Molecular classification of colorectal cancer

[1,2]**Romina Briffa** B. Pharm. (Hons.) (Melit.), M.Sc. Pharm. (Melit.), Ph.D. (Edin.)
[1]**Godfrey Grech** B.Sc. (Melit.), M.Phil. (Melit.), Ph.D. (Erasmus)

[1]Department of Pathology, Faculty of Medicine and Surgery, University of Malta
[2]School of Medicine, University of St Andrews, The Gateway, North Haugh, St Andrews Fife KY16 9RJ Scotland, UK
Email: rbrif01@um.edu.mt

[1]Department of Pathology, Faculty of Medicine and Surgery, University of Malta
Email: godfrey.grech@um.edu.mt

Educational aims

- To describe the current molecular markers used in colorectal cancer
- To give an overview of the various colorectal cancer molecular classifications published in the past five years
- To describe the first consortium colorectal cancer taxonomy

Key words

colorectal cancer, molecular classification, gene expression, biomarkers, colorectal cancer taxonomy

Abstract

Colorectal cancer (CRC) is a heterogeneous disease with several clinical, pathological, and molecular presentations. A comprehensive and unifying molecular classification would be useful for genotype-phenotype correlations, to better understand disease progression, and to predict responses to treatment. Such a classification would be helpful for quickly and efficiently translating results from the laboratory to the clinic and closing the gap between research breakthroughs and actually implementing them clinically. In November 2015, an international consortium consisting of six expert groups published the first consensus on molecular subtypes of colorectal cancer, by bringing together six previously published CRC classifications.

Introduction

Cancers have traditionally been classified clinically and pathologically based on stage and grade. Stage is closely associated with patient prognosis, generally defined as progression-free survival (PFS) and overall survival (OS). Although prognosis has been shown to be dependent on local tumour involvement, regional lymph node metastasis, lympho-vascular invasion, positive surgical margins, preoperative elevation of CEA (a circulating tumour marker), high tumour grade, and tumour budding, responses to treatment are still difficult to predict in specific scenarios, particularly the metastatic setting.[1]

The Union of International Cancer Control (UICC) and American Joint Committee on Cancer (AJCC) Tumour Node Metastasis (TNM) systems are the most commonly used classifications for colorectal cancer staging. The TNM classification was developed in the 1940s and codes the depth of primary tumour (T) invasion into or beyond the colorectum (invasion of or adherence to adjacent organs or structures), the number of regional lymph nodes involved (N), and the presence or absence of distant metastasis (M).[2] The classification allows grouping of these three indicators and provides a "stage-grouping".[3] The TNM classification is a dual system and comprises a clinical (pre-treatment) classification (usually referred to as cTNM) and a pathological (post-surgical histopathological) classification (pTNM).[4] As a rule, surgical decisions are based on cTNM classification, while post-surgical management and prognosis are established using the pTNM classification.

In 1932, the Dukes' staging system was proposed for rectal tumours that classified tumours pathologically into three different stages: A to C.[5] Dukes established that the extent of local tumour invasion (A – least, C-most) was prognostic and reliably predicted operative mortality. This classification was then adapted to colon cancer and further variations were introduced, namely the introduction of Stage D (distant metastasis) and the Modified Astler-Coller (MAC) classification.[5] Today, the Dukes' staging system and the MAC classification have been superseded by the UICC and AJCC TNM systems and subsequently their use is highly discouraged.

Molecular markers currently used in colorectal cancer

With further advances in biomedicine, it has become possible to classify tumours based on their molecular characteristics. Three main cancer pathways are implicated in CRC molecular classification: microsatellite instability (MSI), chromosomal instability (CIN), and the CpG island methylation phenotype (CIMP). MSI is characterised by small insertions or deletions in repetitive DNA sequences (microsatellites).[6] When MSI is not present, a tumour is defined as microsatellite stable (MSS). CIN is defined as an elevated frequency of whole-chromosome missegregation, and CIMP occurs when there is methylation of the CpG islands within the promoter region and is associated with transcriptional silencing.[7,8] In 2008, Ogino and Goel proposed a revised molecular classification based on the five subgroups proposed by Jass in 2007 that consisted of six subgroups: Group 1: MSI-H CIMP-High (10%); Group 2: MSI-H CIMP – Low/0 (5%); Group 3: MSI-L/MSS CIMP-High (5-10%); Group 4: MSI-L CIMP-Low (5%); Group 5: MSS CIMP-Low (30 to 35%); Group 6: MSI-L/MSS CIMP-0 (~40%).[6,9] This molecular classification depends on common DNA markers together with *BRAF* and *KRAS* mutational status.[10]

