Clinical applications of gene expression in colorectal cancer

Elrasheid A.H. Kheirelseid, Nicola Miller, Kah Hoong Chang, Mary Nugent, Michael J. Kerin

Department of Surgery, National University of Ireland Galway, Ireland
Corresponding to: Elrasheid A.H. Kheirelseid. Department of Surgery, National University of Ireland Galway, Clinical Science Institute, Costello Road, Galway, Ireland. Email: rashmed1111@gmail.com.

Background: Despite developments in diagnosis and treatment, 20% of colorectal cancer (CRC) patients present with metastatic disease and 30% of cases recur after curative surgery. Furthermore, the molecular factors involved in prognosis and response to therapy in CRC is poorly understood. The aims of this study were to quantitatively examine the expression of target genes in colorectal cancer and to correlate their expression levels with clinico-pathological variables.

Methods: A detailed analysis of published CRC microarray data was performed to identify the most prominent genes. The selected genes were validated in fifty-two pairs of fresh colorectal tumour and associated normal tissue specimens by RQ-PCR using TaqMan® assays. Statistical analysis and correlation with clinicopathological data was performed using SPSS software.

Results: Expression levels of *CXCL12* (P=0.000), *CDH17* (P=0.026), *MUC2* (P=0.000), *L-FABP* (P=0.000) and *PDCD4* (P=0.000) were down regulated and *IL8* (P=0.000) was upregulated in tumours compared to normal colorectal tissues. No significant differences were noted in expression of *CEACAM5*, *CXCR4*, *CXCR7*, *TGFB1*, *TGFBR1* and *TGFBR2*. Furthermore, we found significant associations of gene expression levels and clinicopathological variables such as tumour size, grade, invasion and lymph node status.

Conclusions: We identified a comprehensive list of genes with highly differential expression patterns in colorectal cancer that could serve as molecular markers to complement existing histopathological factors in diagnosis, follow up and therapeutic strategies for individualised care of patients.

Key Words: Gene expression in colorectal cancer; molecular profiling in colorectal cancer; neoadjuvant chemoradiation therapy in rectal cancer

Introduction

Colorectal carcinoma is one of the most common types of cancer worldwide with increasing incidence especially in developed countries (1). Despite advances in diagnosis and treatment, this disease remains a serious threat to life for millions of people globally, with approximately 20% of patients presenting with metastatic disease, and 30% of colorectal cancers recurring (2). At the molecular level, activation of oncogenes and inactivation of tumour suppressor genes are processes known to be involved in colorectal carcinogenesis (3). Nevertheless, exactly how those genetic alterations bring about the development and progression of colorectal carcinomas remains to be resolved. To complicate this picture, accumulation of mutated genes in neoplasms tends to be accompanied by other genetic and epigenetic changes including loss of heterozygosity, inactivation of key genes by methylation or loss of imprinting or gene amplifications, all of which have potential to alter gene expression profiles (4). Genome-wide monitoring of gene expression profiles has greatly advanced our understanding of the numerous and diverse events associated with carcinogenesis thusfar. By harnessing recent technological advances in molecular profiling techniques, it is anticipated that greater insight to the various combinations of genetic events or alternative pathways underlying carcinogenesis will be gained.

In order to identify molecules that could serve as biomarkers of disease and therapeutic targets in colorectal cancer we set this study to quantitative candidate genes expression in colorectal cancer tissues using RT-PCR in order to ddetermine the expression levels of candidate genes in tumour and tumour-associated normal colorectal tissue. In

    www.thejgo.org

Table 1 Candidate genes

Gene name	Gene symbol	Location	Assay ID	bp	References
Cadherin 17	CDH17	8q22.1	Hs00184865_m1	72	(8,9)
Carcinoembryonic antigen related cell adhesion molecule 5	CEACAM5	19q13.2	Hs00944023_m1	71	(10,11)
Chemokine ligand 12	CXCL12	10q11.1	Hs00171022_m1	77	(8,12)
Chemokine, cxc motif, receptor 4	CXCR4	2q21	Hs00237052_m1	78	(13,14)
Chemokine, cxc motif, receptor 7	CXCR7	2q37	Hs00171022_m1	129	(13,15)
Fatty acid binding protein 1, liver	FABP1	2p11	Hs00155026_m1	71	(5,16-19)
Interleukin-8	IL-8	4q13-q21	Hs99999034_m1	81	(7,17,20,21)
Mucin2	MUC2	11p15.5	Hs03005094_m1	64	(7,22,23)
Programmed cell death 4	PDCD4	10q24	Hs00205438_m1	94	(24,25)
Transforming growth factor beta1	TGFB1	19q13.1	Hs00998133_m1	57	(5,7,26-28)
Transforming growth factor-beta receptor type 1	TGFBR1	9q22	Hs00610320_m1	73	(29,30)
Transforming growth factor-beta receptor type 2	TGFBR2	3p22	Hs00234253_m1	70	(26,31)

addition, we aimed to investigate correlation between serum carcinoembryonic antigen (CEA) and tissue CEACAM5 levels. Secondary objectives were to ccorrelate candidate genes expression levels and clinicopathological variables.

Materials and methods

Candidate genes

In order to identify a list of genes associated with deregulated expression in colorectal cancer and thereby might have a role in colorectal cancer tumourogenesis, we carried out a detailed analysis of published colorectal cancer microarray data and identify the most prominent genes. Furthermore, a literature review was performed to identify mRNA highly associated with cancer to identify their role in colorectal cancer pathogenecity and progression (5-7). *Table 1* showed the list of candidate genes selected for analysis in this study

Study groups

Clinicopathological data on all patients were examined in order to select suitable samples for study groups appropriate to address specific questions. A heterogeneous group of 107 patients with colorectal tumours, all of which had matched tumour-associated normal (TAN) samples was selected for gene expression profiling experiment using real-time quantitative (RQ)-PCR (*Table 2*). Tissue samples were gathered from consenting patients at the time of diagnostic procedures or at primary curative surgical resection at Galway University Hospital, Ireland. The cohort comprised of 101 colorectal tumour specimens, 8 polyps and 107 TAN tissues. Following retrieval, all samples were subject

to histopathological review prior immediate snap-freezing in liquid nitrogen and archival at −80 ℃ until further use. Concomitant clinicopathological data on patients and specimens was obtained through patient interview and review of clinical notes. Ethical approval for this study was granted by the Clinical Research Ethics Committee, Galway University Hospitals.

RNA extraction and analysis

Tissue samples (50-100 mg) were homogenised using a hand-held homogenizer (Polytron PT1600E) in 1-2 mL of QIAzol reagent (Qiagen) as described previously (32). In brief, tumour and TAN samples were homogenised separately but on the same day. RNA was extracted using RNeasy Plus Mini kits (Qiagen) according to the manufacturer's instructions. RNA was eluted in 60 µL volumes and stored at −80 ℃. RNA concentration and purity was assessed in duplicate samples using a using a NanoDrop ND-1000 spectrophotometer (NanoDrop). RNA integrity was evaluated using the RNA 6000 Nano Chip kit (Series II) and the Agilent 2100 Bioanalyzer (Agilent Technologies). An RNA integrity number (RIN) was generated for each sample using the Agilent 2100 Expert Software (Version B.02.03) based on the ratio of ribosomal bands and also the presence or absence of degradation products on the electrophoretic and gel-like images. A threshold value of RIN ≥7 was applied and RNA purity was verified by an average A260/A280 ratio of 1.98 (range, 1.97-2.01) and A260/A230 ration of 1.7 (range, 1.5-1.83).

Reverse transcription

RNA was reverse transcribed to first strand cDNA using

Table 2 Clinico-pathological data for patients used for gene expression analysis

Clinicopathological variable	Number of patients N=[107]
Tissue type	
Carcinoma	101
Polyp	8
Gender	
Male	67
Female	40
Mean age (SD)	69.72 (11.89)
Tumour location	
Colon	43
Rectum	58
Tumour location	
Proximal	27
Distal	74
Tumour thickness (mm)	
<10	23
10-15	33
>15	22
Unknown	23
Tumour diameter (mm)	
<30	29
30-40	26
>40	31
Unknown	15
Distant metastasis	
M0	80
M1	21
Nodal status	
N0	22
N1	11
N2	9
UICC stage	
Stage 0	2
Stage I	17
Stage II	28
Stage III	28
Stage IV	21
Stage V	5
pCR	
Tumour differentiation	
Grade 1: well differentiated	11
Grade 2: moderate differentiated	72
Grade 3: poor differentiated	10
Not applicable	8
Mucin secretion	
Mucinous	19
Non-mucinous	82

Superscript III reverse transcriptase (Invitrogen) and random primers (N9; 1 µg, MWG Eurofins). Negative control samples were included in each set of reactions. Reactions were incubated at 25 ℃ for 5 minutes followed by 50 ℃ for 1 hour and final denaturation at 72 ℃ for 15 minutes. Samples were subsequently diluted to 100 µL in nuclease-free water and stored at −20 ℃.

