Integrative analysis of the colorectal cancer proteome: potential clinical impact

Abdo Alnabulsi[1,2] and Graeme I Murray[1]

[1]Pathology, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK. [2]Vertebrate Antibodies, Zoology Building, Tillydrone Avenue, Aberdeen, UK.

Address correspondence to: Professor Graeme I Murray

Email g.i.murray@abdn.ac.uk

Phone: +44(0)1224 553794

Fax: +44(0)1224 663002

Conflict of interest.

Abdo Alnabulsi is a PhD student supported by Vertebrate Antibodies. Graeme Murray is a scientific advisor to Vertebrate Antibodies.

Abstract

Introduction: Colorectal cancer (CRC) is one of the commonest types of cancer that affects a significant proportion of the population and is a major contributor to cancer related mortality. The relatively poor survival rate of CRC could be improved through the identification of clinically useful biomarkers.

Area covered: This review highlights the need for biomarkers and discusses recent proteomics discoveries in the aspects of CRC clinical practice including diagnosis, prognosis, therapy, screening, and molecular pathological epidemiology (MPE). Studies have been evaluated in relation to biomarker target, methodology, sample selection, limitations, and potential impact. Finally, the progress in proteomic approaches is briefly discussed, and the main difficulties facing the translation of proteomics biomarkers into the clinical practice are highlighted.

Expert opinion: The establishment of specific guidelines, best practice recommendations and the improvement in proteomic strategies will significantly improve the prospects for developing clinically useful biomarkers.

Keywords: biomarkers, colorectal cancer, diagnosis, prognosis, proteomics, predictive screening

1. Introduction

1.1 Colorectal cancer background

Colorectal cancer (CRC) is one of the commonest types of cancer and a major cause of cancer related death [1]. The survival rate is still relatively poor particularly for patients presenting with distant metastases [2]. Established primary CRC can generally be diagnosed based on the histopathological characteristics of tissue biopsies obtained at colonoscopy [3]. However, it can be more difficult to diagnose early CRC or CRC that presents as metastatic disease [4, 5]. Screening programmes for CRC using either colonoscopy or faecal occult blood testing have shown that they may reduce the mortality rate from CRC [6, 7]. Nevertheless, current screening methods suffer from several drawbacks including lack of sensitivity and poor participation rates that impede their potential benefits [8].

Prognosis using the current staging system which is based on the histopathological examination of resected CRC does not necessarily reflect the biological heterogeneity of CRC and thus patients with the same tumour stage often have variation in clinical outcome [9]. This staging system is also the main method whereby therapeutic options are determined, yet patients with the same stage often respond differently to the same treatment [10, 11]. Therefore, reliable and easily measurable biomarkers are urgently required to assist clinicians to overcome current difficulties in clinical practice (Figure 1).

1.2. Proteomics and genomics perspective on biomarker discoveries

A biomarker is defined as a “characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [12]. In recent years, there have been a noticeable shift in the biomarker literature towards genomics and transcriptomics technologies, which

undoubtedly have increased our understanding of cancer biology and have led to a variety of biomarker discoveries [13, 14]. However, genotype is not necessarily reflected in phenotype because of the influence of a range of factors including epigenetic changes, alternative splicing, non-coding RNAs (including microRNAs), post translation modifications and protein-protein interactions [15, 16]. Moreover, the analysis of large genomic data sets is a challenging task that demands complex and sophisticated analytical tools and software [17, 18]. This can make the interpretation of data generated using different samples, array types, sequencing platforms difficult [19, 20].

However, proteomics can assess proteins which directly reflect a pathogenic phenotype and potentially is more likely to provide accurate information on disease state and clinical outcome [15, 21, 22]. Current proteomics technologies are able to assess protein modifications such as post-translation modifications and sequence variants [23]. Advances in proteomic technologies have enabled more accurate and in-depth identification of individual proteins within complex protein mixtures [24, 25]. Furthermore, improvements in protein extraction and separation have made proteomics analysis suitable for use on formalin fixed paraffin embedded (FFPE) tissues, thus potentially exploiting a larger number of archival samples necessary for protein biomarker validation [26]. However, current proteomics technologies are still potentially lacking in terms of their ability to detect very low abundance proteins [27].

2. Biomarkers for CRC diagnosis

2.1. Diagnosis of malignant polyps

Currently, the risk of malignant transformation of polyps is determined through the pathological analysis of polyp characteristics such as increasing size, degree of epithelial cell

dysplasia and greater “villousness” [28]. Patient management guidelines (e.g. time interval between surveillance colonoscopies) are based on the identification of these histopathological characteristics [29]. Since only less than 5% of polyps proceed to malignancy, it is imperative to identify novel protein biomarkers to further assist in identifying high-risk adenomas and therefore allow rational use of endoscopic resources [5]. Proteomic studies have identified several candidates proteins associated with the malignant transformation of adenomatous polyps (Table 1).