With the advent of high-throughput molecular technology, we are now generating a vast amount of genomic, transcriptomic, proteomic, and metabolomic data that are facilitating the comprehensive molecular characterisation of cancer. The aim is to identify a robust molecular signature that can be applied in the clinic and help to identify the best treatment and care for the individual. Furthermore, it would be helpful to establish molecular signatures capable of predicting individuals at higher risk of developing the disease who hence must take preventative measures (diagnostic biomarkers). Molecular signatures would also be beneficial in guiding clinicians with respect to predicting recurrence of the disease following surgery (surveillance biomarkers). Other beneficial molecular signatures are prognostic biomarkers, which give an indication of the likely progression of disease, and predictive biomarkers, which predict responses to treatment. This has been best demonstrated in breast cancer, in which the molecular classification has been useful for both diagnosis and treatment.[11-14]

However, the molecular characterisation of colorectal cancer has lagged behind breast cancer in which, in some instances, diagnosis and treatment are based on the expression of particular molecules, e.g. expression of the estrogen receptor (ER) to indicate anti-estrogen therapy or amplification of the human epidermal growth factor receptor 2 (HER2) for HER2-directed therapy.[15] This progress in breast cancer has in no small part been driven by its molecular classification. Nonetheless, there are still clinical dilemmas in CRC that would benefit from a molecular classification approach: a) who might develop metastases even though they have a favourable pathological stage? b) who needs adjuvant therapy even if there is no nodal involvement? c) who is likely to best respond to specific chemotherapy or targeted therapy? To this end, there have now been a number of efforts to develop a better molecular classification of colorectal cancer. Over the past decade, a number of complex classification studies producing variable results have been published. Although the different cohorts, methods, group sizes, and clinical information have meant that these classification systems have seemed different, in fact there is emerging commonality between them. In this brief review, we highlight six of the most recent classifications that together form the basis of the recently published consensus on molecular classification of CRC developed by the CRC Subtyping Consortium (CRCSC).

Classification based on mesenchymal or epithelial expression signature

In 2012, Schlicker *et al.* stratified and pharmacologically characterised a panel of 74 different CRC cell lines.[16] Furthermore, using iterative clustering, this analysis revealed five CRC subtypes (1.1, 1.2, 1.3, 2.1, 2.2) that were successfully validated on over 1600 CRC tumour samples from publicly available gene expression datasets.

Type 1 tumours exhibited a mesenchymal expression signature, had a poor prognosis, and similar number of MSI and MSS tumours (out of the annotated cohorts (n = 229), 58 Type 1 tumours exhibited MSI and 59 exhibited MSS). Type 2 tumours exhibited an epithelial expression signature, had a good prognosis, and were enriched with MSS tumours. Further subtyping of Type 1 tumours revealed that subtype 1.1 was strongly mesenchymal, enriched in late-stage CRC, and the up-regulated genes were mainly involved in Ca-signalling and SRF-targeted. Moreover, subtype 1.1 was characterized by pathways involved in angiogenesis, inflammation, and proliferation. Subtype 1.2 contained more female than male patients, were mostly MSI enriched, and activated similar pathways to 1.1 but also strongly activated the JAK-STAT signalling pathway. Furthermore, the main up-regulated genes were immune-system related. Subtype 1.3 exhibited high expression of transporter genes and were mainly MSS.

Additional subtyping of type 2 tumours resulted in two further subtypes – subtype 2.1 and subtype 2.2. In the former, pathways related to inflammation, angiogenesis and proliferation were activated. Moreover, stress response and immune system-related genes were up-regulated. In subtype 2.2, a number of genes that were up-regulated were involved in cell cycle and amino acid synthesis. Some other genes up-regulated in subtype 2.2 were located on a number of cytobands in chromosome 20q and 13q.

