq \text{ reference gene}_{\text{normal sample}})] \\ - [(Cq \text{ target gene}_{\text{cancer sample}}) - (Cq \text{ reference gene}_{\text{cancer sample}})].$$

**Statistical Analysis:** Statistical analysis was carried out using *IBM SPSS Statistics 21* (SPSS, Inc.) The mean of triplicates was used for statistical analysis where the mean difference between the matched samples, the standard deviation, standard error, and 95% confidence interval of this mean difference is evaluated, respectively. The distribution of the data is tested for normality using histograms for both normal and tumour samples and also using the Shapiro-Wilk test (used where  $n < 50$ ) for normality to determine whether the distribution of the differences in the dependent variable between the two related groups was approximately normally distributed. Whisker plots were generated and interpreted to indicate if there were any “outliers” in the data sets. Outliers were removed on the criteria of being two times the standard deviation of the dataset. Parametric tests (Paired-sample student *t*-tests) and non-parametric tests (Wilcoxon signed-rank) were used to assess related samples based on the distribution of the data. If homogeneity of variances was violated, then separate parametric and non-parametric tests were performed to assess effect variations. A two-tailed *p*-value  $< 0.05$  was considered statistically significant. Correlations between parameters and ECM gene expression levels in the samples analysed were determined by Pearson product-moment correlation coefficient, *r*. *P* values  $< 0.05$  were considered statistically significant.

## RESULTS

**Gene Selection criteria:** A detailed analysis of extracellular matrix literature reviews and published colorectal cancer gene expression profiling studies was undertaken to identify key ECM genes involved in colorectal cancer formation and development. Genes that code for constituents such as integrins, collagens, proteases, and growth factors were selected to be included in the panel. Furthermore, an analysis of current ECM gene expression arrays such as the Taqman<sup>®</sup> and Qiagen<sup>®</sup> ECM and Adhesion Molecule arrays (Catalogue number: 4414133 and PAHS-013ZA)<sup>30,31</sup> was reviewed to support hypotheses developed for which genes could be most influential in the metastatic spread of the disease. A list of the selected genes for analysis in this study can be seen in Table II with references for studies which show specific genetic dysregulation.

## DIFFERENTIAL GENE EXPRESSION IN MATCHED SAMPLES

RNA expression levels of the panel of thirty-four ECM genes were measured by qPCR in the twenty-four malignant tissue samples and their corresponding normal tissue. The proportion of tissue represented by epithelium or stroma (i.e., the epithelial-stromal ratio) can affect the expression of genes. Expression levels of Cytokeratin-18 (KRT18) and Vimentin (VIM) were assessed by qPCR to approximate the epithelial and stromal tissue components of each sample<sup>9,32</sup> with average composition for the twenty four samples shown to be consistent and can be seen in Figure 1. Differential gene expression greater than two-fold difference between normal and malignant tissue was observed in twenty-two genes. The differences identified in expression

TABLE II. Candidate cancer-related genes that may play roles in regulating the composition of the ECM and be involved in the progression and development of cancer.

Gene name	Gene symbol	Location	Assay ID	bp	References
Cadherin-2	CDH2	18q11.2	Hs00983056_m1	66	65 and 66
Carcinoembryonic antigen-related cell adhesion molecule 1	CEACAM1	19q13.2	Hs00989786_m1	75	58 and 67
CD9 Molecule	CD9	12p13.3	Hs00233521_m1	72	68 and 69
Chemokine (C-X-C motif) receptor 2	CXCR2	2q35	Hs01891184_s1	64	70 and 71
Collagen, type I, alpha 1	COL1A1	17q21.33	Hs00164004_m1	66	30 and 72
Collagen, type III, alpha 1	COL3A1	2q31	Hs00943809_m1	65	50 and 73
Epidermal growth factor	EGF	4q25	Hs01099999_m1	70	74 and 75
Epidermal growth factor receptor	EGFR	7p12	Hs01076090_m1	57	76 and 77
Insulin-like growth factor 1	IGF1	12q23.2	Hs01547656_m1	68	78 and 79
Insulin-like growth factor 1 receptor	IGF1R	15q26.3	Hs00609566_m1	64	55, 80, and 81
Insulin-like growth factor 2	IGF2	11p15.5	Hs04188276_m1	83	82 and 83
Insulin-like growth factor 2 receptor	IGF2R	6q26	Hs00974474_m1	59	84 and 85
Insulin-like growth factor binding protein 2	IGFBP2	2q35	Hs01040719_m1	54	86 and 87
Insulin-like growth factor binding protein 3	IGFBP3	7p12.3	Hs00365742_g1	79	87 and 88
Insulin-like growth factor binding protein 4	IGFBP4	17q21.2	Hs01057900_m1	81	87
Integrin, alpha 2	ITGA2	5q11.2	Hs00158127_m1	67	77 and 89
Integrin, alpha 5	ITGA5	12q11-q13	Hs01547673_m1	54	50 and 90
Integrin, alpha 8	ITGA8	10p13	Hs00233321_m1	89	30
Integrin, alpha V	ITGAV	2q31-q32	Hs00233808_m1	64	77 and 90
Integrin, beta 1	ITGB1	10p11.2	Hs00559595_m1	75	30 and 91
Integrin, beta 4	ITGB4	17q25	Hs00173995_m1	58	30 and 92
Integrin, beta 5	ITGB5	3q21.2	Hs00174435_m1	78	30 and 50
Integrin, beta-like 1	ITGBL1	13q33	Hs01557019_m1	71	31
Interleukin-8	IL8	4q13-q21	Hs00174103_m1	101	93 and 94
Laminin, alpha 1	LAMA1	18p11.3	Hs00300550_m1	77	30 and 95
Matrix metalloproteinase 2	MMP2	16q12.2	Hs01548727_m1	65	96–98
Matrix metalloproteinase 3	MMP3	11q22.3	Hs00968305_m1	126	93 and 99
Matrix metalloproteinase 7	MMP7	11q22.2	Hs01042796_m1	64	39 and 100
Matrix metalloproteinase 9	MMP9	20q13.12	Hs00234579_m1	54	101 and 102
Periostin, osteoblast specific factor	POSTN	13q13.3	Hs01566750_m1	67	103 and 104
Protein tyrosine kinase 2	PTK2	8q24.3	Hs01056457_m1	76	105 and 106
Receptor of activated protein kinase C	RACK1	5q35.3	Hs00272002_m1	66	107 and 108
Secreted protein, acidic, cysteine-rich	SPARC	5q31.3-q32	Hs00234160_m1	76	90 and 109
Vascular endothelial growth factor A	VEGFA	6p12	Hs00900055_m1	59	109 and 110

were statistically significant in the case of fourteen genes (Figure 2). These genes have varying biological functions and code for cell surface proteins, growth factors, and proteins involved in tissue remodelling and cell adhesion. The greatest difference in gene expression was observed in *MMP7* (+31 fold,  $p < 0.001$ ) between the matched tumour and normal mucosa in this study. *IL-8* expression was significantly up-regulated (+18-fold,  $p < 0.001$ ) and *ITGBL1* expression was shown to be increased (+11-fold,  $p = 0.003$ ) in malignant versus matched normal tissue. *MMP9* expression was observed to be upregulated (+7-fold,  $p = 0.02$ ) in the malignant tissue as was *COL1A1* expression (+6-fold,  $p = 0.039$ ) and *ITGA2* (4-fold,  $p = 0.007$ ). *CXCR2* (4-fold,  $p = 0.022$ ), *SPARC* (3.5-fold,  $p = 0.033$ ), and *RACK1* (2.5 fold,  $p = 0.019$ ) were also identified as being upregulated in the matched comparison. Conversely, decreased expression levels were observed in six of the genes in the ECM panel. *IGF-1* expression levels were lower (−8-fold,  $p = 0.007$ ) in cancer tissue than in the normal tissue. Decreases in

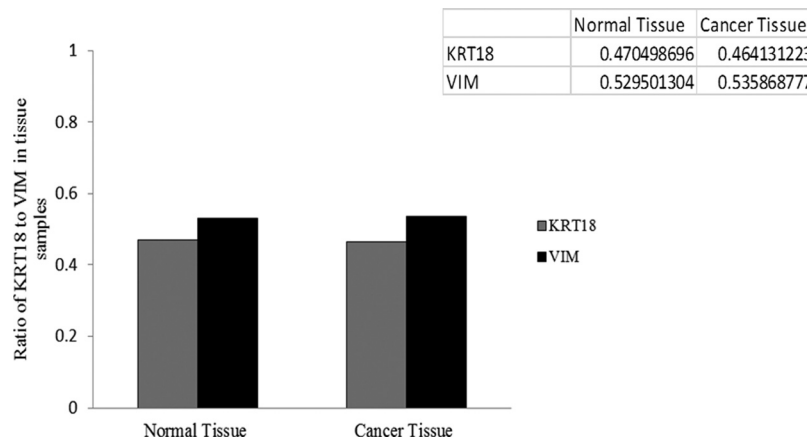


FIG. 1. Figure shows the determination of epithelial and stromal content of the samples. Legend shows the average percentage of each in the normal and cancer samples. Minimal difference in levels of each is observed.

expression levels for *EGF* ( $-5$  fold,  $p=0.014$ ) and *EGFR* ( $-3$ -fold,  $p=0.05$ ) were also observed. *CEACAMI* was also identified as being decreased ( $-4$ -fold,  $p=0.048$ ). *ITGA8* gene expression levels were shown to be downregulated ( $-3.6$ -fold,  $p=0.003$ ) in the cohort. This information is tabulated in Table V.