Real-time quantitative PCR

Amplification efficiency

The amplification efficiency of each assay is an important consideration in the determination of relative quantities of gene expression by RQ-PCR. PCR efficiency impacts greatly on the accuracy of the calculated expression result and is influenced by PCR reaction components. For 100% efficiency a doubling of the amount of DNA will occur at each cycle, while for 80% and 70% the amount of DNA will increase from 1 to 1.8 and 1.7, respectively. Resultantly, small differences in efficiency can greatly affect the calculation parameters involved in the determination of gene expression values. Amplification efficiencies for each gene assay in this study were calculated applying the formula $E=(10^{-1}/slope^{-1}) \times 100$, using the slope of the plot of Ct versus log input of cDNA (10-fold dilution series). A threshold of 10% above and below 100% efficiency was applied.

Endogenous control

Relative quantification is the most widely adopted approach whereby quantification of gene expression is normalised relative to an endogenously expressed control (EC) gene(s). Central to the reliable determination of gene expression is the choice of control gene. *B2M* and *PPIA* have previously been identified as the most stably expressed genes in a large cohort of colorectal tissues (33) and were used to normalise expression values in the present study.

RT-PCR of mRNA

The expression of each EC gene was analysed by RQ-PCR using TaqMan gene expression assays using a 7900HT instrument (Applied Biosystems). All reactions were performed in 10 µL reactions, in triplicate within the same PCR run. Negative controls were included for each gene target under assay. On each plate, an interassay control was included to account for any variations between runs. For each well 2 µL of cDNA from each sample was added to 18 µL of PCR reaction mix which consisted of 10 µL TaqMan master mix, nuclease free water and 1 µL gene expression assay primer-probe mix (Applied Biosystems). Standard fast thermal cycling parameters of 40 cycles of

Table 3 Clinicopathological correlations of candidate genes expression in CRC

Variable	CXCL12	CXCR4	CXCR7	IL8	TGFB1	TGFBR1	TGFBR2	CDH17	CEA	FABP1	MUC2	PDCD4
Tumour diameter	0.481	0.860	0.035*	0.285	0.766	0.189	0.155	0.043*	0.213	0.449	0.271	0.674
Tumour thickness	0.094	0.242	0.036*	0.616	0.348	0.317	0.234	0.035*	0.282	0.049	0.654	0.052
Wall involvement	0.019*	0.005*	0.002*	0.055	0.201	0.000*	0.006*	0.002*	0.002*	0.949	0.400	0.886
Tumour location	0.020*	0.381	0.021*	0.285	0.347	0.621	0.003*	0.473	0.858	0.466	0.711	0.007*
Tumour differentiation	0.043*	0.043*	0.596	0.008*	0.335*	0.443	0.560	0.045*	0.016*	0.109	0.910	0.209
Mucin secretion	0.342	0.272	0.679	0.115	0.585	0.225	0.870	0.407	0.782	0.398	0.013*	0.217
Depth of invasion	0.001*	0.093	0.485	0.319	0.828	0.217	0.116	0.587	0.442	0.389	0.645	0.487
Lymph node status	0.040*	0.059	0.287	0.015*	0.238	0.688	0.049*	0.175	0.071	0.976	0.716	0.934
Distant metastasis	0.163	0.044*	0.138	0.062	0.861	0.161	0.235	0.434	0.547	0.373	0.938	0.443
Perineural invasion	0.389	0.126	0.904	0.670	0.792	985	0.030*	0.180	0.154	0.057	0.443	0.969
LV invasion	0.033*	0.132	0.020*	0.0687	0.035	0.208	0.012*	0.016*	0.019*	0.515	0.131	0.600
Tumour stage	0.016*	0.253	0.749	0.676	0.490	0.534	0.370	0.949	0.681	0.704	0.911	0.649
Recurrence	0.169	0.058	0.124	0.036*	0.476	0.126	0.213	0.891	0.801	0.170	0.685	0.023*

*Statistically significant (P<0.05)

95 ℃ for 15 seconds and 60 ℃ for 60 seconds were applied in accordance with the manufacturer's recommendations.

Relative quantification

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. QBasePlus software (Biogazelle) was used to calculate expression values of each chemokine target. Relative quantities were corrected for efficiency of amplification and fold change in gene expression between groups was calculated as $E^{-\Delta\Delta Ct} \pm$ s.e.m. The lowest expressed sample was used as a calibrator.

$-\Delta\Delta Ct = (C_t$ target gene, test sample $- C_t$ endogenous control, test sample) $- (C_t$ target gene, calibrator sample $- C_t$ endogenous control, calibrator sample).

Statistical analysis

Statistical analysis was carried out with IBM SPSS Statistics 17.0 (SPSS Inc.). Data was tested for normal distribution graphically using histograms and also using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Parametric tests were used where appropriate. One-way ANOVA and independent t-test were used to determine association and comparisons between independent groups. Correlation analysis used Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively. Univariate analysis and paired-T test were used to assess related samples. The statistical significance of differences in survival between groups was determined by log rank which compares differences along all points of the curve and multivariate analysis was done using Cox regression. P values <0.05 were considered statistically significant.

Results

Gene expression and clinicopathological parameters

The expression of CDH17 was significantly lower in colorectal cancer compared to TAN tissues (P<0.001, t-test, *Figure 1*). Regarding the clinicopathological variables, the CDH17 expression significantly increased with increased tumour diameter (P=0.043) and tumour thickness (P=0.035), however, its expression reduced with increased bowel wall involvement (P=0.002) (*Table 3*). This finding could be explained by CDH17 adhesion function. Its expression was also reduced in poorly differentiated tumours (P=0.045) and in patients with increased CA 19.9 serum level (P=0.014) (Kruskal-Wallis and Mann-Whitney tests, *Table 3*).

Reduced expression of *FABP1* was observed in a progressive manner from TAN, to polyp, to tumour (P<0.001, Kruskal-Wallis t-test, *Figure 1*). Between groups analysis revealed significant differences in FABP1 expression levels between tumour and TAN (P<0.001) and between polyps and TAN (P=0.001), but not between tumours and polyp (P=0.055). There was no significant association of FABP1 with other clinicopathological variables of the colorectal tumours (*Table 3*).