The authors also adopted a comprehensive cell line model to investigate the relationship between the CRC subtypes and cell signalling and, additionally, associate the molecular features with drug responses. They showed that CRC cell lines classified as subtype 1.2 were highly sensitive to glycogen synthase kinase, Src, and Wnt signalling inhibitors. On the other hand, CRC cell lines designated as type 2 when compared to those classified as type 1 were considerably more sensitive to aurora kinase inhibitors.

Classification based on CIN subtype, MSI/CIMP positive subtype and *KRAS*/*BRAF* mutation subtype

In 2013, two colorectal classification papers were published in Nature Medicine. In the first article De Sousa E Melo F *et al* described three molecularly distinct CRC subtypes (CCS1, CCS2 and CCS3), where one of the subtypes (CCS3) was characterized for the first time.[17] An integral part of this study was the analysis of the gene expression data of 1,164 CRC patients using an unsupervised classification strategy. All the CRC samples, xenografts, cell lines, and precursor lesions were classified into the three subtypes using a 146-gene classifier.

Colon Cancer Subtype 1 (CCS1) mainly consisted of chromosomal instable (CIN) cancers – the majority of the samples in this group generally had *KRAS* and/or *TP53* mutations. Moreover, the tumours in this group were principally located on the left side of the colon. From a metastasis point of view, the authors concluded that Wnt target genes were highly expressed in CCS1 and that they metastasize less frequently compared to CCS3 tumours.

Colon Cancer Subtype 2 (CCS2) mainly consisted of MSI/CIMP-positive tumours located on the right side of the colon. The authors focused principally on CCS3 since it was the least well characterized subtype and, furthermore, compared it mainly with CCS1 tumours; hence, there is relatively little information on CSS2.

Colon Cancer Subtype 3 (CCS3) was not enriched with either CIN or MSI tumours but contained a relatively large number of patients with *KRAS* or *BRAF* mutations. The tumours were distributed throughout the colon and tended to be poorly differentiated. Based on the microsatellite stable and CIMP+ status, together with relative over-representation of *BRAF* mutants, the authors hypothesised that these tumours can arise from pre-neoplastic lesions associated with the serrated pathway. This was also confirmed using Principal Component Analysis on the patient set, classifier, and independent cohorts. Furthermore, when comparing CCS1 with CCS3 using Gene Set Enrichment Analysis (GSEA), it appeared that in the CCS3 subtype the up-regulated genes were involved in epithelial to mesenchymal transition, matrix remodelling, cell migration, and transforming growth factor β signalling. The authors concluded that the least characterized subtype (CCS3) is highly malignant when compared to CCS1 and CCS2.

Disease-free survival (DFS) was statistically significantly lower in CCS3 compared to CCS1 and CCS2 tumours. Furthermore, over half of the CCS3 patients had a recurrence within two years and, overall, the patients in this subgroup had a poorer DFS. In addition to the proposed classification, the authors examined the responses of the different subtypes to targeted therapy. The authors inferred that CCS3 metastatic colorectal cancer patients were resistant to cetuximab, independent of the *KRAS* mutational status. This difference in response was also observed *in vitro* when comparing CCS1 cell lines versus CCS3 cell lines.[17]

Classification based on intestinal stem cells and their respective differentiated cells

In the second paper, Sadanandam *et al* described six CRC subtypes (stem-like, inflammatory, transit-amplifying cetuximab resistant, transit-amplifying cetuximab sensitive, goblet-like, enterocyte) together with their response to cetuximab, standard of care chemotherapy, and DFS.[18] The CRC subtypes were generated using consensus-based unsupervised clustering of the gene expression profiles of 1,290 CRC samples.

The least differentiated stem-like sub-type phenotype was associated with the base of the colon crypt and had a poor DFS. Patients with this subtype had higher expression of Wnt signalling, stem cell, myoepithelial, and mesenchymal genes while having lower expression of differentiation markers. This group was predicted to benefit from chemotherapy (preferably FOLFIRI – a chemotherapy regimen containing leucovorin calcium, 5-fluorouracil and irinotecan) – both as an adjuvant treatment and also in the metastatic setting.