#### ANALYSIS OF EXPRESSION COORDINATION AND GENE NETWORKING

After the initial gene expression levels were determined by experimental means, gene pairs which had similar expression patterns from the fourteen gene panel were sought. The expression levels in the 24 patient cohort were assessed using a Pearson's correlation coefficient test for significance ( $p$ -value  $< 0.05$ ). The strongest positive relationships were observed between *COL1A1* and *SPARC* expression ( $\rho=0.763$ ,  $p$ -value  $= < 0.001$ ) and *COL1A1* and *RACK1* ( $\rho=0.451$ ,  $p$ -value  $= 0.032$ ). *MMP7* and *ITGA2* also showed a correlation ( $\rho=0.552$ ,  $p$ -value  $= 0.018$ ). Correlations between *ITGBL1* and both *SPARC* expression ( $\rho=-0.720$ ,  $p$ -value  $= 0.001$ ) and *EGF* ( $\rho=-0.699$ ,  $p$ -value  $= < 0.001$ ) also emerged as being statistically significant. A strong positive correlation was observed between *EGF* and its receptor *EGFR* expression ( $\rho=-0.470$ ,  $p$ -value  $= 0.032$ ). The associations between the significant gene expression arrays are shown in Table III.

Subsequently, the patient cohort was stratified according to UICC staging to identify further differential expression patterns (Figure 3). The cohort was categorised based on the histopathological diagnoses as follows; a) Group 1; benign disease samples including diverticulosis, hyperplasia, tubulovillous adenoma (TVA) polyps and Stage I samples ( $n=6$ ) b) Group 2; stage II carcinomas ( $n=11$ ) c) Group 3; stage III and stage IV advanced carcinomas ( $n=7$ ). Gene expression was compared between malignant and matched normal tissue. The number of genes differentially expressed increased sequentially from Group 1 to Group 3. **Group 1;** ECM gene dysregulation between the normal and cancer samples for patients was noted only in an upregulation of *IL-8* (15-fold,  $p=0.05$ ). **Group 2;** Statistically significant up-regulation was observed for *MMP7* (48 fold,  $p=0.004$ ), *IL-8* (17-fold,  $p=0.015$ ), *COL1A1* (8.7 fold,  $p=0.049$ ), *LAMAI* (11-fold,  $p=0.033$ ), *ITGBL1* (8-fold,  $p=0.016$ ), *MMP9* (5.2 fold,  $p=0.05$ ), *ITGA2* (5.1 fold,  $p=0.039$ ), and *COL3A1* (4-fold,  $p=0.041$ ). Genes that were observed to be statistically downregulated for Group 2 samples included *EGF* ( $-7.4$  fold,  $p=0.004$ ), *IGF1* ( $-4.7$  fold,  $p=0.034$ ), and *CEACAMI* ( $-4.5$ -fold,  $p=0.044$ ). **Group 3;** Upregulation of *IL-8* (25 fold,  $p=0.002$ ), *ITGBL1* (22.5-fold,  $p=0.021$ ), *MMP7* (17.8-fold,  $p=0.035$ ), *SPARC* (5.6 fold,  $p=0.05$ ), *ITGA2* (5.1 fold,  $p=0.029$ ), *RACK1* (5 fold,  $p=0.006$ ), and *ITGAV* (2 fold,  $p=0.012$ ). Downregulation of *IGF1* ( $-11.7$  fold,  $p=0.04$ ), *LAMAI* ( $-9.4$  fold,  $p=0.08$ ), *ITGA8* ( $-7.8$  fold,  $p=0.048$ ), and *EGF* ( $-7$  fold,  $p=0.03$ ). Patient demographics, including histopathological diagnoses, are summarised in Table IV.

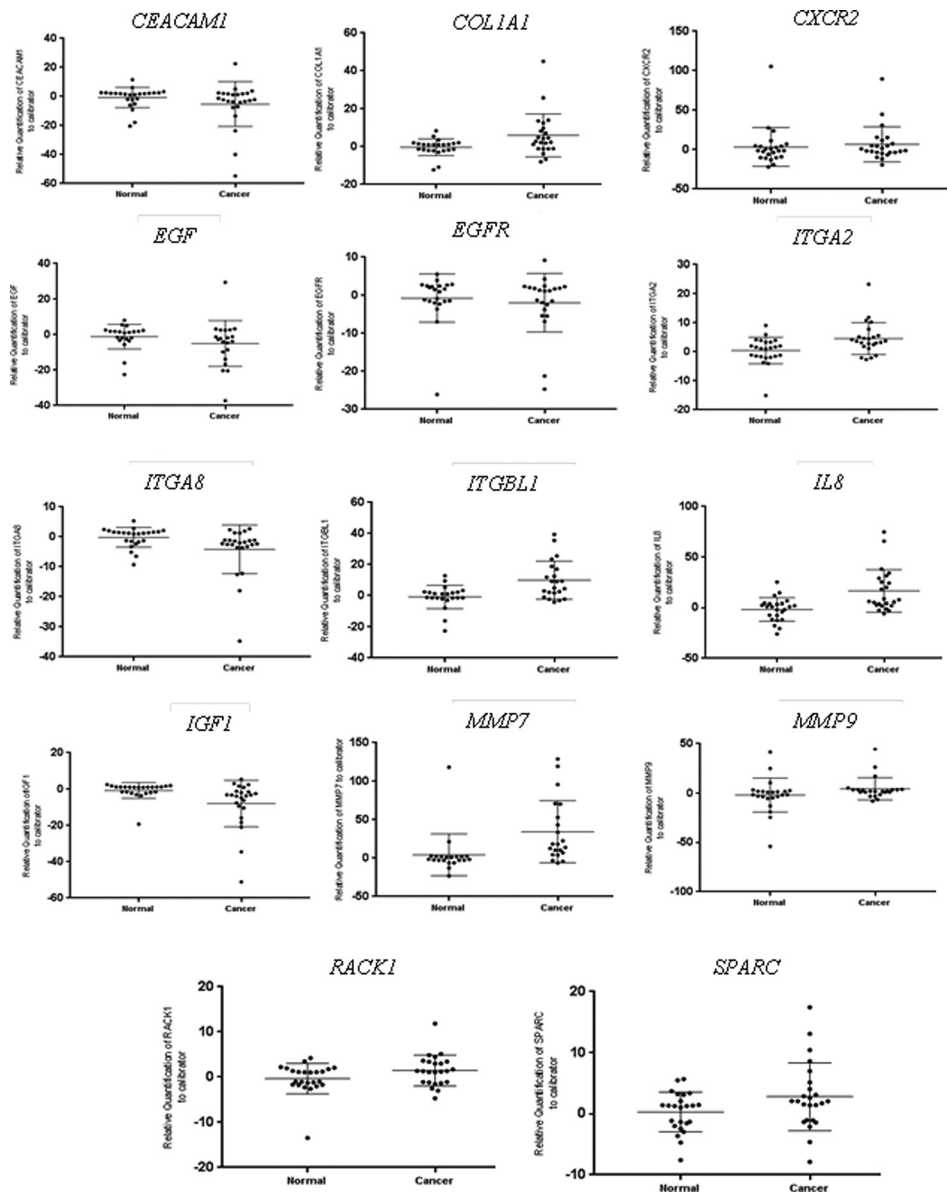


FIG. 2. Gene expression fold change levels for the patients included in the study showing statistically significant differences between normal mucosa and matched tumour samples from our patient cohort. We took the average fold change relative to normal tissue sample and normalised to the geometric mean of the reference genes.