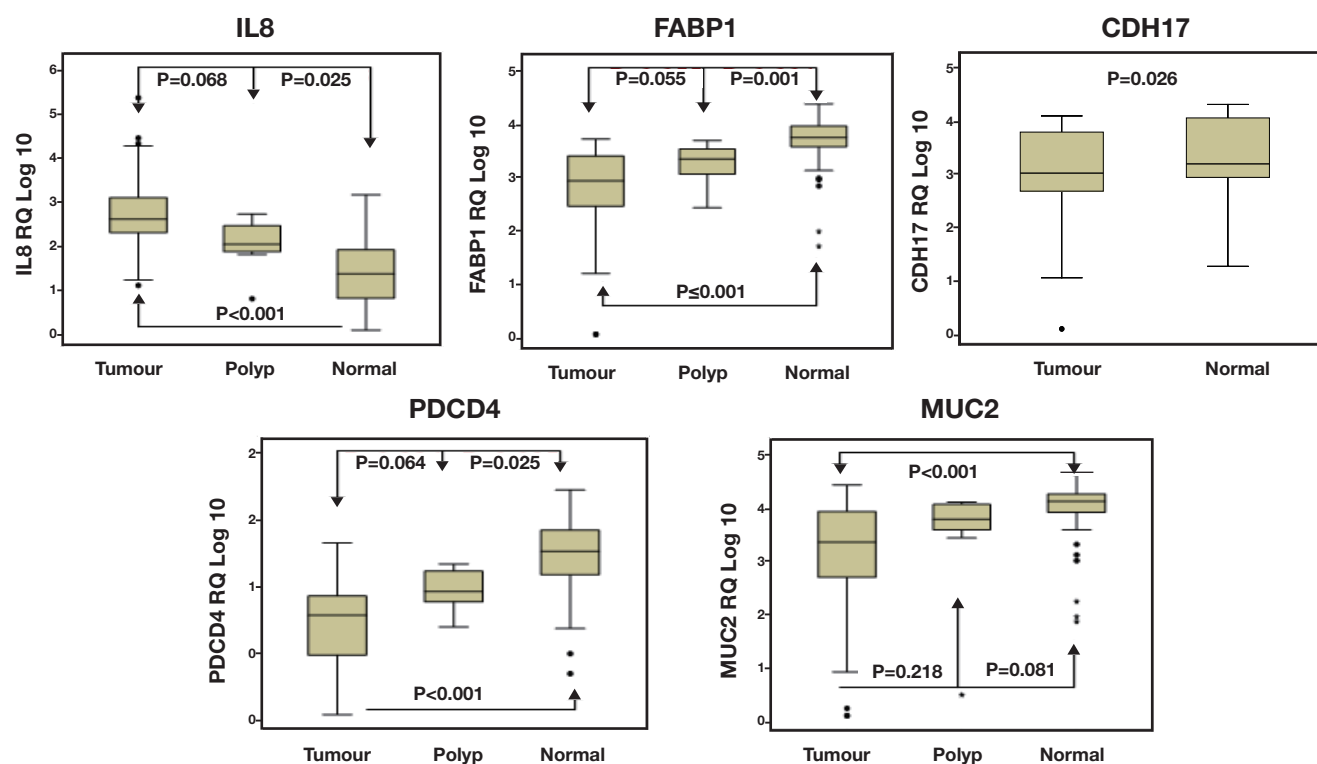Expression levels of *IL-8* increased progressively from

Figure 1 Gene expression in CRC tumour & normal tissue

tumour-associated normal, to polyps, to tumours (P<0.001, ANOVA). Post-Hoc Tukey analysis revealed significant differences in *IL-8* expression levels between tumour and TAN (P<0.001) and between polyps and TAN (P=0.025), but not between tumours and polyp (P=0.068) (*Figure 1*).

Although the expression of *IL-8* increased in tumours compared to normal colorectal tissues, its reduced expression was significantly associated with poor differentiation (P=0.008), advanced nodal stage (P=0.015) and disease recurrence (P=0.036) (ANOVA, *Table 3*). A non-significant trend of reduced *IL-8* expression was also associated with perineural (P=0.670) and lymphovascular invasion (P=0.687), advanced Dukes' stage (P=0.425) and distal metastasis (P=0.062) (ANOVA, *Table 3*).

Again a progressive manner of expression from tumour, to polyp, to tumour associated normal was observed in *MUC2* (P<0.001, Kruskal-Wallis t-test, *Figure 1*). Further analysis confirmed a significant differences in *MUC2* expression levels between tumour and TAN (P<0.001) but not between polyps and TAN (P=0.081), and between tumours and polyp (P=0.218). MUC2 expression was higher in mucinous tumours compared to non-mucinous (P=0.013, Mann-Whitney test); however, it was reduced in patients with high CA 19-9 serum level (P=0.037) (Mann-Whitney test, *Table 3*).

PDCD4 showed step-wise increase in expression from

tumours, to polyps, to tumour associated normal tissues (P<0.001, ANOVA, *Figure 1*). Further between groups analysis (Post-Hoc Tukey test) identified significant differences in expression between tumour and TAN (P<0.001) and between polyp and TAN (P=0.002) but not between tumour and polyp (P=0.065). Additionally, down-regulation of PDCD4 was significantly associated with proximal colon tumours (P=0.007), tumour recurrence (P=0.023) and raised CA19.9 serum level (P=0.003) (*t*-test, *Figure 1*, *Table 3*)

Paired t tests were used to investigate differences in gene expression between 101 paired tumour and normal colorectal tissues. *CXCR4* expression levels were thus found to be higher in tumours in contrast to *CXCL12* which was expressed at lower levels in tumour versus normal tissue. However, these differences only reached statistical significance in relation to *CXCL12* (P<0.001) (*Figure 2*). No difference in *CXCR7* expression was noted between tumour and TAN tissue (*Figure 2*). Although a significant difference was observed in *CXCL12* expression in tumour and polyp compared to TAN tissue (P<0.001 and P<0.003, respectively), no difference was found between tumours and polyps (P=0.907) (*Figure 2*, ANOVA).

The relationship between *CXCL12*, *CXCR7* and *CXCR4* was further investigated using Pearson correlation. Preliminary analysis was performed to ensure no violation
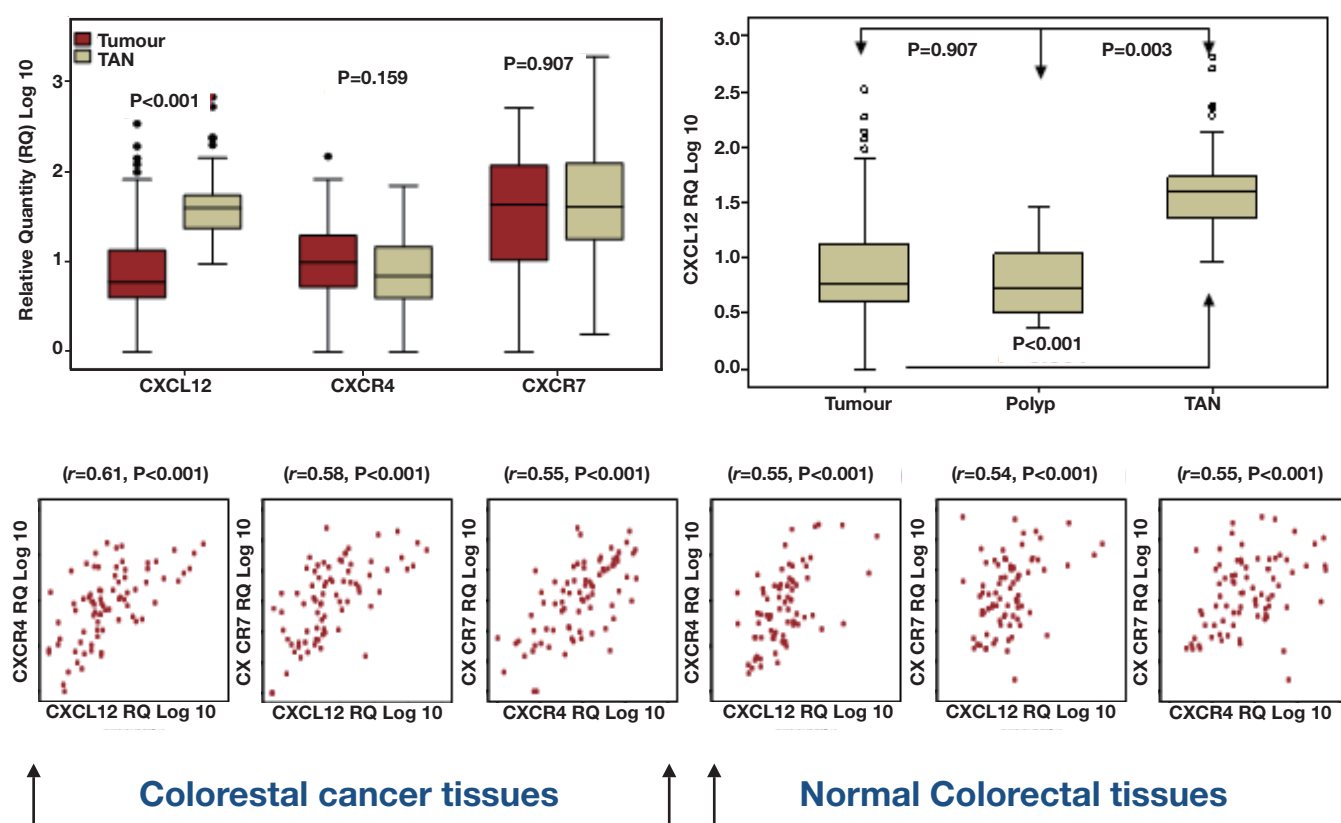
Figure 2 Chemokine expression in CRC tumour & normal tissue

of the assumption of normality, linearity and homogenecity. Strong positive correlation between all variables in both tumour and normal was observed, with high expression of the ligand associated with high expression of its receptors (*Figure 2*).