Kininogen-1 was identified as a marker for the diagnosis and/or screening of the malignant transformation of adenomas [30]. This study analysed the serum samples of 110 participants using matrix assisted laser desorption ionization time-of-flight (MALDI-TOF/TOF-MS). Kinongen-1 levels were significantly lower in normal mucosa compared with CRC and advanced colorectal adenoma, which is consistent with a previous finding [36]. Nevertheless, the exact role of kininogen-1 in CRC pathogenesis remains unclear. Moreover, it is not clear whether the protein levels vary between different types of adenomas. Therefore, there is a need for further experimental validation of the results especially in the presence of contradictory results regarding the levels of kininogen-1 detected in adenoma and carcinoma [37].

In another study, the analysis of plasma samples using MALDI-TOF/TOF-MS has enabled the identification of a peptide signature to monitor and predict the malignant transformation of polyps in familial adenomatous polyposis (FAP) [31]. This study was the first to show that peptide profiling can be used to monitor CRC development in FAP patients. However, a significant concern is the specificity of the peptide signature for CRC because some of the identified peptides are derived from proteins (e.g. complement C3 and C4) involved in inflammation. Moreover, enzyme linked immunosorbent assay (ELISA) validation of complement proteins revealed contradictory results to the proteomic findings.

This highlights the need for further evaluation of the results in an independent and much larger cohort since the ELISA was performed only on a small cohort (10 FAP, 8 adenoma, and 36 CRC). The contradictory ELISA results were attributed to the increased activity of proteases rather than the circulating level of precursor proteins. The inability to use ELISA for assessing and monitoring the peptide signature makes it much more difficult to validate the findings and may compromise the potential clinical utility of this proposed biomarker signature. For the detection of peptides using ELISA, it may be possible to use monoclonal antibodies produced against short synthetic peptides [38].

A diagnostic portfolio of proteins that could be used to distinguish adenomas from CRC or healthy controls was identified using MALDI-TOF/TOF-MS [32]. This research investigated plasma samples from healthy controls, and from patients with colorectal adenomas or invasive disease. Blood-based biomarkers which can accurately diagnose potentially malignant adenomas could have great clinical utility. However, the proteins identified in this study included inflammatory cytokines [39]. Moreover, the validation was only conducted on a small cohort (30 adenomas and 30 carcinomas). Hence, additional studies are required before these proteins can be considered as potential diagnostic markers.

Proteomic technologies can also be utilised to assess urine for the purpose of identifying biomarkers associated with high-risk adenomas [33]. This study has found that urinary levels of prostaglandin metabolites (PGE-M) measured using liquid chromatography–mass spectrometry (LC/MS) are associated with high-risk adenomas. The study cohort contained a relatively large number of controls and adenomas that were from a prospective cohort. There are limitations in the design and the composition of the cohort used in this study. Firstly, the cohort only included females. Secondly, and more importantly, there were no colorectal carcinomas in the samples analysed. Nevertheless, the findings are consistent with other proteomics studies that have measured urinary PGE-M using the same method [40,

41]. The study by Shrubsole and colleagues [40] used 224 cases with at least one advanced adenoma, 152 small tubular adenomas, 300 single small tubular adenoma, and 364 controls. Whereas the study by Johnson *et al.* [41] assessed PGE-M in 58 CRC, 70 polyps, 28 Crohn's disease, and 72 healthy controls. In addition, Nakanishi et al. [42] found that the inhibition of PGE_2 production suppresses intestinal carcinogenesis in an APC-mutant mouse model. A significant concern is that PGE-M are also involved in several inflammatory pathways and also other malignancies, hence there is a need to assess the specificity of urinary PGE-M in CRC versus other cancers and inflammatory diseases (e.g. Crohn's disease, ulcerative colitis) [43]. It is therefore difficult to interpret the association between PGE-M levels and development of CRC without further study of well-defined cohorts.

In another study, nuclear magnetic resonance (NMR) spectrum analysis of urinary samples of 988 high-risk individuals who required colonoscopy identified metabolomics signatures associated with CRC [34]. Using an algorithmic classifier of metabolic signature (4 metabolites and 4 clinical questions), the study showed it was possible to predict individuals who required colonoscopy with an accuracy better than faecal occult blood test. This is potentially a useful addition to the clinical practice because it could ensure patients avoid colonoscopy. Additionally, the learning algorithmic classifier may represent a novel method of transforming complex proteomics data into clinically relevant parameters. However, the study classified both participants with hyperplastic polyps or adenomas into one group requiring colonoscopic follow-up even though the risk of progression to CRC significantly differs between the two groups. The risk of malignant transformation should be the main determinant of who needs colonoscopy. Furthermore, it is difficult to interpret these findings without further study since only two carcinomas were included in the study. The lack of early CRC from the samples may results in a failure to identify important changes in proteins associated with the early stages of malignant transformation.