### ASSOCIATION BETWEEN MALIGNANT SAMPLE GENE EXPRESSION AND HISTOPATHOLOGICAL PARAMETERS

**Tumour diameter:** The gene expression of three ECM genes in the panel, *IGF1* ( $r = -0.596$ ,  $p = 0.009$ ), *ITGBL1* ( $r = 0.669$ ,  $p = 0.005$ ), and *VEGF* ( $r = -0.487$ ,  $p = 0.04$ ), correlated with the diameter of the samples which were recorded during surgical examination. The tumour size at the largest point was recorded in millimetres (mm) with three categories of  $<30$  mm,  $30\text{--}40$  mm, and  $>40$  mm. *IGF1* decreased expression ( $-13$  fold) and *ITGBL1* increased expression (26.4-fold) were observed to be present in tumours with a diameter greater than 40 mm.

**Mucinous Component:** From the cohort, four of the twenty tumours were documented as having a mucinous component and all were Stage III carcinomas. Colorectal mucinous carcinomas present at a more advanced stage, predominantly in men, with higher right colon location

TABLE III. Genes with dysregulated expression from Figure 1 are shown. Statistically significant correlated (co-expressed) genes are highlighted in bold.

		Correlations													
		ITGA2	ITGA8	ITGBL1	IGF1	SPARC	EGF	EFGR	MMP7	MMP9	COL1A1	CEACAM1	RACK1	IL8	CXCR2
ITGA2	Pearson Correlation	1													
	Sig. (2-tailed)														
ITGA8	Pearson Correlation	0.082	1												
	Sig. (2-tailed)	0.711													
ITGBL1	Pearson Correlation	0.157	-0.339	1											
	Sig. (2-tailed)	0.495	0.123												
IGF1	Pearson Correlation	0.012	0.371	-0.285	1										
	Sig. (2-tailed)	0.955	0.074	0.199											
SPARC	Pearson Correlation	0.211	-0.264	<b>.493<sup>a</sup></b>	-0.235	1									
	Sig. (2-tailed)	0.333	0.213	0.02	0.269										
EGF	Pearson Correlation	0.086	0.077	<b>.487<sup>a</sup></b>	-0.072	0.047	1								
	Sig. (2-tailed)	0.718	0.739	0.029	0.756	0.84									
EFGR	Pearson Correlation	0.227	0.161	0.172	0.32	-0.277	<b>.470<sup>a</sup></b>	1							
	Sig. (2-tailed)	0.298	0.453	0.445	0.127	0.191	0.032								
MMP7	Pearson Correlation	<b>.522<sup>a</sup></b>	0.081	-0.003	-0.102	0.197	-0.051	-0.049	1						
	Sig. (2-tailed)	0.018	0.726	0.988	0.659	0.391	0.836	0.834							
MMP9	Pearson Correlation	0.146	-0.053	-0.053	0.041	0.298	0.017	-0.169	0.051	1					
	Sig. (2-tailed)	0.527	0.814	0.82	0.857	0.178	0.946	0.453	0.831						
COL1A1	Pearson Correlation	0.079	-0.323	0.174	-0.292	<b>.763<sup>b</sup></b>	-0.13	-0.283	0.131	0.277	1				
	Sig. (2-tailed)	0.728	0.133	0.438	0.176	0.01	0.585	0.191	0.571	0.212					
CEACAM1	Pearson Correlation	-0.161	0.225	-0.097	0.215	-0.203	-0.194	0.037	-0.224	-0.295	-0.143	1			
	Sig. (2-tailed)	0.462	0.291	0.667	0.312	0.341	0.401	0.865	0.329	0.183	0.515				
RACK1	Pearson Correlation	0.274	-0.125	0.09	-0.108	0.357	-0.002	-0.077	0.066	0.354	<b>.451<sup>a</sup></b>	0.258	1		
	Sig. (2-tailed)	0.206	0.561	0.692	0.614	0.087	0.993	0.721	0.777	0.106	0.031	0.223			
IL8	Pearson Correlation	0.141	-0.313	0.21	-0.311	-0.099	-0.35	-0.073	0.352	-0.058	-0.073	0.145	0.094	1	
	Sig. (2-tailed)	0.521	0.136	0.349	0.139	0.644	0.12	0.734	0.117	0.798	0.742	0.5	0.663		
CXCR2	Pearson Correlation	0.392	-0.1	0.403	0.124	0.141	0.054	0.369	0.105	0.017	0.289	-0.11	0.25	0.118	1
	Sig. (2-tailed)	0.064	0.642	0.063	0.564	0.512	0.815	0.076	0.649	0.941	0.181	0.61	0.238	0.583	

<sup>a</sup>Statistically significant ( $p < 0.05$ ), Correlation is significant at the 0.05 level (2-tailed).<sup>b</sup>Correlation is significant at the 0.01 level (2-tailed).



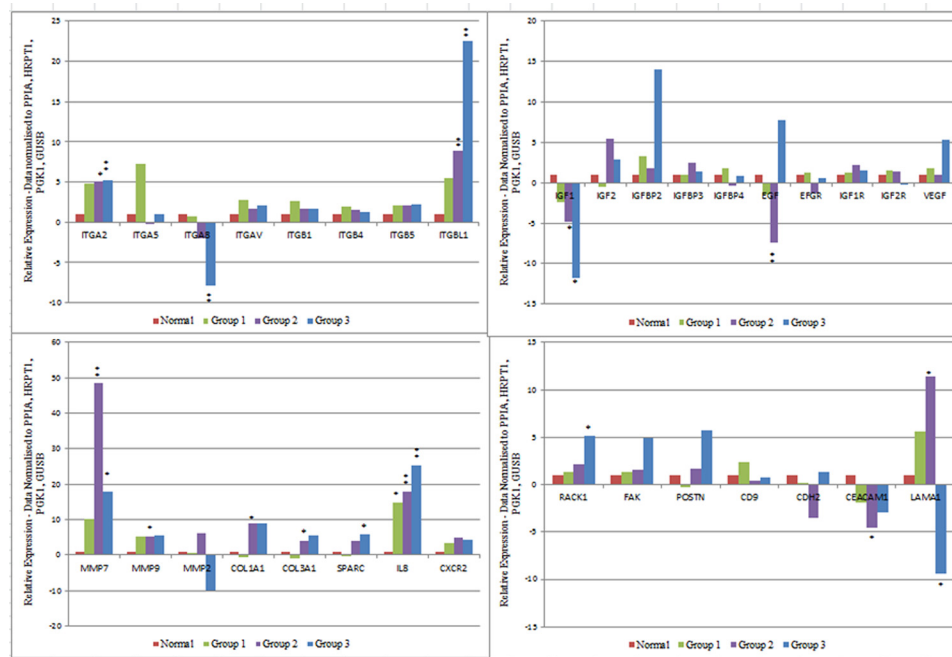


FIG. 3. The figure shows expression fold-difference for the gene panel between normal and cancer samples when samples are categorised according to UICC staging. Normal samples are normalised to one and highlighted in red, Group 1 is highlighted in green, Group 2 is highlighted purple, and Group 3 is highlighted in blue. Statistical significance is also shown (p-value; \* =  $<0.05$ ; \*\* =  $<0.01$ ).

TABLE IV. Clinicopathological data of patients used for gene expression analysis qPCR in our study.

Number of patients (n = 24)	
TVA/other pathology	6(25%)
Carcinoma	18(75%)
<b>Age(years)</b>	
Range	45–83
Mean + SD	65.87(9.85)
Median	65.5
Mode	56
<b>Gender(n = 24)</b>	
Male	15
Female	9
<b>UICC staging(n = 24)</b>	
N/A	5 (21%)
Stage I	1(4%)
Stage II	11(46%)
Stage III	6(25%)
Stage IV	1(4%)
<b>Location(n = 24)</b>	
Right colon (male/female)	9 (6/3)
Left colon (male/female)	5(4/1)
Rectum/Sigmoid(male/female)	9(4/5)
Total Colectomy	1(1/0)
<b>Tumour diameter(n = 24)</b>	
N/A	6

TABLE IV. (Continued.)

Number of patients (n = 24)	
<30 mm	6
30–40 mm	3
>40 mm	7
Unknown	2
<b>Tumour grade(n = 24)</b>	
N/A	5
T1	1
T2	4
T3	11
T4	2
Unknown	1
<b>Nodal status(n = 24)</b>	
NX	4
N0	12
N1	6
N2	2
<b>Differentiation(n = 24)</b>	
Well-moderate	9
Moderate	10
Moderate-poor	5
<b>Invasive(n = 24)</b>	
Yes	16
No	8
<b>Mucin secretion(n = 24)</b>	
Mucinous	5
Not stated	19
<b>Lymphovascular invasion(n = 24)</b>	
N/A	6
Present	3
Absent	9
Suspicious	2
Not stated	4

rate, and a worse overall 5-year survival rate than the non-mucinous colorectal cancer.<sup>33</sup> Three of the four patients with mucinous component were male. *ITGBL1* and *ITGA8* expressions correlated with whether or not a mucinous component was present. Increased expression in *ITGBL1* ( $r=0.669$ ,  $p=0.005$ ) was observed in mucin positive tumours and average expression of *ITGA8* was decreased ( $r=0.655$ ,  $p=0.002$ ). Three of the four patients received chemotherapy.