One-way ANOVA and t-tests were conducted to explore the relationship between chemokine expression and clinicopathological parameters. Both *CXCL12* and *CXCR*7 were significantly under-expressed in proximal colon. Reduced expression of *CXCL12* and both receptors was significantly associated with survival (P=0.010), advanced stage (P=0.040), poor differentiation (P=0.043), and tumour size (P<0.05), invasion and metastasis (P=0.044) (*Figure 3*).

Significant differences in overall patient survival were observed in tumours with higher (above median) *CXCR*7 expression in comparison to those with lower *CXCR*7 expression (below median) (log rank test P<0.010, *Figure 3*). With median follow up of 15 months, *CXCR*7 under-expressers (below median) had a high mortality from colorectal cancer with mean survival of 27 months compared to 46 months in over-expressers (*CXCR*7 above median). A multivariate Cox regression analysis was used to determine the prognostic factors for overall survival. After simultaneous adjustment of all these variables there

continue to be a significant difference in survival between both groups (P=0.044).

TGFB1 expression levels were higher in tumour compared to TAN tissues (*P=0.109*, paired t-test, *Figure 4*) in contrast to the expression of its receptors *TGFBR1* and *TGFBR2* whish showed low expression trends in tumour compared to TAN (P=0.044 and *0.460* respectively, paired t-test, *Figure 4*). Interestingly, TGFB1 expression showed step-wise increase from polyp, to normal, to tumour (P=0.016, ANOVA). Further analysis (Post-Hoc Tukey test) pointed out significant differences in expression between tumours and polyps (P=0.029), but not between tumours and TAN (P=0.345) and between polyps and TAN (P=0.914) (*Figure 4*).

The relationship between *TGFB1*, *TGFBR1* and *TGFBR2* was further investigated using Pearson correlation. No violation of the assumption of normality, linearity and homogenecity was ensured before conducting further analysis. There was positive correlation between all the variables in both tumour and TAN colorectal tissues with high expression level of the ligand associated with high expression of the receptors (*Table 3*).

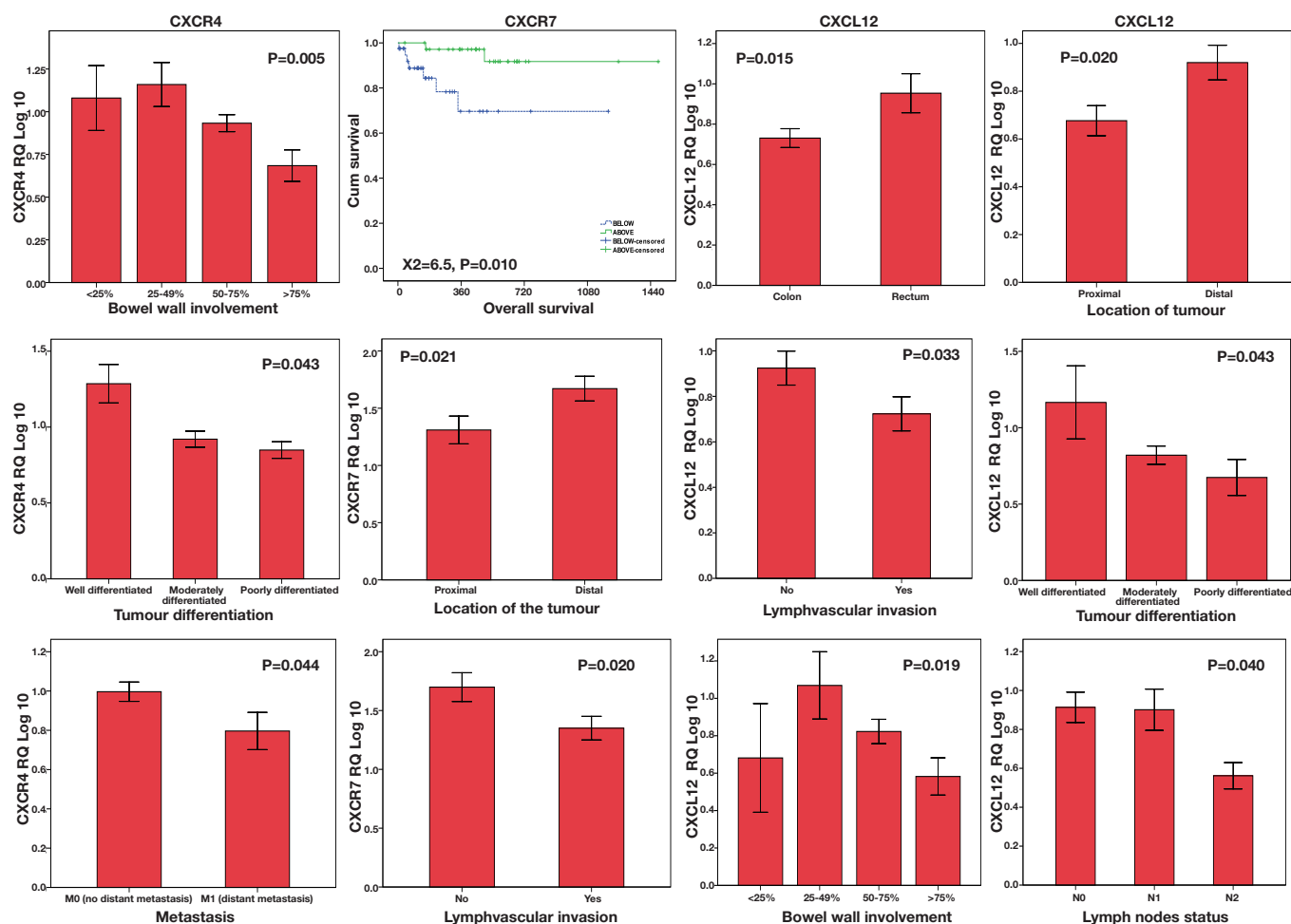The relation of *TGFB1* and its receptors expression levels and the clinico-pathological parameters were

Figure 3 Chemokine expression levels and clinicopathological parameters in CRC

examined using ANOVA and t-test (*Figure 4*). Although high level of *TGFB1* was documented in tumours compared to normal colorectal tissues, we noticed an association of TGFB1 down-regulation and lymphovascular invasion (P=0.035). Both *TGFBR1* and *TGFBR2* were under-expressed in proximal colon, however, the difference was only significant for *TGFBR2* (P=0.003). *TGFBR1* showed reduced expression in association with advanced disease clinicopathological parameters like tumour size, poor differentiation, advanced nodal stage, advanced Dukes' stage and tumour invasion and metastasis (*Table 3*), However, these associations were only significant in relation to bowel wall involvement (P<0.001), and raised CEA serum level (P=0.045). Down-regulation of *TGFBR2* was significantly associated with increased bowel wall involvement (P=0.006), in colon cancer compared to rectal cancer (P=0.031) and in association with perineural (P=0.030) and lymphovascular invasion (P=0.012).

No significant differences were identified in *CEACAM5* expression levels in tumour compared to TAN colorectal tissues (P=0.981, t-test). In addition, no significant

correlations were found between *CEACAM5* expression and the CEA serum level (r=−134, n=79, P=0.240). Higher expression of CEACAM 5 was associated with moderately differentiated tumours (P=0.016) and local (P=0.002) and lymphovascular invasion (P=0.019) (Kruskal-Wallis and Mann-Whitney tests, *Table 3*).