The second subtype was referred to as inflammatory and was not associated with crypt-top or base phenotypes. The main upregulated genes in this group were interferon-related and chemokines. With regards to treatment, it was suggested that patients in the adjuvant setting should ideally be treated with chemotherapy, more specifically with FOLFIRI. On the other hand, in the metastatic setting, both chemotherapy and cetuximab could be ineffective. This subtype also had an intermediate DFS.

The goblet-like subtype was equated with the top part of the colon crypt and had a good DFS. This subtype exhibited high mRNA expression of goblet-specific MUC2 and TFF3. The patients in this subtype would possibly be unresponsive to treatment in the adjuvant setting and hence it was suggested that watchful surveillance should instead be implemented following surgical resection. On the other hand, the authors advised that, in the metastatic stage, chemotherapy (preferably FOLFIRI) or another therapy (that still to be determined) should be administered.

The enterocyte-like sub-type was characterized by high expression of enterocyte-specific genes. Patients assigned to this group had an intermediate DFS. Based on the gene signature of the enterocyte-like sub-type, it was determined that these tumours identify with the colon-crypt top phenotype. Furthermore, when compared to the stem or progenitor cell phenotype, the enterocyte-like sub-type had a more differentiated phenotype. As a result of these association studies, the authors recommended that patients receiving adjuvant treatment in this sub-group should be treated with chemotherapy (preferably FOLFIRI) or other therapy (excluding cetuximab or c-MET therapy). On the other hand, for metastatic disease, the authors did not recommend treatment with chemotherapy, cetuximab, or a c-MET inhibitor.

The final sub-type described was the transit-amplifying (TA) sub-type which, to a certain extent, was considered a heterogeneous group since both the stem cells and Wnt target genes were irregularly expressed. Furthermore, this subtype also exhibited a mixed phenotype, since 59% of these tumours had a crypt top signature with low expression of Wnt signalling targets. On the other hand, the remainder of the TA subtype were significantly associated with crypt base and over-expressed stem and progenitor markers such as *LGR5* and *ASCL2*. In this subtype, the authors recommended that both adjuvant chemotherapy and chemoradiotherapy should be avoided, since they identified a trend that patients in this subtype had a lower DFS when treated. Following a series of proliferation assays to monitor drug responses, this subtype was further divided into another two sub-subtypes – cetuximab-sensitive transit-amplifying (CS-TA) and cetuximab-resistant transit-amplifying (CR-TA). CS-TA exhibited statistically significant higher expression of *EREG* and *AREG* compared to CR-TA. In contrast, CR-TA demonstrated higher expression of filamin A, which is involved in c-MET regulation. In effect, this subtype (CR-TA) was sensitive to *in vitro* treatment with a c-MET inhibitor. The authors hence concluded that, in the metastatic setting, CS-TA patients should be treated with cetuximab while CR-TA should be treated with a c-MET inhibitor.

Classification based on clinicopathological variables and commonly used DNA markers

In a third paper in 2013, Marisa *et al*. published in PLoS Medicine a gene expression classification of six molecular subtypes based on clinicopathological variables and commonly used DNA markers.[10] The relevance of this classification is that the subtypes were associated with different prognoses. The study was performed using a discovery set of 443 patients and a validation set of 1,029 patients. In this classification, the subtypes were named in accordance with their biological characteristics as described below.

The first subtype C1, termed $CIN_{Immune\text{-}Down}$, shared a lot of similarities with $C5_{CINWntUp}$. Both subtypes fell under the