**Differentiation:** The expression levels of three genes showed a moderate correlation with the level of differentiation which was recorded during surgical resection. Upregulation of *COL3A1* was significantly associated with poorer differentiation ( $r=0.469$ ,  $p=0.024$ ) along with upregulation of *MMP9* ( $r=0.423$ ,  $p=0.05$ ). It has been shown<sup>34</sup> that increased levels of *MMP-9* correlate with poorer differentiation. A slight increase in *IGF-1R* ( $r=0.533$ ,  $p=0.049$ ) levels also correlated with poorer differentiation.

**Recurrence:** Three of the samples in the cohort were taken from patients who later presented with a recurrence of the carcinoma. *PTK2* ( $r=0.435$ ,  $p=0.043$ ) levels in these patients were shown to be elevated in comparison with patients whose cancers had not recurred.

**Age:** Patients were categorised into two groups of being up to and including sixty-four years of age and sixty-five and older. Increased *MMP7* expression correlated with increased age

in the cohort ( $r=0.488$ ,  $p=0.018$ ). Patients aged sixty-four and under had an average increase of 10-fold expression of *MMP7* while patients aged sixty-five and over had an average of 40-fold increase. *CDH2* levels were also correlated ( $r=0.490$ ,  $p=0.015$ ) with age where patients under 64 had six fold decreased *CDH2* expression than people aged over 65. Increased *IGFBP3* expression (+2 fold) correlated ( $r=0.473$ ,  $p=0.023$ ) with an increased age compared to the younger group.

**Gender:** *IGFBP3* expression levels were noted to correlate ( $r=0.424$ ,  $p=0.044$ ) with patient gender with male patients having an increased expression (+4-fold) compared to females

**Lymph Node Status:** In nonmetastatic colorectal cancer, lymph node status is the strongest pathologic predictor of patient outcome and is used for determining the most appropriate adjuvant treatment for patients. In the cohort, patient's lymph nodes were examined for abnormal cellular behaviour and inflammation. The nodal status was determined using the *American Joint Committee on Cancer (AJCC)* TNM system where the "N" portion of this classification system refers to the nodal status indicating the presence, if any, of cancer spread to other lymph nodes. The system is: NX: Regional lymph nodes cannot be assessed, N0: Cancer has not spread to regional lymph nodes, N1: Cancer has spread to 1 to 3 lymph nodes, and N2: Cancer has spread to 4 to 9 lymph nodes. Four patients were classified as NX, twelve patients were classified as N0, six were classified as N1, and 2 were classified as N2. The expression of *PTK2*, or Focal Adhesion Kinase (*FAK*) as it is more commonly known, showed a step-wise increase from N0 to N1 to N2 ( $r=0.587$ ,  $p=0.006$ ). A stepwise decrease was also observed for *ITGA8* for N0 to N2 ( $r=-0.523$ ,  $p=0.018$ ).

**Invasiveness:** The invasiveness of the tumour is closely linked to the size and the grade of the tumour and is determined by whether or not the tumour penetrates the basement membrane and/or muscle layers in the lumen of the colon. For this study, a tumour was categorised as being invasive or not. In the twenty-four patient cohort, nineteen samples are documented to be invasive malignancies of varying degree. The depth of invasion was not recorded. These samples were analysed to determine if changes in expression for the ECM genes correlated with the invasiveness of a cancer. *SPARC* and *COL1A1* expression levels were shown to be increased significantly in samples which were invasive ( $r=0.465$ ,  $p=0.022$ ) ( $r=0.494$ ,  $p=0.017$ ).

**Perineural Invasion (PNI):** PNI is grossly underreported in CRC and could serve as an independent prognostic factor of outcomes in these patients.<sup>35</sup> The PNI absence was reported in six of the 20 tumour samples. Perineural invasion was not stated for fourteen others. *MMP-9* levels were significantly decreased in patients with no perineural invasion ( $r=0.481$ ,  $p=0.037$ ). Increases in *MMP-9* aid a cancer cell's progression to metastasis and have been shown<sup>36</sup> to correlate with increases in perineural invasion in the literature.

**Stage:** The clinical staging of the patients disease was also determined by the AJCC TNM system. Stage grouping can be seen in Table IV. *ITGBL1* expression ( $r=0.588$ ,  $p=0.013$ ) was identified as showing a stepwise increase in expression which correlated with the staging of the patient cohort with an ~8 fold increase in expression for patients with Stage II and ~22 fold increase for Stage III in comparison to Stage I. *LAMA1* expression showed an increase in Stage II (+20-fold) but a significant decrease in expression in Stage III (-40 fold) ( $r=0.509$ ,  $p=0.037$ ).

**Grade:** *ITGA2*, *ITGB1*, and *CD9* gene expression levels were correlated with the grade of the malignant samples analysed albeit with a small sample size. *ITGA2* was observed to be lower in T3 and T4 tumour samples in comparison to T1 and T2 tumour samples ( $r=-0.657$ ,  $p=0.004$ ), while the levels of *ITGB1* replicated this pattern ( $r=-0.501$ ,  $p=0.041$ ). *CD9* expression levels also correlated with the grade of the cancer samples with a stepwise decrease in expression ( $r=-0.613$ ,  $p=0.007$ ) from T1 (+10 fold), T2 (+2.14 fold), T3 (+0.05 fold) to T4 (-11.5 fold). In addition, the T1 sample and the two T4 samples were removed from the dataset and T2 and T3 correlations were evaluated independently. When this was done, the correlation between *ITGA2* and *ITGB1* expression with the grade of the tumour remained statistically significant. However, the correlation between *CD9* expression and grade did not remain statistically significant.

## DISCUSSION

Identifying differentially expressed genes acts as an additional clinical tool that may be used to diagnose patients and may influence the potential treatment decisions. The global dysregulation observed in this gene expression study of the ECM indicates that substantial dysregulation and remodelling of the ECM occur during colorectal pathogenesis and this can lead to increased metastatic potential. The differential expression of fourteen genes which were statistically significant was identified. The largest dysregulation was observed in *MMP7*. *MMP7* cleaves cell surface proteins, promotes adhesion of cancer cells, and increases the potential of tumour metastasis<sup>37</sup> and has been widely shown to be significantly higher than that of normal cells in numerous studies both for colorectal carcinoma<sup>38–40</sup> and other cancers.<sup>41–43</sup> *IL-8* was also dysregulated and its overexpression has been detected in many carcinomas, including colorectal cancer<sup>44</sup> and is associated with poor prognosis.<sup>45</sup> It has been suggested that tumour cells produce IL-8 as an autocrine growth factor that promote tumour growth, tissue invasion, and metastatic spread.<sup>46</sup> *ITGBL1* is a gene that is not frequently highlighted in the literature but has been shown to have a role in breast cancer bone metastasis by activating TGF $\beta$  signalling pathway.<sup>47</sup> *MMP9* is a second protease from the initial panel that emerged as having differential expression in malignant tissue in comparison to the matched normal tissue. Its overexpression has been observed in colon cancer studies where it is largely associated with metastatic progression of colorectal cancer and overall poor prognosis.<sup>48,49</sup> *COL1A1* was observed to be increased which may contribute to abnormal tissue remodelling<sup>50</sup> and increased tumour invasion potential. *ITGA2* has been linked to having a role in cell migration and metastasis by increasing cofilin phosphorylation<sup>51,52</sup> and was observed to be upregulated in the cohort, while *CXCR2*, which is the receptor for *LAMA1* and also a critical mediator of cellular senescence, was observed to be overexpressed. Upregulated expression of *SPARC* was shown which conforms with literature.<sup>53</sup> *RACK1* overexpression was also evident between samples and our research group has shown it interacts with *PTK2* and *IGF1R* to promote cellular migration.<sup>54</sup> *IGF1* gene expression has been reported to be higher in colorectal cancer than that of normal mucosa.<sup>55,56</sup> However, decreased levels were observed in this study in the malignant tissue. *CEACAM1*, a gene shown to be upregulated in gastric cancer and has been shown to promote angiogenesis,<sup>57</sup> was also observed to be overall decreased between normal and cancerous tissue. Restoration of *CEACAM1* expression in tumour cell lines often abolishes their oncogenicity *in vivo*, and therefore, this adhesion molecule has been regarded as a tumour suppressor.<sup>58</sup> Decreased levels of *ITGA8* levels were observed in the cohort. *ITGA8* has been shown to be one of a panel of genes on an enriched molecular pathway in colorectal cancer.<sup>59</sup> *EGF* and its receptor *EGFR* were both downregulated in the study and both have been shown to be implicated in the growth and invasion of colon cancer.<sup>60,61</sup> The overexpression of *EGFR* in colon cancers without a *KRAS* mutation has led to targeted therapies by blocking *EGFR* which has led to increased patient benefits.<sup>62</sup> When analysed with regard to UICC classification of the disease, the pattern of dysregulation became more pronounced and allowed for potential molecular markers to be identified. By showing the differential expression with regard to UICC staging (Figure 3), a number of genes emerge as being potential biomarkers for Stage I, Stage II, and Stage III/IV colorectal cancer. These genes could become targets for specific treatment in patients presenting with these stages of colorectal cancer or become an influencing factor in determining a patients' diagnosis. The stages at which genes become more influential in disease development and progression is shown (Figure 4). Once expression levels were determined by experimental means, the panel of fourteen genes was assessed for co-expressed genes (genes with similar expression patterns) using Pearson's Correlation test. Strength of relationship is determined by closeness to 1 (direct correlation) or  $-1$  (inverse correlation). *COL1A1* was identified to be co-expressed with both *SPARC* and *RACK1*. *COL1A1* and *SPARC* genes participate in ECM and cytoskeletal control and are also involved in the carcinogenesis of many malignancies and both have been shown to be relevant for tumour prognosis.<sup>63</sup> Co-expression was also significant between *MMP7* and *ITGA2*, *ITGBL1* and *SPARC*, *ITGBL1* and *EGF*, and *EGF* and *EGFR* (Table III). A similar pattern of expression can be used to identify co-expression networks which are of biological interest as co-expressed genes can be (a)