Neoadjuvant therapy and colorectal cancer genes expression

In the cohort of rectal cancer patients (n=58) we analysed the differences in gene expression in patients who had neoadjuvant chemoradiation (n=25) compared to those who did not (n=33) using t-test. Univariate analysis of variance was further conducted to test for interaction effect and to control for confounding factors. We demonstrated decrease expression of *CDH17* (P=0.020) and *CEACAM5* (P=0.032) and increase expression of *CXCL12* (P<0.001), *CXCR4* (P=0.004) and *MUC2* (P=0.041) in response to neoadjuvant therapy. However, the differences only persisted for *CXCL12* (P=0.035) and CXCR4 (P=0.001) after univariate analysis (*Figure 5*).
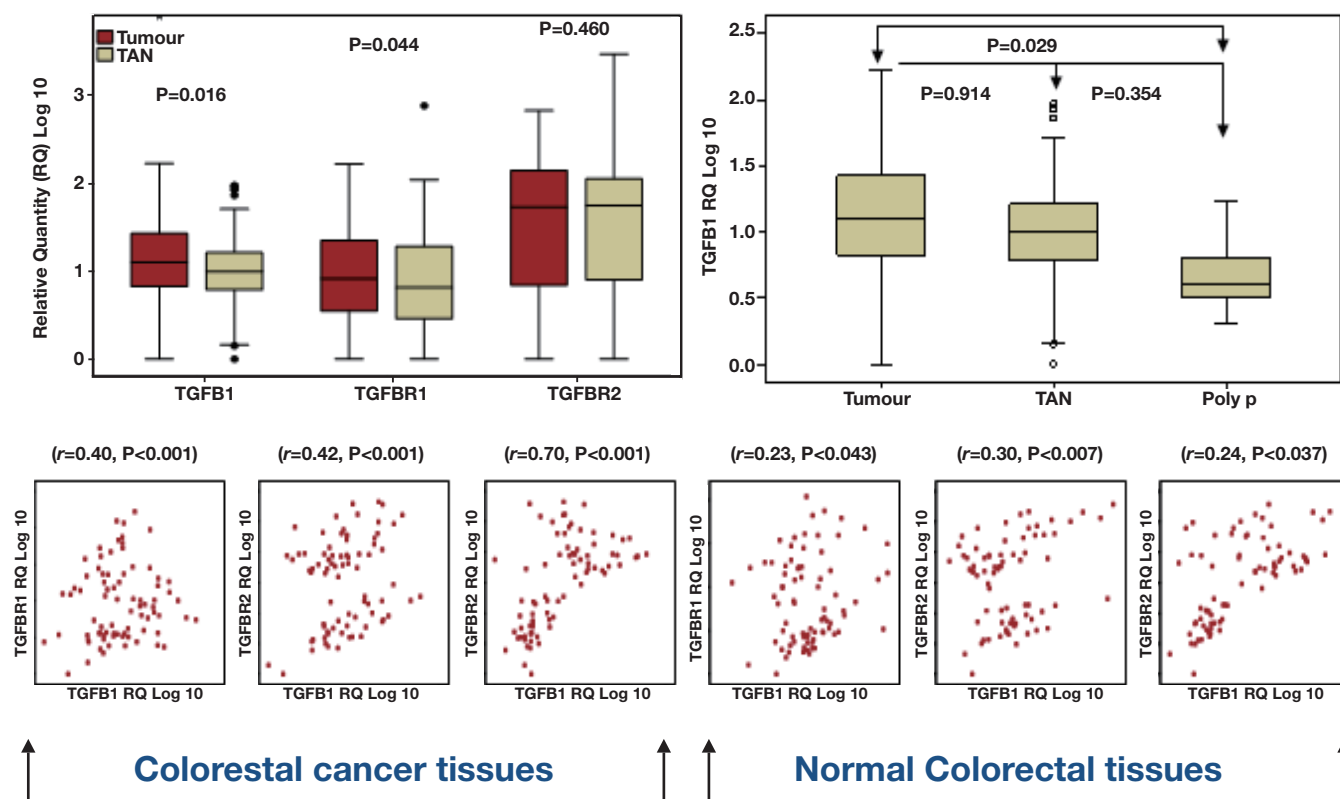
Figure 4 TGFB1 and its receptors expression in CRC tumour & normal tsssue
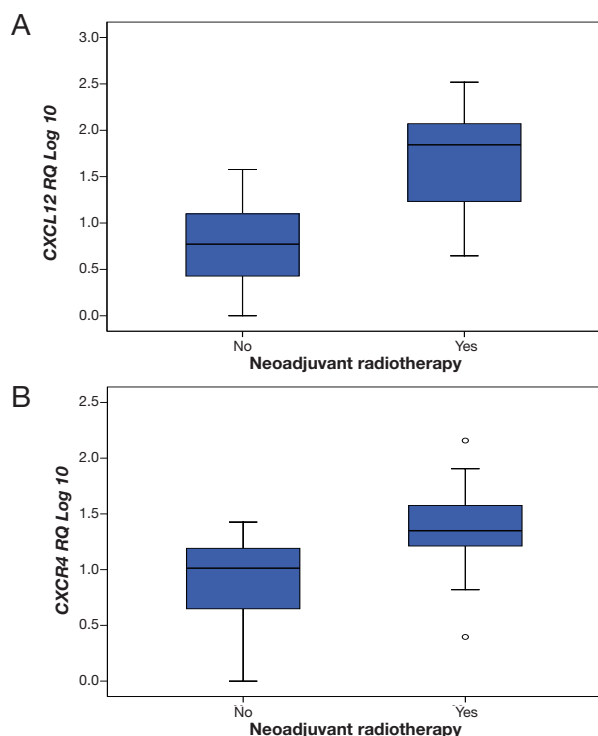

Figure 5 Dysregulation of gene expression in response to neoadjuvant CRT. Neoadjuvant chemoratiation associated with sigificat up-regulation of CXCL12 (A, univariate analysis, P=0.0035) and CXCR (B, univariate analysis, P=0.001) expression

Interestingly, expression levels of *CDH17* (P=0.003), *CEACAM5* (P=0.036), *CXCL12* (P≤0.001) and *CXCR4* (P=0.003) significantly correlated with Mandard tumour regression grade (TRG). Higher expression of *CXCL12* and *CXCR4* was noticed in good responders (TRG1, TRG2 and TRG3) compared to poor responders (TRG4 and TRG5) in contrast to the expression levels of *CDH17* and *CEACAM5* which were lower in good responders (ANOVA test, *Figure 6*)

Discussion

Colorectal cancer is the fourth most common cancer in men and the third most common cancer in women worldwide (34). In the USA, colorectal cancer is the second most common cause of cancer death among men aged 40 to 79 years and accounts for 9% of all cancer related deaths (35). In Ireland, the National Cancer Registry predicts that the incidence of colorectal cancer will increase from 2,111 cases in 2005 to 5,537 in 2035 (36), indicating a more than 100% increase over the next 30 years. In this setting of increasing disease burden, translational research is of vital importance to clinical advancement. At the molecular level, activation of oncogenes and inactivation of tumour suppressor genes (3) are processes known to be involved
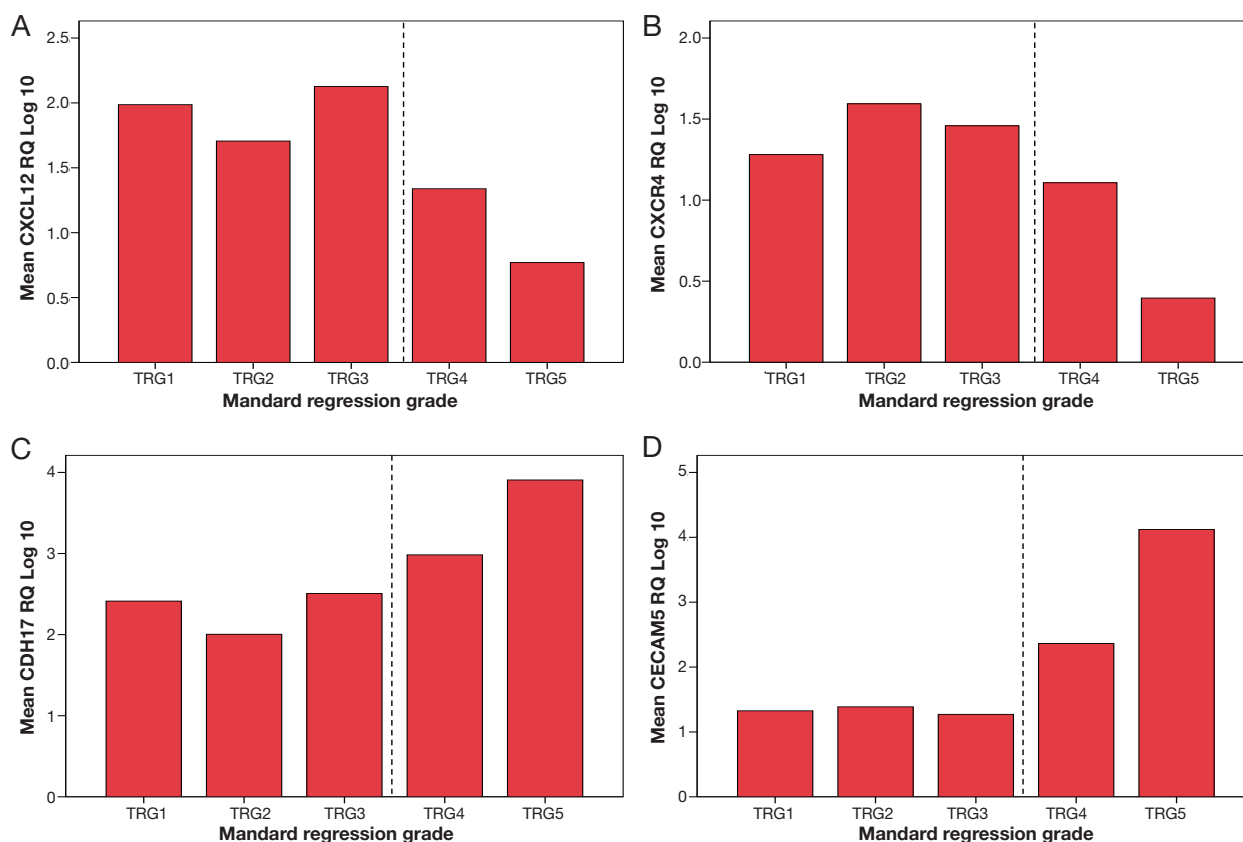
Figure 6 Correlation of gene expression with tumour regression grade. Increased expression of CXCL12 (A, P<0.001) and CXCR4 (B, P=0.003) was associated with lower TRG (good response) in constrast to CDH17 (C, P=0.003) and CEACAM5 (D, P=0.036)