conventional precursor neoplasia pathway, with very low frequencies of mismatch repair deficient genes (dMMR) and CIMP in contrast to very high frequencies of CIN. Furthermore, in both subtypes, the majority were located in the distal colon. With respect to mutational status, both C1-$CIN_{Immune\text{-}Down}$ and $C5_{CINWntUp}$ exhibited intermediate *KRAS* and *TP53* mutation frequency but very low *BRAF* mutation frequency. When comparing resemblance to supervised gene expression signatures, both subtypes had minimal characteristics with *BRAF* mutant-like supervised signatures and exhibited intermediate likeness to a "normal-like" supervised signature. One of the main differences between the two subtypes was that $C5_{CINWntUp}$ demonstrated an intermediate frequency of tumours with stem-cell phenotype-like gene expression profiles, while C1-$CIN_{Immune\text{-}Down}$ exhibited a very low frequency. The other main difference between these two subtypes was in respect of deregulated signalling pathways. In C1-$CIN_{Immune\text{-}Down}$, most signalling pathways were downregulated, especially cell communication and immune pathways. On the other hand, in $C5_{CINWntUp}$, cell communication, Wnt and metabolism pathways were upregulated.

The other four subtypes, namely, C2 dMMR, C3 *KRAS*m, C4 CSC and C6 CIN_{normL} were linked to the serrated precursor neoplasia pathway. The C2 dMMR subtype consisted mainly of dMMR (68%) and *BRAF* mutant tumours (40%) and were very frequently located in the proximal colon. Furthermore, a very high frequency of CIMP (59%) was recorded in this subtype. *KRAS* mutant tumours were also found at an intermediate frequency. With respect to supervised gene expression, the majority of the C2 dMMR exhibited a *BRAF*m-like signature and a serrated CC-like signature. Moreover, the immune system and cell growth pathways were found to be up-regulated in this subtype, whilst the Wnt pathway was found to be deregulated.

The C3 *KRAS*m subtype was also frequently located in the proximal colon and was mainly enriched for *KRAS* mutant tumours (87%). This subtype displayed an intermediate frequency of CIN+, and 18% of the tumour were CIMP+. Moreover, most signalling pathways in C3 *KRAS*m tumours were down-regulated.

The metastasis-enriched (31%) C4-CSC subtype was the only subtype with a reproducible association between poor prognosis and the supervised stem-cell gene expression signature. This subtype was frequently located in the proximal colon, had intermediate frequency of CIN+, and 34% of the patients were CIMP+. Intermediate frequencies for *KRAS* and *TP53* mutations were reported in this subtype. The majority of the tumours in this subtype (91%) exhibited the stem cell phenotype-like gene expression signature. Besides the stem cell signature, this subtype also displayed the *BRAF*-mutant like gene expression profile and the serrated CC-like signature. The EMT/motility pathways were up-regulated, whilst the cell growth and death pathways were down-regulated.

The final subtype described in this study was C6 CIN_{normL}. Although classified under the serrated precursor neoplasia category, this subtype was mainly CIN+ (86%), CIMP-, *TP53* mutant, and located in the distal colon. In contrast to the other CIN+ subtypes, this subtype exhibited very high frequencies of the normal-like gene expression signature and intermediate frequencies of the serrated CC-like gene expression profile. Moreover, C6 CIN_{normL} showed down-regulation of the proliferation pathways and upregulation of the EMT/motility pathways.

Finally, the authors demonstrated that there was a statistically significant association between subtypes C4-CSC and C6 CIN_{normL} and prognosis. In fact, Stage II/III patients with one of these two subtypes had a worse prognosis, with 5-year relapse-free survival rates of 52% and 61%, respectively.

Classification based on biological motifs, morphology, common clinical variables, and molecular markers

In July, 2013 Budinska and colleagues described an additional five subgroup classification using a discovery set of 1,113 CRC gene expression profiles and a validation set of 720 CRC transcriptomic profiles.[19] The authors used unsupervised clustering of genome-wide transcriptome analysis and moreover, described each subtype with respect to biological motifs, morphology, common clinical variables, and molecular markers. The five major subtypes characterized were: A – surface crypt-like; B – lower crypt-like; C – CIMP-H-like; D – mesenchymal-like and E – mixed. These subtypes do not replace classification using current clinicopathological variables or molecule markers but merely complement them.

The surface crypt-like subtype (A) was significantly enriched in *KRAS*-mutant tumours, having predominantly a papillary or serrated histopathology and a low percentage of β-catenin-positive nuclei at the invasive front. The authors also observed that tumours in this subtype were well differentiated and most comparable to normal colonic epithelium by gene expression profiling. Moreover, gene expression analysis showed up-regulation of genes involved in the top of the colonic crypt, secretory cells, and metallothioneins, whilst genes involved with EMT/stroma, Wnt, putative colon cancer stem cells (CSCs), Chr20q, and proliferation were down-regulated. In this subtype, survival after relapse (SAR) was 28.9 months.