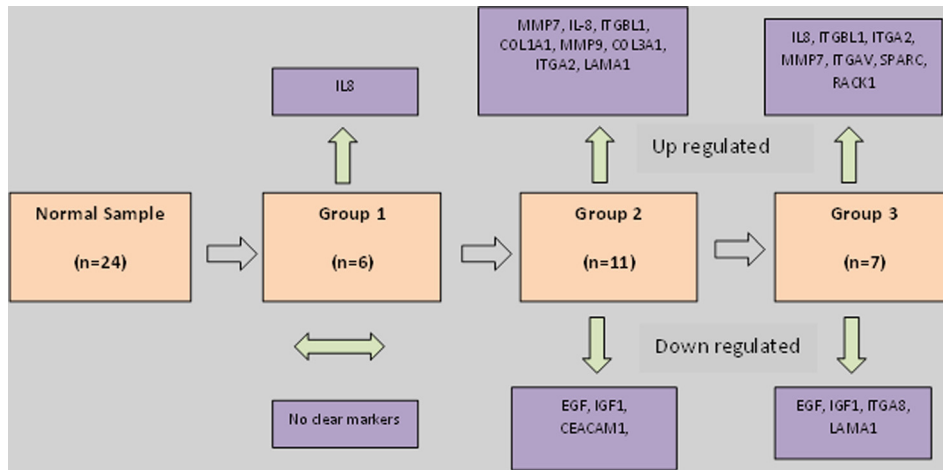


FIG. 4. Potential genetic biomarkers for CRC diagnosis based on UICC staging. Patients were grouped into categories depending on the staging of the carcinoma.

controlled by the same transcriptional regulatory program (b) functionally related or (c) members of the same pathway or protein complex.<sup>64</sup> Furthermore, expression coordination between genes involved in different functional categories suggests that the functional pathways can modulate each other. Associations between the gene expression patterns of all genes and histopathological parameters recorded was also assessed. Strong associations between these parameters were observed highlighting ECM genes which potentially could be used in aiding tumour phenotype predictions. Graphical representation of the association of expression analysis with histopathological parameters can be seen in Figure 5 where the genes identified could be used alone or as complement to existing CRC diagnostic tools for cancer development and/or progression.

**CONCLUSION**

This study highlights fourteen ECM genes, which show differential expression between tumour tissue and matched normal colon mucosa. The genes which emerged as being different have been previously shown in the literature to play key roles in the carcinoma process for a range of individual cancers and in particular, colorectal carcinoma. A pairwise correlation

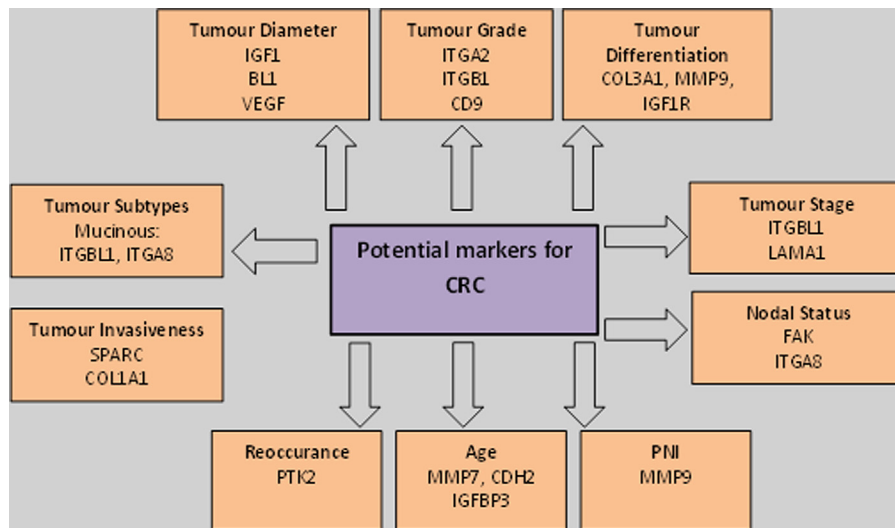


FIG. 5. Association between gene expression and histopathological parameters. Pathological parameters can be found in Table IV.

TABLE V. The dysregulated gene panel (Figure 2) was analysed as a combined cohort. Values shown in the table show expression fold change between normal and cancerous samples and associated p-values (statistical significance).

Gene name	Patient cohort (n = 24)	
	Fold change	Stat sig.
CEACAM1	-4.1	0.048 <sup>a</sup>
COL1A1	+6.2	0.039 <sup>a</sup>
CXCR2	+4.0	0.02 <sup>a</sup>
EGF	-5.0	0.014 <sup>a</sup>
EGFR	-3.0	0.05 <sup>a</sup>
ITGA2	+4.0	0.007 <sup>a</sup>
ITGA8	-3.6	0.003 <sup>a</sup>
ITGBL1	+11.8	0.003 <sup>a</sup>
IGF1	-7.7	0.007 <sup>a</sup>
IL-8	+18.45	<0.001 <sup>a</sup>
MMP7	+31.1	<0.001 <sup>a</sup>
MMP9	+7.4	0.02 <sup>a</sup>
RACK1	+2.4	0.019 <sup>a</sup>
SPARC	+3.8	0.033 <sup>a</sup>

<sup>a</sup>Statistically significant ( $p < 0.05$ ).

pattern emerged also which showed that certain genes were co-expressed. Further to this, correlations emerged between the gene expression pattern and clinic-pathological parameters which were obtainable from pathology reports such as patient gender, tumour diameter, lymphovascular invasion, and nodal status. The genes highlighted in this study may represent important targets in colorectal carcinogenesis and could provide useful clinicopathological tools in the management and treatment of colorectal cancer. Expanding the patient cohort and genetic panel could yield further potential biomarkers for metastasis and other treatment targets in colorectal cancer. The successful utilisation of novel technologies such as microfluidics to solve biological questions potentially allows us to maximise biomarker discovery rates and improve cancer patient prognosis. qPCR integrated microfluidics provides numerous advantages for gene expression: economies of scale, parallelisation and automation, and increased sensitivity and precision that comes from small volume reactions. The microfluidic droplet technology used in this study has the ability to analyse the expression levels of hundreds of genes across hundreds of samples in a continuous flow regime while using reduced volumes of both sample and reagents. The instrumentation will assist in the development of further microliter and nanoliter microfluidic platforms which will progress molecular diagnostic advances in the biological research field. Using this technology, we have shown that differential gene expression levels greatly influence many biological conditions in both normal and abnormal circumstances and can provide valuable information into the tissue and cell developmental behaviour in response to environmental stimuli. The three-dimensional matrix, which supports, protects, and stabilises all structures and cells, has diverse functions in the body. Accumulating evidence supports the theory that the ECM is mechanically, biochemically, and functionally influential in the metastatic behaviour of cancer cells. All structures within the matrix environment are inherently linked. By identifying ECM related genes that influence the metastatic potential of cells, we could potentially use these genes to predict prognosis. Comparing these findings with current array-based studies supports the emerging panel of ECM genes as being dysregulated in colorectal cancer. This highlights the potential of using microfluidic-based approaches for candidate biomarker studies and is an example of one of the first applications of microfluidic technology to colorectal cancer studies. The results of this gene expression profiling could stimulate an interest in the direct use of expression profiles in the routine clinical setting. However, the data outlined is preliminary. Further studies involving larger patient populations (including different

tissue types) and confirmed biomolecular networks are merited giving the potential for more informed diagnosis and early identification of metastasis in malignancies.