in colorectal carcinogenesis. Additionally, abrogation of mismatch repair systems (37) contributes to some colorectal cancers. Nevertheless, exactly how these genetic alterations bring about the development and progression of colorectal carcinomas remains to be resolved. To complicate the picture, accumulation of mutant genes in neoplasms tends to be accompanied by other genetic and epigenetic changes including loss of heterozygosity, inactivation of important genes by methylation or loss of imprinting (4) or gene amplifications, all of which can alter gene expression profiles. Therefore, genome wide monitoring of gene expression is of great importance if we are to disclose the numerous and diverse events associated with carcinogenesis. Molecular profiling, a tool of genome monitoring, is an attempt to identify the different combinations of genetic events or alternative pathways that may be represented by cancers of a similar type.

The principle of an adenoma-carcinoma sequence, described in 1990, postulates that the transition from adenoma to carcinoma is associated with an accumulation of genetic events in key regulatory genes that confer a growth advantage to a clonal population of cells (38). Since then, although molecular detection methods based on gene

mutation determination have been carried out for several years, the clinical utility of the many molecular markers and their clinical applications remain limited for colorectal cancer patients. Therefore, there is real need for new molecular markers to improve tumour subclassification and prediction of clinical outcome.

Microarray technology and gene expression profiling studies in colorectal cancer stimulated an interest in potential results that could be directly used in the routine clinical setting. Gene expression signatures predictive of disease outcome and response to adjuvant therapy have been generated and are being evaluated in the clinical setting. Such molecular diagnostics and their promise of tailored therapy generated much excitement among researchers however they have yet to be fully incorporated into today's standard of care as they are limited by difficulties in reproducibility, standardisation and lack proof of significance beyond traditional prognostic tools.

One of the primary aims of this study was to characterise the expression profiles of candidate genes in colorectal tissue. Rigourous evaluation of appropriate genes with which to normalise real-time quantitative PCR data identified *PPIA* and *B2M* as the most stably expressed genes

in colorectal tissue samples. This enabled the development of a robust experimental approach which ensured that subsequent profiling of gene expression levels would be measured accurately and reproducibly in colorectal tissue. As a result, a comprehensive list of genes with highly differential expression patterns was derived.

CXCL12 and its receptors CXCR4 and CXCR7

The first candidates to be examined were the chemokine *CXCL12* and its receptors *CXCR4* and *CXCR*7, whose gene expression levels were, determined in 107 tumour and tumour associated normal colorectal tissues, the largest patient cohort reported to date. Significant down-regulation of *CXCL12* in tumour compared to normal colorectal tissue was found, in contrast to *CXCR4*, which showed non-significant up-regulated expression levels in tumour tissues. The reduced expression of *CXCL12* was noticed in both polyps and tumours. This could be explained by the role of *CXCL12* in tumour immunology; however, it may highlight a possible tumour suppressor function of this gene. Investigation of the interaction between *CXCL12*, *CXCR4* and *CXCR*7 may provide some understanding of their functions and the role of each gene in regulating the expression of the others. Despite the reciprocal patterns of expression, strong positive correlation of *CXCL12/CXCR4* and *CXCL12/CXCR*7 in both tumour and normal colorectal tissue was found. Moreover, *CXCR4* and *CXCR*7 expression patterns correlated in the same manner. Saigusa *et al.* also reported significant positive correlation between expression levels of *CXCL12* and *CXCR4* in patients with rectal cancer who underwent preoperative CRT. Moreover, the expression of *CXCR*7 in *CXCR4* positive cells appears to enhance the responsiveness to *CXCL12* as reported by Sierro (39). These findings suggest a possible receptor interaction in tumour and normal colorectal tissues.

Correlation of gene expression levels with clinicopathological data indicated that levels of *CXCL12* and *CXCR*7 were lower in the proximal colon. This may indicate a possible role of this axis in microsatellite instability (MSI), as tumours associated with MSI arise mainly in the proximal colon. Down-regulation of *CXCL12* and its receptors was also found to be associated with increased tumour size, local invasion, poor differentiation, advanced nodal stage, advance tumour stage and lymphovascular invasion. Of further interest, we identified for the first time the prognostic significance of *CXCR*7 mRNA in colorectal cancer. We found that patients with high expression of *CXCR*7 in their tumour cells lived longer than their counterparts with lower *CXCR*7 gene expression. This was further confirmed by multivariate analysis.

TGFB1 and its receptors TGFBR1 and TGFBR2

Although no significant differences were identified in gene expression levels of the chemokine receptor molecules *TGFBR1* and *TGFBR2* in tumour versus normal tissue, the expression of their ligand *TGFB1* was found to be significantly lower in polyps and higher in tumours compared to normal tissue. These findings confirm previous work by Daniel *et al.* [2007], investigating TGFB1 protein expression by IHC in colorectal cancer. The authors demonstrated than in high-grade dysplastic polyps, than in low-grade dysplastic polyp (40). Matsushita *et al.* [1999] found that *TGFB* receptor mRNA was expressed mainly by normal and adenoma colorectal tissues whereas *TGFB1* expressed by cancer (41). Moreover, the significant positive correlation between *TGFB1* and the expression levels of its receptors in both tumour and normal tissue confirms that their role in colorectal cancer is more complex than a simple legend-receptor feedback.

Interestingly, we identified for the first time the relationship of *TGFB* pathway and some established prognostic clinicopathological parameters. Low expression of *TGFBR1* was found to be associated with raised CEA serum level and local tumour invasion. In addition, *TGFBR2* down-regulation was associated with local, perineural and lymphovascular invasion and advanced nodal stage. These findings will further confirm the role of *TGFB* receptors as tumour suppressor. The down-regulation of *TGFBR2* in proximal compared to distal tumours was described before and highlights the role of this gene in microsatellite instable tumours.

Tumours of proximal and distal parts of the colon may form different but related groups of tumours because of their different embryological origin, different exposure to bowel contents and differences in clinical presentation, progression and possible genetic and environmental epidemiology (42).

Many previous studies have examined the relationship between *TGFB* pathway and the disease progression in colorectal cancer. Nevertheless, this is the first study to explore the relation of *TGFB1* and its receptors mRNA in colorectal cancer using RT-PCR. Moreover, the large cohort of patients in this study gives it further advantage compared to the other studies.