The well-differentiated lower crypt-like subtype (B) displayed complex tubular morphology and had a high percentage of β-catenin-positive nuclei at the invasive front. A higher copy number gain/amplification was reported in Chr20q. Moreover, over-expression of genes involved in the top of the colonic crypt, proliferation (mainly *EREG*), and Wnt pathway was detected in the lower crypt-like subtype. In contrast, gene expression regulating the EMT/stroma, immune system, and secretory cells were down-regulated. From a clinical point of view, the tumours in this subtype were mainly located on the left side and were grade 2. The SAR was the highest when compared to the other subtypes (50.4 months). Furthermore, this subtype was significantly prognostic with respect to RFS, OS, and SAR.

Subtype C – CIMP-H-like was commonly MSI+ and *BRAF* mutant (87%) and histopathologically characterized by solid/trabecular or mucinous growth patterns. Commonly, tumours falling in this subtype were located on the right hand side and were high grade. At the invasive front, subtype C tumours did not show any β-catenin nuclear immunoreactivity, although transcriptomically they had high expression of immunity-associated genes, metallothioneins, and homeobox gene module. However, this subtype had low expression of gut development, top colon crypt, *EREG*, Chr20q genes, and genes differentially expressed in the CRC (GDC) gene module. The CIMP-H-like subtype was associated with a poor OS (SAR of only 6.9 months). The authors speculated that this subtype had a transcriptomic signature of a group of tumours that, once metastasized, would become resistant to chemotherapy.

Histopathological examination of the fourth CRC subtype (mesenchymal) revealed a desmoplastic pattern and immunochemistry (IHC) showed that only a low percentage of β-catenin-positive nuclei were present at the invasive front. Among other upregulated gene expression signatures, this subtype also had a high EMT/stroma gene expression. In contrast, lipid synthesis gene expression and the canonical Wnt signalling target signatures were among those down-regulated in the mesenchymal subtype. This subtype had a high frequency of *BRAF* mutant tumours and was correlated with poor OS, possibly as a result of the high EMT expression and low expression of proliferation-associated genes. Moreover, this subtype was also significantly prognostic with respect to RFS and SAR.

The final subtype described in this study was the mixed subtype (E), which was mainly located on the left side of the colon. Histopathologically, this subtype appeared as complex tubular, and IHC showed that a high percentage of β-catenin-positive nuclei were present at the invasive front. Among other upregulated gene expression signatures, this subtype also had high EMT/stroma gene expression, canonical Wnt signalling pathway target signatures, *EREG* gene module expression, and homeobox gene module expression.

Classification based on three biological hallmarks of cancer

The last classification we review was published in July 2013 by Roepman *et al.*.[20] This classification proposed three different intrinsic subtypes based on the three biological hallmarks of cancer – epithelial to mesenchymal transition, deficiency in mismatch repair genes, and cellular proliferation. This molecular CRC classification was derived by unsupervised clustering of multi-omic data generated from 188 stages I – IV CRC patients (discovery set). The validation set comprised of 320 Stage II and 223 Stage III CRC patients and was associated with prognosis and response to chemotherapy.

Thirty-five percent and 22% of patients were classified as MMR-deficient epithelial A-type in the discovery cohort and validation cohort, respectively. This subtype had a significantly higher percentage of MSI+ tumours (49%) and was also enriched for *BRAF* mutants. The authors speculated that this subtype may be more prevalent in early-stage cancers. Sixty-eight percent of the patients in this cluster exhibited an MSI/dMMR expression profile. Differential gene expression analysis for mesenchymal and epithelial markers revealed that A-type tumours could be deemed epithelial. From a proliferative point of view, this subtype exhibited the highest expression of *MKI67*. When compared to the other two subtypes, the prognosis for A-type-MSI patients was very good (93% with 10-year distant metastasis-free survival). Response to chemotherapy was assessed in 222 Stage III patients, where patients in this group had a better OS from adjuvant chemotherapy with a hazard ratio (HR) of 0.39 (p = 0.18).