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Carmel Hannan for conversations regarding the statistical analysis of the data and thank colleagues in Stokes Laboratories and the Laboratory of Cellular and Molecular Biology for helpful discussions. We acknowledge funding from the Irish Cancer Society Grant CRS12DOW (to CMD) and from Science Foundation Ireland award 13/CDA/2228 (to PAK). This work was also funded in part by donations made by patients to the Symptomatic Breast Donations Fund in the University Hospital Limerick. We also acknowledge financial support from the University of Limerick which made this research possible.

- <sup>1</sup>J. Ferlay *et al.*, “Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012,” *Int. J. Cancer* **136**(5), E359–E386 (2015).
- <sup>2</sup>R. Higuchi *et al.*, “Simultaneous amplification and detection of specific DNA sequences,” *Bio/Technol.* **10**(4), 413–417 (1992).
- <sup>3</sup>X. Solé *et al.*, “Discovery and validation of new potential biomarkers for early detection of colon cancer,” *PLoS One* **9**(9), e106748 (2014).
- <sup>4</sup>K. M. Clark-Langone *et al.*, “Biomarker discovery for colon cancer using a 761 gene RT-PCR assay,” *BMC Genomics* **8**(1), 279 (2007).
- <sup>5</sup>B. J. Esteban, M. Cronin, M. L. Liu, M. G. Llamas, M. G. Walker, R. Mena, and S. Shak, “Tumor gene expression and prognosis in breast cancer: multi-gene RT-PCR assay of paraffin-embedded tissue [abstract],” *Proc. Am. Soc. Clin. Oncol.* **22**, 850 (2003).
- <sup>6</sup>S. Ocak *et al.*, “High-throughput molecular analysis in lung cancer: Insights into biology and potential clinical applications,” *Euro. Respir. J.* **34**(2), 489–506 (2009).
- <sup>7</sup>J. Liu, C. Hansen, and S. R. Quake, “Solving the “World-to-Chip” interface problem with a microfluidic matrix,” *Anal. Chem.* **75**(18), 4718–4723 (2003).
- <sup>8</sup>J. Hogan *et al.*, “Adipophilin (ADFP) is a novel and independent prognostic biomarker in colorectal cancer,” in *Ir. J. Med. Sci.* **183**, S28 (2014).
- <sup>9</sup>C. M. Dowling *et al.*, “Protein kinase C beta II suppresses colorectal cancer by regulating IGF-1 mediated cell survival,” *Oncotarget* **7**, 20919 (2016).
- <sup>10</sup>J. Hogan *et al.*, “Overall survival is improved in mucinous adenocarcinoma of the colon,” *Int. J. Colorectal Dis.* **29**(5), 563–569 (2014).
- <sup>11</sup>J. Hogan *et al.*, “Introducing a novel and robust technique for determining lymph node status in colorectal cancer,” *Ann. Surg.* **260**(1), 94–102 (2014).
- <sup>12</sup>R. P. Mecham, “Overview of extracellular matrix,” in *Current Protocols in Cell Biology* (Wiley-Blackwell, 2001).
- <sup>13</sup>E. D. Hay, “Extracellular matrix alters epithelial differentiation,” *Curr. Opin. Cell Biol.* **5**(6), 1029–1035 (1993).
- <sup>14</sup>C. Bonnans, J. Chou, and Z. Werb, “Remodelling the extracellular matrix in development and disease,” *Nat. Rev. Mol. Cell Biol.* **15**(12), 786–801 (2014).
- <sup>15</sup>P. Lu *et al.*, “Extracellular matrix degradation and remodeling in development and disease,” *Cold Spring Harbor Perspect. Biol.* **3**(12), a005058 (2011).
- <sup>16</sup>D. Stickens, “Altered endochondral bone development in matrix metalloproteinase 13-deficient mice,” *Development* **131**(23), 5883–5895 (2004).
- <sup>17</sup>J. A. Joyce and J. W. Pollard, “Microenvironmental regulation of metastasis,” *Nat. Rev. Cancer* **9**(4), 239–252 (2009).
- <sup>18</sup>P. Lu, V. M. Weaver, and Z. Werb, “The extracellular matrix: A dynamic niche in cancer progression,” *J. Cell Biol.* **196**(4), 395–406 (2012).
- <sup>19</sup>M. W. Pfaffl, “A new mathematical model for relative quantification in real-time RT-PCR,” *Nucleic Acids Res.* **29**(9), 45e (2001).
- <sup>20</sup>E. A. H. Kheirleiseid *et al.*, “Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer,” *BMC Mol. Biol.* **11**(1), 12 (2010).
- <sup>21</sup>L. A. A. Sorby *et al.*, “Evaluation of suitable reference genes for normalization of real-time reverse transcription PCR analysis in colon cancer,” *J. Exp. Clin. Cancer Res.* **29**(1), 144 (2010).
- <sup>22</sup>J. Vandesompele *et al.*, “Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes,” *Genome Biol.* **3**(7), 0034.1 (2002).
- <sup>23</sup>C. M. Dowling *et al.*, “The importance of selecting the appropriate reference genes for quantitative real time PCR as illustrated using colon cancer cells and tissue,” *F1000 Research* **5**, 99 (2016).
- <sup>24</sup>C. J. Hayes and T. M. Dalton, “Microfluidic droplet-based PCR instrumentation for high-throughput gene expression profiling and biomarker discovery,” *Biomol. Detect. Quantif.* **4**, 22–32 (2015).
- <sup>25</sup>K. Curran, S. Colin, L. Baldes, and M. D. Davies, “Liquid bridge instability applied to microfluidics,” *Microfluid. Nanofluid.* **1**, 336–345 (2005).
- <sup>26</sup>C. T. Wittwer, G. B. Reed, and K. M. Ririe, “Rapid cycle DNA amplification,” in *The Polymerase Chain Reaction* (Springer Science + Business Media, 1994), pp. 174–181.
- <sup>27</sup>S. A. Bustin *et al.*, “The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments,” *Clin. Chem.* **55**(4), 611–622 (2009).
- <sup>28</sup>J. R. Dijkstra *et al.*, “MicroRNA expression in formalin-fixed paraffin embedded tissue using real time quantitative PCR: the strengths and pitfalls,” *J. Cell. Mol. Med.* **16**(4), 683–690 (2012).