Other genes shown to be potential biomarkers in this study included *CDH1*7, *FABP1*, *IL8*, *MUC2* and *PDCD4*. In colorectal cancer, *CDH1*7 expression was only investigated at protein level using IHC and immunoblotting. Hinoi *et al.* examined the protein expression in human colorectal cancer cell lines. In their study, *CDH1*7 was not detected in cell lines showing dedifferentiated phenotypes (43). This was further confirmed by Takamura *et al.* who examined the

*CDH1*7 expression in four cell lines and 45 human primary colorectal carcinoma using monoclonal antibodies. In cell lines the protein was expressed in differentiated but not the dedifferentiated phenotypes while in tissues reduced *CDH1*7 expression was associated with high tumour grade, advanced stage and lymphatic invasion and metastasis (44). Moreover, Kwak *et al.* found reduced expression in 51% of the 207 colorectal cancers he studied using immunohistochemistry and he significantly correlated down-expression of *CDH1*7 with poor survival and lymph nodes metastasis (45). To our knowledge, this is the first study to investigate *CDH1*7 mRNA in colorectal cancer using RQ-PCR. Our findings support the above reports and confirm that down-regulation of *CDH1*7 in colorectal cancer is associated with poor differentiation, raised CA19.9 tumour marker serum level and local tumour invasion indicated by increase bowel wall involvement. Interestingly, *CDH1*7 expression correlated with increased tumour diameter and tumour thickness (indices of intraluminal tumour growth) and decreased with increased bowel wall involvement (index of local tumour invasion). Those findings could be explained by the adhesion function of the protein. Generally, for the tumour to grow in diameter and thickness it needs to retain adhesion molecules expression, while loss or inactivation of those adhesion molecules correlate with inhibition of cell aggregation and promotion of tumour invasiveness. This finding may highlight the potential role of *CDH1*7 as a marker for rectal cancer surgical management planning. In other wards, decrease level of *CDH1*7 may indicate local invasion of tumour and therefore total mesorectal excision (TME) will be indicated.

Evidence of dysregulated *FABP1* gene expression has been reported in colorectal gene expression array datasets (5,46), however, little is known of its expression profile with regard to clinical data. Lawrie *et al.* identified consistent loss of *FABP1* in tumour compared to normal colon and also noted the association of decreased protein expression and poorly differentiated tumours and large adenomas (47). Moreover, *FABP1* expression was found to be associated with good prognosis after liver resection of colorectal cancer metastasis (48). Although no statistically significant correlation between *FABP1* expression and clinicopathological parameters was identified in this study, we observed that *FABP1* is differentially expressed in normal-adenoma-carcinoma sequence and its loss occurred early in colorectal cancer tumourogenesis. This indicates tumour suppressor function of *FABP1* in colorectal cancer. The loss of *FABP1* in colorectal cancer contrast with the findings in other tumours types which might be explained by the organ-specific distribution and the different role of *FABP1* through distinct intracellular interacting molecules.

In keeping with the previous reports, we noted overexpression of *IL8* in tumour compared to normal colorectal tissue. In addition, we identified a progressive manner of increase gene expression from normal, to polyps, to tumour. The early dysregulation of *IL8* in colorectal cancer suggest that the gene may play a role in carcinogenesis in addition to its confirmed role in tumour progression. Correlations with clinicopathological parameters revealed significant association of reduced *IL8* expression and poor tumour differentiation, advanced nodal stage and disease recurrence. Although the significant of these findings is unclear, it should be considered when planning *IL8* targeting therapy.

Furthermore, we confirmed *MUC2* mRNA down-regulation in non-mucinous and over-regulation in mucinous colorectal cancer. We also showed decreased expression of *MUC2* in a progressive manner from tumour-associated normal, to polyps, to tumours. No significant association of *MUC2* and clinicopathological variables other than CA19.9 serum levels has been determined in this study. Regarding *PDCD4* mRNA, its expression was significantly lower in tumour and polyp compared to tumour-associated tissue in keeping with the protein expression levels described before (46,49,50). Furthermore, we identified the novel association of reduced *PDCD4* expression with disease recurrence and raised CA19.9 serum level. These findings suggest that *PDCD4* involves in both tumour promotion and tumour progression and represent a potential biomarker for evaluating the transition of normal colorectal tissue to adenoma and carcinoma. Reduced expression of PDCD4 in proximal compared to distal colon may indicate a potential role in microsatellite instability (MSI) and Lynch syndrome.

Measurement and quantifying of tumour response to neoadjuvant CRT is an important parameter in order to elucidate factors that may allow for response prediction and planning of next step of treatment in rectal cancer patients. Clinical response (cCR), pathological response (pCR) and tumour downstaging are the commonly used methods to measure response. Both clinical response and tumour downstaging compared the tumour characteristics before and after treatment clinically and using radiological tools like magnetic resonance imaging (MRI) and trans-rectal ultrasound (TRUS). Whereas pathological response (regression grade) stratifies response base on biological effect of radiation on tumours. Mandard tumour regression grade, originally described for oesophageal cancer, is the most commonly used (51). It consists of five different grades based on ratio of fibrosis to tumours. We identified, for the first time, a group of genes that can be used as markers to quantify tumour response following neoadjuvant therapy in rectal cancer patients.

- Biomarkers for diagnosis and screening:
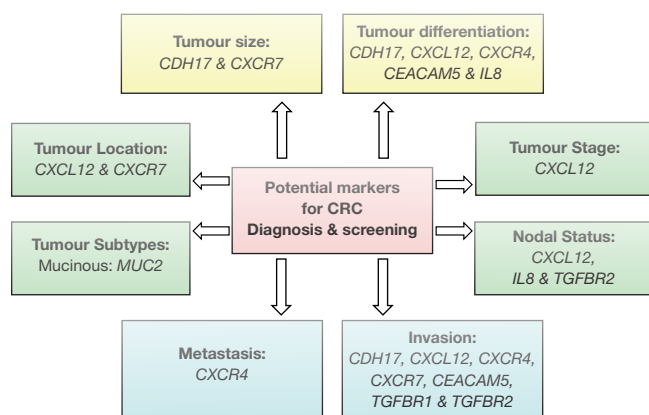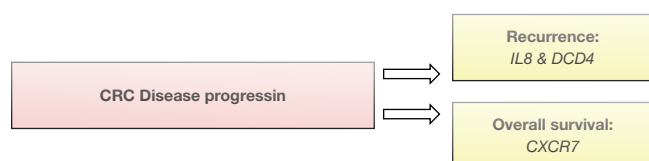

- Biomarkers for disease progression:

Figure 7 Potential biomarker for CRC. Genes identified in the study as potential biomarket for CRC screening, diagnosis and disease progression

Figure 8 Correlation gene expression and CRC management stratigies

Conclusions

The list of the genes identified in this study could serve as molecular markers to complement existing histopathological factors in screening, diagnosis, follow up and therapeutic strategies for individualised care of patients (*Figures 7,8*).