The proliferative epithelial-like B subtype consisted of 52% of the discovery cohort patients and 62% of the validation cohort patients. This subtype encompassed mainly *BRAF* wild-type tumours (98%) and was almost entirely MSS and pMMR phenotype (99%). In this group, *CDH2*, *FGFR1*, and *TGFB1* (mesenchymal markers) were down-regulated, while four epithelial markers were up-regulated. This subtype was characterized as highly proliferative as a result of up-regulation of *MKI67* and even stronger with respect to *AURKA*. These patients had a relatively poor baseline prognosis but patients in this category would be expected to benefit from adjuvant chemotherapy with 5-fluorouracil (5-FU) (HR = 0.42, p = 0.014).

The mesenchymal-like C subtype was the smallest subtype with 13% of the discovery cohort and 17% of the validation cohort. A small proportion of this subtype was *BRAF* mutant (16%) and 36% were dMMR. In this group, all the mesenchymal markers excluding *FLT1* were considerably up-regulated compared to *CDH*, *EGFR*, and *MET* (epithelial markers) that were down-regulated. Moreover, low proliferative activity was recorded in this group. These patients had a poor baseline prognosis and in addition, were resistant to chemotherapy with 5-FU. In the C-type-MSI patients, only 50% had 10-year distant metastasis-free survival.

The consensus molecular subtypes of colorectal cancer

At the end of 2015, an international consortium published a consensus classification for CRC.[21] In this major collaborative study, the authors evaluated the results of the above CRC classifications and, by utilizing a network-based approach, four robust consensus molecular subtypes were identified. The eighteen datasets utilized in the above characterization and subtyping studies amounted to a cohort of 4,151 patients. The authors divided the cohort into two equivalent groups for training and validation.

The four consensus subtyping clusters comprised 3,104 samples. Additionally, 858 samples did not fall into any of the subtypes and hence were described as unlabelled non-consensus samples. The clusters were first biologically characterized and associated with both clinical variables and prognostic values. The new taxonomy for CRC is described below.

Consensus Molecular Subtype 1 (CMS1): This subtype was referred to as microsatellite instability immune and comprised 14% of the cohort. CMS1 had an MSI+, CIMP-H, and hypermutated phenotype and was highly enriched for *BRAF* mutants. The MSI immune subtype was characterized by increased expression of genes associated with immune infiltrates together with strong activation of the immune evasion pathways. Clinically, these tumours were frequently diagnosed in females, located on the right side, and presented at later stages. Furthermore, when compared to the other subtypes, CMS1 had a worse survival after relapse.

Consensus Molecular Subtype 2 (CMS2): Thirty-seven per cent of the cohort fell in this canonical-epithelial subtype. CMS2 had the highest copy number gains and copy number losses when compared to the other subtypes. Additionally, this subtype had high expression of WNT and MYC downstream targets. Clinically these tumours commonly presented on the left side. In contrast to CMS1, patients in this subtype had better survival rates after relapse.

Consensus Molecular Subtype 3 (CMS3): Thirteen per cent of the cohort fell into this subtype. The authors referred to this as the metabolic subtype on account of the metabolic dysregulation that characterized the group. The majority of the tumours in this subtype were reported to be of mixed MSI status, mainly CIMP-L, and had a low frequency of copy number alterations. Moreover, *KRAS* mutations were common to this particular CMS.

Consensus Molecular Subtype 4 (CMS4): This subtype was characterised as having prominent transforming growth factor-β activation, stromal invasion, and angiogenesis. As a consequence, this subgroup was also referred to as the mesenchymal subtype and comprised 23%

Key points

- Colorectal cancer is a heterogeneous disease with several clinical, pathological, and molecular presentations
- Traditionally, cancers have been classified clinically and pathologically based on stage and grade
- To date there are still clinical dilemmas in colorectal cancer that would benefit from a molecular classification approach
- The integration of genomic, epigenomic, transcriptomic and proteomic data using applied bioinformatics and computational biology is one way of understanding better colorectal carcinogenesis and response to treatment
- Molecular classification is crucial for the discovery of actionable diagnostic, surveillance, prognostic and predictive biomarkers

of the cohort. CMS4 had a high frequency of copy number gains and losses. Clinically, tumours in this subtype tended to be diagnosed at a later stage and had the worst relapse free and overall survival.