- <sup>29</sup>K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method," *Methods* **25**(4), 402–408 (2001).
- <sup>30</sup>Thermo Fisher Scientific, see <https://www.thermofisher.com/order/catalog/product/4414133> for TaqMan<sup>®</sup> Extracellular Matrix & Adhesion Molecule Gene Expression Array.
- <sup>31</sup>SABiosciences, see <https://www.qiagen.com/us/shop/pcr/primer-sets/rt2-profiler-pcr-arrays?catno=PAHS-013Z#geneglobe> for RT2-Profiler PCR array extracellular matrix and adhesion molecules, Qiagen, Germantown, MD.
- <sup>32</sup>A. Latil *et al.*, "Evaluation of androgen, estrogen (ER alpha and ER beta), and progesterone receptor expression in human prostate cancer by real-time quantitative reverse transcription-polymerase chain reaction assays," *Cancer Res.* **61**(5), 1919–1926 (2001).
- <sup>33</sup>V. N. Papadopoulos *et al.*, "Prognostic significance of mucinous component in colorectal carcinoma," *Tech. Coloproctol.* **8**(1), s123 (2004).
- <sup>34</sup>Y. Zhang *et al.*, "Expression of MMP-9 and WAVE3 in colorectal cancer and its relationship to clinicopathological features," *J. Cancer Res. Clin. Oncol.* **138**(12), 2035–2044 (2012).
- <sup>35</sup>C. Liebig *et al.*, "Perineural invasion is an independent predictor of outcome in colorectal cancer," *J. Clin. Oncol.* **27**(31), 5131–5137 (2009).
- <sup>36</sup>L. Duan *et al.*, "GPC-1 may serve as a predictor of perineural invasion and a prognosticator of survival in pancreatic cancer," *Asian J. Surg.* **36**(1), 7–12 (2013).
- <sup>37</sup>K. Edman *et al.*, "The discovery of MMP7 inhibitors exploiting a novel selectivity trigger," *ChemMedChem* **6**(5), 769–773 (2011).
- <sup>38</sup>Y. Adachi *et al.*, "Contribution of matrilysin (MMP-7) to the metastatic pathway of human colorectal cancers," *Gut* **45**(2), 252–258 (1999).
- <sup>39</sup>Y. Miyata *et al.*, "Expression of matrix metalloproteinase-7 on cancer cells and tissue endothelial cells in renal cell carcinoma: Prognostic implications and clinical significance for invasion and metastasis," *Clin. Cancer Res.* **12**(23), 6998–7003 (2006).
- <sup>40</sup>A. Said, J.-P. Raufman, and G. Xie, "The role of matrix metalloproteinases in colorectal cancer," *Cancers* **6**(1), 366–375 (2014).
- <sup>41</sup>F.-Q. Wang *et al.*, "Matrilysin (MMP-7) promotes invasion of ovarian cancer cells by activation of progelatinase," *Int. J. Cancer* **114**(1), 19–31 (2005).
- <sup>42</sup>T. Kitoh *et al.*, "Increased expression of matrix metalloproteinase-7 in invasive early gastric cancer," *J. Gastroenterol.* **39**(5), 434–440 (2004).
- <sup>43</sup>K. Ohashi *et al.*, "Increased expression of matrix metalloproteinase 7 and 9 and membrane type 1-matrix metalloproteinase in esophageal squamous cell carcinomas," *Cancer* **88**(10), 2201–2209 (2000).
- <sup>44</sup>E. A. H. Kheirleheid *et al.*, "Clinical applications of molecular profiling of colorectal cancer," *Euro. J. Cancer Suppl.* **8**(5), 197 (2010).
- <sup>45</sup>Y. Ning *et al.*, "Interleukin-8 is associated with proliferation, migration, angiogenesis and chemosensitivity *in vitro* and *in vivo* in colon cancer cell line models," *Int. J. Cancer* **128**(9), 2038–2049 (2011).
- <sup>46</sup>K. Xie, "Interleukin-8 and human cancer biology," *Cytokine Growth Factor Rev.* **12**(4), 375–391 (2001).
- <sup>47</sup>X. Q. Li *et al.*, "ITGBL1 Is a Runx2 transcriptional target and promotes breast cancer bone metastasis by activating the TGF signaling pathway," *Cancer Res.* **75**(16), 3302–3313 (2015).
- <sup>48</sup>B. Yang *et al.*, "Matrix metalloproteinase-9 overexpression is closely related to poor prognosis in patients with colon cancer," *World J. Surg. Oncol.* **12**(1), 1–6 (2014).
- <sup>49</sup>C. Y. Li *et al.*, "Matrix metalloproteinase 9 expression and prognosis in colorectal cancer: A meta-analysis," *Tumour Biol.* **34**(2), 735–741 (2013).
- <sup>50</sup>Y. A. Wu *et al.*, "Transcriptome profiling of the cancer, adjacent non-tumor and distant normal tissues from a colorectal cancer patient by deep sequencing," *PLoS One* **7**(8), e41001 (2012).
- <sup>51</sup>A. Ferraro, T. Boni, and A. Pintzas, "EZH2 regulates cofilin activity and colon cancer cell migration by targeting ITGA2 gene," *PLoS One* **9**(12), e115276 (2014).
- <sup>52</sup>J. L. Luque-García *et al.*, "Differential protein expression on the cell surface of colorectal cancer cells associated to tumor metastasis," *Proteomics* **10**(5), 940–952 (2010).
- <sup>53</sup>T. D. Said, "Secreted protein acidic and rich in cysteine (sparc) in cancer," *J. Carcinogene Mutagene* **4**, 151 (2013).
- <sup>54</sup>P. A. Kiely *et al.*, "RACK1-mediated integration of adhesion and Insulin-like Growth Factor I (IGF-I) signaling and cell migration are defective in cells expressing an IGF-I receptor mutated at Tyrosines 1250 and 1251," *J. Biol. Chem.* **280**(9), 7624–7633 (2005).
- <sup>55</sup>S. Freier *et al.*, "Expression of the insulin-like growth factors and their receptors in adenocarcinoma of the colon," *Gut* **44**(5), 704–708 (1999).
- <sup>56</sup>G.-L. X. Rong Zhang, Y. Li, L.-J. He, L.-M. Chen, G.-B. Wang, S.-Y. Lin, G.-Y. Luo, X.-Y. Gao, and H.-B. Shan, "The role of insulin-like growth factor 1 and its receptor in the formation and development of colorectal carcinoma," *J. Int. Med. Res.* **41**, 1228–1235 (2013).
- <sup>57</sup>J. F. Shi *et al.*, "Expression of carcinoembryonic antigen-related cell adhesion molecule 1(CEACAM1) and its correlation with angiogenesis in gastric cancer," *Pathol., Res. Pract.* **210**(8), 473–476 (2014).
- <sup>58</sup>V. Fiori, M. Magnani, and M. Cianfriglia, "The expression and modulation of CEACAM1 and tumor cell transformation," *Ann. Ist. Super Sanita* **48**(2), 161–171 (2012).
- <sup>59</sup>T. Kok-Sin *et al.*, "Identification of diagnostic markers in colorectal cancer via integrative epigenomics and genomics data," *Oncol. Rep.* **34**, 22 (2015).
- <sup>60</sup>R. Shaheen *et al.*, "Effects of an antibody to vascular endothelial growth factor receptor-2 on survival, tumor vascularity, and apoptosis in a murine model of colon carcinomatosis," *Int. J. Oncol.* **18**, 221 (2001).
- <sup>61</sup>J. Mendelsohn, "The epidermal growth factor receptor as a target for cancer therapy," *Endocr. Relat. Cancer* **8**(1), 3–9 (2001).
- <sup>62</sup>C. S. Karapetis *et al.*, "K-ras mutations and benefit from cetuximab in advanced colorectal cancer," *New England J. Med.* **359**(17), 1757–1765 (2008).
- <sup>63</sup>J. Helleman *et al.*, "Association of an extracellular matrix gene cluster with breast cancer prognosis and endocrine therapy response," *Am. Assoc. Cancer Res.* **14**(17), 5555–5564 (2008).