Acknowledgements

References

1. Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.
2. Comber H, Walsh PM. Patterns of care and survival of cancer patients in Ireland 1994 to 2004. Summary report. National Cancer Registry. Cork, Ireland,2008.
3. Forrester K, Almoguera C, Han K, et al. Detection of high incidence of K-ras oncogenes during human colon tumorigenesis. Nature 1987;327:298-303.
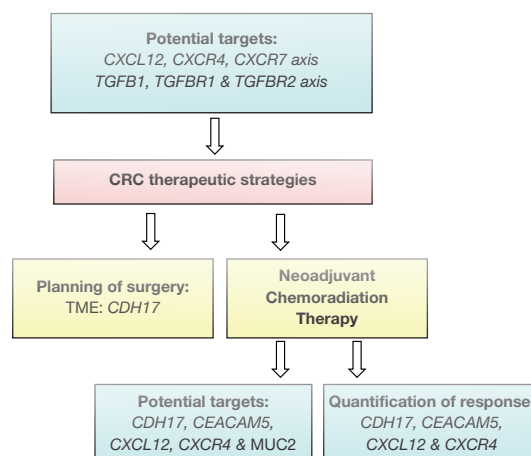4. Cui H, Horon IL, Ohlsson R, et al. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. Nat Med 1998;4:1276-80.
5. Chan SK, Griffith OL, Tai IT, et al. Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. Cancer Epidemiol Biomarkers Prev 2008;17:543-52.
6. Nannini M, Pantaleo MA, Maleddu A, et al. Gene expression profiling in colorectal cancer using microarray technologies: results and perspectives. Cancer Treat Rev 2009;35:201-9.
7. Birkenkamp-Demtroder K, Olesen SH, Sørensen FB, et al. Differential gene expression in colon cancer of the caecum versus the sigmoid and rectosigmoid. Gut 2005;54:374-84.
8. Varghese S, Burness M, Xu H, et al. Site-specific gene expression profiles and novel molecular prognostic factors in patients with lower gastrointestinal adenocarcinoma diffusely metastatic to liver or peritoneum. Ann Surg Oncol 2007;14:3460-71.
9. Wang Y, Jatkoe T, Zhang Y, et al. Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer. J Clin Oncol 2004;22:1564-71.
10. Xu D, Li XF, Zheng S, et al. Quantitative real-time RT-PCR detection for CEA, CK20 and CK19 mRNA in peripheral blood of colorectal cancer patients. J Zhejiang Univ Sci B 2006;7:445-51.
11. Iinuma H, Okinaga K, Egami H, et al. Usefulness and clinical significance of quantitative real-time RT-PCR to detect isolated tumor cells in the peripheral blood and tumor drainage blood of patients with colorectal cancer. Int J Oncol 2006;28:297-306.
12. Ki DH, Jeung HC, Park CH, et al. Whole genome analysis for liver metastasis gene signatures in colorectal cancer. Int J Cancer 2007;121:2005-12.
13. Kollmar O, Rupertus K, Scheuer C, et al. CXCR4 and CXCR7 regulate angiogenesis and CT26.WT

tumor growth independent from SDF-1. Int J Cancer 2010;126:1302-15.

14. Ingold B, Schulz S, Budczies J, et al. The role of vascular CXCR4 expression in colorectal carcinoma. Histopathology 2009;55:576-86.

15. Meijer J, Ogink J, Roos E. Effect of the chemokine receptor CXCR7 on proliferation of carcinoma cells in vitro and in vivo. Br J Cancer 2008;99:1493-501.

16. Birkenkamp-Demtroder K, Christensen LL, Olesen SH, et al. Gene expression in colorectal cancer. Cancer Res 2002;62:4352-63.

17. Grade M, Hörmann P, Becker S, et al. Gene expression profiling reveals a massive, aneuploidy-dependent transcriptional deregulation and distinct differences between lymph node-negative and lymph node-positive colon carcinomas. Cancer Res 2007;67:41-56.

18. Pei H, Zhu H, Zeng S, et al. Proteome analysis and tissue microarray for profiling protein markers associated with lymph node metastasis in colorectal cancer. J Proteome Res 2007;6:2495-501.

19. Kitahara O, Furukawa Y, Tanaka T, et al. Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res 2001;61:3544-9.

20. Barrier A, Lemoine A, Boelle PY, et al. Colon cancer prognosis prediction by gene expression profiling. Oncogene 2005;24:6155-64.

21. Colliver DW, Crawford NP, Eichenberger MR, et al. Molecular profiling of ulcerative colitis-associated neoplastic progression. Exp Mol Pathol 2006;80:1-10.

22. Friederichs J, Rosenberg R, Mages J, et al. Gene expression profiles of different clinical stages of colorectal carcinoma: toward a molecular genetic understanding of tumor progression. Int J Colorectal Dis 2005;20:391-402.

23. Lin YM, Furukawa Y, Tsunoda T, et al. Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. Oncogene 2002;21:4120-8.

24. Bianchini M, Levy E, Zucchini C, et al. Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa. Int J Oncol 2006;29:83-94.

25. Kim IJ, Lim SB, Kang HC, et al. Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer. Dis Colon Rectum 2007;50:1342-53.

26. Xu WQ, Jiang XC, Zheng L, et al. Expression of TGF-beta1, TbetaRII and Smad4 in colorectal carcinoma. Exp Mol Pathol 2007;82:284-91.

27. Xu Y, Pasche B. TGF-beta signaling alterations and

susceptibility to colorectal cancer. Hum Mol Genet 2007;16:R14-20.

28. Saltzman BS, Yamamoto JF, Decker R, et al. Association of genetic variation in the transforming growth factor beta-1 gene with serum levels and risk of colorectal neoplasia. Cancer Res 2008;68:1236-44.

29. Skoglund Lundin J, Vandrovcova J, Song B, et al. TGFBR1 variants TGFBR1(*)6A and Int7G24A are not associated with an increased familial colorectal cancer risk. Br J Cancer 2009;100:1674-9.

30. Zeng Q, Phukan S, Xu Y, et al. Tgfbr1 haploinsufficiency is a potent modifier of colorectal cancer development. Cancer Res 2009;69:678-86.

31. Biswas S, Trobridge P, Romero-Gallo J, et al. Mutational inactivation of TGFBR2 in microsatellite unstable colon cancer arises from the cooperation of genomic instability and the clonal outgrowth of transforming growth factor beta resistant cells. Genes Chromosomes Cancer 2008;47:95-106.

32. Davoren PA, McNeill RE, Lowery AJ, et al. Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer. BMC Mol Biol 2008;9:76.

33. Kheirelseid EA, Chang KH, Newell J, et al. Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer. BMC Mol Biol 2010;11:12.

34. Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.

35. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. CA Cancer J Clin 2010;60:277-300.

36. Cancer Projections 2005-2035. National Cancer Registry. 2008: Cork, Ireland.

37. Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993;363:558-61.

38. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-67.

39. Sierro F, Biben C, Martínez-Muñoz L, et al. Disrupted cardiac development but normal hematopoiesis in mice deficient in the second CXCL12/SDF-1 receptor, CXCR7. Proc Natl Acad Sci U S A 2007;104:14759-64.

40. Daniel P, Wagrowska-Danilewicz M, Danilewicz M, et al. Transforming growth factor beta 1 and metalloproteinase-9 overexpression in colorectal cancer (CC) and adenoma. Int J Colorectal Dis 2007;22:1165-72.

41. Matsushita M, Matsuzaki K, Date M, et al. Down-regulation of TGF-beta receptors in human colorectal cancer: implications for cancer development. Br J Cancer 1999;80:194-205.

42. Bufill JA. Colorectal cancer: evidence for distinct genetic categories based on proximal or distal tumor location. Ann Intern Med 1990;113:779-88.

43. Hinoi T, Lucas PC, Kuick R, et al. CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. Gastroenterology 2002;123:1565-77.

44. Takamura M, Ichida T, Matsuda Y, et al. Reduced expression of liver-intestine cadherin is associated with progression and lymph node metastasis of human colorectal carcinoma. Cancer Lett 2004;212:253-9.

45. Kwak JM, Min BW, Lee JH, et al. The prognostic significance of E-cadherin and liver intestine-cadherin expression in colorectal cancer. Dis Colon Rectum 2007;50:1873-80.

46. Lee S, Bang S, Song K, et al. Differential expression in normal-adenoma-carcinoma sequence suggests complex molecular carcinogenesis in colon. Oncol Rep 2006;16:747-54.

47. Lawrie LC, Dundas SR, Curran S, et al. Liver fatty acid binding protein expression in colorectal neoplasia. Br J Cancer 2004;90:1955-60.

48. Yamazaki T, Kanda T, Sakai Y, et al. Liver fatty acid-binding protein is a new prognostic factor for hepatic resection of colorectal cancer metastases. J Surg Oncol 1999;72:83-7.

49. Mudduluru G, Medved F, Grobholz R, et al. Loss of programmed cell death 4 expression marks adenoma-carcinoma transition, correlates inversely with phosphorylated protein kinase B, and is an independent prognostic factor in resected colorectal cancer. Cancer 2007;110:1697-707.

50. Göke R, Barth P, Schmidt A, et al. Programmed cell death protein 4 suppresses CDK1/cdc2 via induction of p21(Waf1/Cip1). Am J Physiol Cell Physiol 2004;287:C1541-6.

51. Silliman RA, Guadagnoli E, Rakowski W, et al. Adjuvant tamoxifen prescription in women 65 years and older with primary breast cancer. J Clin Oncol 2002;20:2680-8.