Samples with mixed features: 13% of the samples did not stratify in any of the four consensus molecular subtypes. The authors have speculated that the samples in this group might be representative of either a transition phenotype or intratumoral heterogeneity.

Conclusion

Molecular classification of colorectal cancer is of high significance since the colorectal subtypes are useful for generating correlations with a number of variables namely clinicopathological parameters, overall survival, progression free survival, and response to treatment and relapse. These correlations can facilitate our understanding of colorectal cancer biology and provide evidence regarding carcinogenesis.[9] Furthermore, these correlations can be used to discover new molecular subtypes. Moreover, association studies can be helpful in highlighting potential confounding factors and preventing incorrect associations.[9] Finally, robust molecular classification is of paramount importance for the discovery and clinical implementation of prognostic and predictive biomarkers - the precursors for successfully implementation of point of care genomics and making personalised healthcare a reality.

Acknowledgments

Romina Briffa is supported by a REACH HIGH Scholars Programme – Post-Doctoral Grants. The grant may be part-financed by the European Union, Operational Programme II — Cohesion Policy 2014-2020 Investing in human capital to create more opportunities and promote the wellbeing of society - European Social Fund.

References

1. Compton CC, Fielding LP, Burgart LJ, Conley B, Cooper HS, Hamilton SR, et al. Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med. 2000 Jul;124(7):979–94.
2. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A, editors. AJCC cancer staging manual. 7th ed. New York, NY: Springer; 2010.
3. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. Ann Surg Oncol. 2010;17(6):1471–4.
4. Labianca R, Beretta GD, Kildani B, Milesi L, Merlin F, Mosconi S, et al. Colon cancer. Crit Rev Oncol Hematol. 2010 May;74(2):106–33.
5. Haq AI, Schneeweiss J, Kalsi V, Arya M. The Dukes staging system: a cornerstone in the clinical management of colorectal cancer. Lancet Oncol. 2009;10(11):1128.
6. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50(1):113–30.
7. Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C. The significance of unstable chromosomes in colorectal cancer. Nat Rev Cancer. 2003;(9):695–701.
8. Lao VV, Grady WM. Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol. 2011;8(12):686–700.
9. Ogino S, Goel A. Molecular Classification and Correlates in Colorectal Cancer. J Mol Diagn JMD. 2008;10(1):13–27.
10. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. PLoS Med. 2013;10(5):1–13.
11. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh I-T, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010;11(1):55–65.
12. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene Expression and Benefit of Chemotherapy in Women With Node-Negative, Estrogen Receptor–Positive Breast Cancer. J Clin Oncol. 2006;24(23):3726–34.
13. Liu Z, Zhang X-S, Zhang S. Breast tumor subgroups reveal diverse clinical prognostic power. Sci Rep. 2014;4:4002.
14. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy. Clin Cancer Res. 2005;11(16):5678–85.
15. Forda SR, Bergstrom R, Chlebus M, Barker R, Andersen PH. Priorities for improving drug research, development and regulation. Nat Rev Drug Discov. 2013;12(4):247–8.
16. Schlicker A, Beran G, Chresta CM, McWalter G, Pritchard A, Weston S, et al. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. BMC Med Genomics. 2012;5(1):1–15.
17. De Sousa E Melo F, Wang X, Jansen M, Fessler E, Trinh A, de Rooij LPMH, et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. Nat Med. 2013;19(5):614–8.
18. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschleger S, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat Med. 2013;19(5):619–25.
19. Budinska E, Popovici V, Tejpar S, D'Ario G, Lapique N, Sikora KO, et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. J Pathol. 2013;231(1):63–76.
20. Roepman P, Schlicker A, Tabernero J, Majewski I, Tian S, Moreno V, et al. Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. Int J Cancer. 2014;134(3):552–62.
21. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350–6.