- <sup>64</sup>M. T. Weirauch, "Gene coexpression networks for the analysis of DNA microarray data," in *Applied Statistics for Network Biology* (Wiley-VCH Verlag GmbH & Co. KGaA, 2011), pp. 215–250.
- <sup>65</sup>J. Ding *et al.*, "LSD1-mediated epigenetic modification contributes to proliferation and metastasis of colon cancer," *Br. J. Cancer* **109**(4), 994–1003 (2013).
- <sup>66</sup>F. Kroepil *et al.*, "Down-regulation of CDH1 is associated with expression of SNAI1 in colorectal adenomas," *PLoS One* **7**(9), e46665 (2012).
- <sup>67</sup>A. Arabzadeh and N. Beauchemin, "Stromal CEACAM1 expression regulates colorectal cancer metastasis," *Oncoimmunology* **1**(7), 1205–1207 (2012).
- <sup>68</sup>M. Mori *et al.*, "Motility related protein 1 (MRP1/CD9) expression in colon cancer," *Clin. Cancer Res.* **4**(6), 1507–1510 (1998).
- <sup>69</sup>H. Hashida *et al.*, "Clinical significance of transmembrane 4 superfamily in colon cancer," *Br. J. Cancer* **89**(1), 158–167 (2003).
- <sup>70</sup>W. Wu *et al.*, "Expression of CXCR2 and its clinical significance in human colorectal cancer," *Int. J. Clin. Exp. Med.* **8**(4), 5883–5889 (2015).
- <sup>71</sup>Y. S. Lee *et al.*, "Interleukin-8 and its receptor CXCR2 in the tumour microenvironment promote colon cancer growth, progression and metastasis," *Br. J. Cancer* **106**(11), 1833–1841 (2012).
- <sup>72</sup>M. K. Bode, T. J. Karttunen, J. Makela, L. Risteli, and J. Risteli, "Type I and III collagens in human colon cancer and diverticulosis," *Scand. J. Gastroenterol.* **35**(7), 747–752 (2000).
- <sup>73</sup>C. Kahlert *et al.*, "Overexpression of ZEB2 at the invasion front of colorectal cancer is an independent prognostic marker and regulates tumor invasion *in vitro*," *Clin. Cancer Res.* **17**(24), 7654–7663 (2011).
- <sup>74</sup>D. S. Salomon *et al.*, "Epidermal growth factor-related peptides and their receptors in human malignancies," *Crit. Rev. Oncol./Hematol.* **19**(3), 183–232 (1995).
- <sup>75</sup>Y. Feng *et al.*, "EGF signalling pathway regulates colon cancer stem cell proliferation and apoptosis," *Cell Proliferation* **45**(5), 413–419 (2012).
- <sup>76</sup>J. P. Spano *et al.*, "Impact of EGFR expression on colorectal cancer patient prognosis and survival," *Ann. Oncol.* **16**(1), 102–108 (2005).
- <sup>77</sup>F. Bertucci *et al.*, "Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters," *Oncogene* **23**(7), 1377–1391 (2004).
- <sup>78</sup>I. Shiratsuchi *et al.*, "Expression of IGF-1 and IGF-1R and their relation to clinicopathological factors in colorectal cancer," *Anticancer Res.* **31**(7), 2541–2545 (2011).
- <sup>79</sup>E. Giovannucci, "Insulin, insulin-like growth factors and colon cancer: A review of the evidence," *J. Nutrition* **131**(11), 3109S–3120S (2001).
- <sup>80</sup>R. Baserga, "The contradictions of the insulin-like growth factor 1 receptor," *Oncogene* **19**(49), 5574–5581 (2000).
- <sup>81</sup>Y. Wang and Y. Sun, "Insulin-like growth factor receptor-1 as an anti-cancer target: Blocking transformation and inducing apoptosis," *Curr. Cancer Drug Targets* **2**(3), 191–207 (2002).
- <sup>82</sup>S. R. Li *et al.*, "Differential expression patterns of the insulin-like growth factor 2 gene in human colorectal cancer," *Tumour Biol.* **25**(1–2), 62–68 (2004).
- <sup>83</sup>T. Oshima *et al.*, "Clinicopathological significance of the gene expression of matrix metalloproteinase-7, insulin-like growth factor-1, insulin-like growth factor-2 and insulin-like growth factor-1 receptor in patients with colorectal cancer: Insulin-like growth factor-1 receptor gene expression is a useful predictor of liver metastasis from colorectal cancer," *Oncol. Rep.* **20**(2), 359–364 (2008).
- <sup>84</sup>D. LeRoith and C. T. Roberts, Jr., "The insulin-like growth factor system and cancer," *Cancer Lett.* **195**(2), 127–137 (2003).
- <sup>85</sup>D. Velázquez-Fernández *et al.*, "Expression profiling of adrenocortical neoplasms suggests a molecular signature of malignancy," *Surgery* **138**(6), 1087–1094 (2005).
- <sup>86</sup>H. Yu *et al.*, "Role of the insulin-like growth factor family in cancer development and progression," *JNCI J. Natl. Cancer Inst.* **92**(18), 1472–1489 (2000).
- <sup>87</sup>G. Walker *et al.*, "Insulin-like growth factor binding proteins IGFBP3, IGFBP4, and IGFBP5 predict endocrine responsiveness in patients with ovarian cancer," *Clin. Cancer Res.* **13**(5), 1438–1444 (2007).
- <sup>88</sup>P. J. Jenkins *et al.*, "Differential expression of IGF-binding protein-3 in normal and malignant colon and its influence on apoptosis," *Endocr. Relat. Cancer* **12**(4), 891–901 (2005).
- <sup>89</sup>M. Bianchini *et al.*, "Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa," *Int. J. Oncol.* **29**, 83 (2006).
- <sup>90</sup>S. Viana Lde *et al.*, "Relationship between the expression of the extracellular matrix genes SPARC, SPP1, FN1, ITGA5 and ITGAV and clinicopathological parameters of tumor progression and colorectal cancer dissemination," *Oncology* **84**(2), 81–91 (2013).
- <sup>91</sup>J. Song *et al.*, "Beta1 integrin modulates tumor growth and apoptosis of human colorectal cancer," *Oncol. Rep.* **32**(1), 302–308 (2014).
- <sup>92</sup>Y. Masugi *et al.*, "Upregulation of integrin  $\beta$ 4 promotes epithelial–mesenchymal transition and is a novel prognostic marker in pancreatic ductal adenocarcinoma," *Lab. Invest.* **95**(3), 308–319 (2015).
- <sup>93</sup>K. Birkenkamp-Demtroder, "Differential gene expression in colon cancer of the caecum versus the sigmoid and rectosigmoid," *Gut* **54**(3), 374–384 (2005).
- <sup>94</sup>A. Barrier *et al.*, "Colon cancer prognosis prediction by gene expression profiling," *Oncogene* **24**(40), 6155–6164 (2005).
- <sup>95</sup>T. L. Moser *et al.*, "Interaction of urinary-type plasminogen activator with extracellular matrix proteins: Specific binding to vitronectin," *Fibrinolysis* **8**, 34 (1994).
- <sup>96</sup>J. C. T. Wong *et al.*, "Absence of MMP2 expression correlates with poor clinical outcomes in rectal cancer, and is distinct from MMP1-related outcomes in colon cancer," *Clin. Cancer Res.* **17**(12), 4167–4176 (2011).
- <sup>97</sup>C. Kahlert *et al.*, "Tumour-site-dependent expression profile of angiogenic factors in tumour-associated stroma of primary colorectal cancer and metastases," *Br. J. Cancer* **110**(2), 441–449 (2014).
- <sup>98</sup>T. Oshima, C. Kunisaki, K. Yoshihara *et al.*, "Clinicopathological significance of the gene expression of matrix metalloproteinases and reversion-inducing cysteine-rich protein with Kazal motifs in patients with colorectal cancer: MMP-2 gene expression is a useful predictor of liver metastasis from colorectal cancer," *Oncol. Rep.* **19**(5), 1285–1291 (2008).

- <sup>99</sup>K. H. Koh *et al.*, “Differential gene expression profiles of metastases in paired primary and metastatic colorectal carcinomas,” *Oncology* **75**(1–2), 92–101 (2008).
- <sup>100</sup>A. Polistena *et al.*, “MMP7 expression in colorectal tumours of different stages,” *In Vivo* **28**(1), 105–110 (2014).
- <sup>101</sup>M. Illemann, “MMP-9 is differentially expressed in primary human colorectal adenocarcinomas and their metastases,” *Mol. Cancer Res.* **4**(5), 293–302 (2006).
- <sup>102</sup>D. W. Murray, “Molecular mechanisms of human colorectal cancer metastasis: Identification of novel biomarkers,” Ph.D. thesis (Dublin City University, 2004).
- <sup>103</sup>S. Bao *et al.*, “Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway,” *Cancer Cell* **5**(4), 329–339 (2004).
- <sup>104</sup>Z. Li *et al.*, “Periostin expression and its prognostic value for colorectal cancer,” *Int. J. Mol. Sci.* **16**(6), 12108–12118 (2015).
- <sup>105</sup>F. J. Sulzmaier, C. Jean, and D. D. Schlaepfer, “FAK in cancer: Mechanistic findings and clinical applications,” *Nat. Rev. Cancer* **14**, 598–610 (2014).
- <sup>106</sup>M. Heffler *et al.*, “FAK and HAS inhibition synergistically decrease colon cancer cell viability and affect expression of critical genes,” *Anticancer Agents Med. Chem.* **13**(4), 584–594 (2013).
- <sup>107</sup>P. A. Kiely *et al.*, “Insulin-like growth factor I controls a mutually exclusive association of RACK1 with protein phosphatase 2A and I integrin to promote cell migration,” *Mol. Cell. Biol.* **26**(11), 4041–4051 (2006).
- <sup>108</sup>A. Saito, “Detection of genes expressed in primary colon cancers by *in situ* hybridisation: Overexpression of RACK 1,” *Mol. Pathol.* **55**(1), 34–39 (2002).
- <sup>109</sup>J.-F. Liang *et al.*, “Relationship and prognostic significance of SPARC and VEGF protein expression in colon cancer,” *J. Exp. Clin. Cancer Res.* **29**(1), 1–11 (2010).
- <sup>110</sup>L. Geng *et al.*, “TGF- $\beta$  suppresses VEGFA-mediated angiogenesis in colon cancer metastasis,” *PLoS One* **8**(3), 25 (2013).
- <sup>111</sup>M. Krzystek-Korpacka *et al.*, “Expression stability of common housekeeping genes is differently affected by bowel inflammation and cancer: implications for finding suitable normalizers for inflammatory bowel disease studies,” *Inflammatory Bowel Dis.* **20**(7), 1147–1156 (2014).