2.2. Diagnosis of metastatic CRC

The identification of the primary origin of an unknown metastatic tumour is still challenging in spite of the availability of a variety of diagnostic methods including histopathology, molecular analysis, imaging, and endoscopy [4]. The majority of cancers of unknown primary are metastatic adenocarcinomas of which around 7 % are of colo-rectal origin [44]. Failure to identify the primary origin of a tumour is a significant problem since the clinical management of patients particularly the selection of appropriate treatment regimens depends on the identification of the specific cancer type. Histopathological assessment using a combination of the immunohistochemical markers cytokeratin 20, cytokeratin 7 and CDX2 is often used to identify CRC, although the typical cytokeratin 20+/CDX2+/cytokeratin 7-ve phenotype is not expressed by all colorectal carcinomas [45]. Recent studies have shown that the assessment of a combination of Stabilin-2 (STAB2) with cytokeratin 7 and cytokeratin 20 can provide a highly sensitive and specific test for CRC diagnosis [46, 47]. Both studies used immunohistochemistry (IHC) to evaluate STAB2, cytokeratin 20 and cytokeratin 7 expression in large cohorts (n = 840 and n = 2696) which included CRC, benign tumours, normal tissues, and other common malignancies. While few studies have focused on the identification of new markers that help differentiate metastatic CRC from other malignancies there is a clear requirement for such biomarkers.

3. Biomarkers for CRC prognosis

3.1. The need for prognostic markers in the clinical practice

A prognostic biomarker can be defined as a (biological) variable that provides prospective information on patient outcome which is complementary to the data obtained by

the pathologist from histopathology and on which therapeutic decisions can be guided [48]. There is unjustifiable scepticism towards prognostic marker studies as they are generally perceived as unnecessary in CRC since histopathological examination of the resected colorectal cancers provide key prognostic information (tumour stage, lymph node stage, extramural venous invasion). However, the clinical outcome can vary considerably between patients who are diagnosed with the same tumour stage especially for patients with stage II and III CRC [3]. Thus, the identification of protein biomarkers on both biopsies of CRC and surgically resected CRC, which reflect the heterogeneity of CRC, will help in providing accurate prediction of the clinical outcome of patients.

3.2. Recent proteomics discoveries

Large numbers of potential prognostic markers have been identified using proteomic-based approaches (supplementary information Table S1). For example, a combination of Nano LC-MS and gene expression analysis of stage IV CRC patients (n = 46) was used to identify metastasis associated markers [49]. Maspin was found to vary between the two groups of patients divided based on time to recurrence (Table S1). Immunohistochemical analysis of maspin expression in a tissue microarray containing 419 stage II and III CRC samples indicated it was an independent prognostic factor of time to recurrence and disease specific survival in stage III CRC. The finding was validated using three transcriptomics data sets (75 stage II, 78 stage III, and 53 stage IV). Consistent with this study, high maspin expression was linked to increased apoptosis resistance in a colon cancer cell line (HCT-116RC) [50]. The discovery of maspin highlights the fact that some proteins might have a stage-specific function and thus have stage-related expression profiles that could be detected through proteomics. Stage-specific markers carry important prognostic information that

could complement or even replace the current staging system when surgical specimens are not available. Several other studies have demonstrated that it is possible to identify protein markers associated with tumour stage using comparative proteomics (studies 1, 2, and 9 described in Table S1). However, it is worth noting that the size of the samples used in those studies were small and therefore further validation using larger and independent cohorts is required.

Stomatin-like 2 (STOML2) protein was identified using high-performance LC-MS analysis of membrane proteome in 28 pairs of normal and CRC tissues [51]. The subsequent assessment of the protein expression by IHC showed a strong association between STOML2 and disease-specific survival. High expression of STOML2 was associated with decreased CRC-related survival. Furthermore, the plasma levels of STOML2 as measured by ELISA were higher in early stage CRC compared with healthy individuals, which suggest STOML2 could potentially be used as a screening marker. The main drawback of this study is the small numbers of CRC samples included in the proteomics analysis (n = 28), ELISA (n = 70) and IHC (n = 205). Additional validation using a large and independent cohort is still needed as well as further investigation of the role of STOML2 in CRC pathogenesis.

Another interesting study has revealed a prognostic protein signature using tailored computational analysis of proteomics data generated via a combination of LC-MS and targeted LC-MS (SRM) assessment of plasma samples [52]. The protein signature (major histocompatibility complex class I-A, complement factor H, CD44, protein tyrosine phosphatase, receptor type J, haptoglobin, and cadherin 5 type 2) was associated with different prognostic parameters and could stratify patients to distinct prognostic subgroups. The study also used data from three external transcriptomics cohorts for additional validation. The findings of this study are encouraging and may have considerable implication on the management of CRC patients because, unlike the key pathological prognostic factors which

require examination of a resected CRC specimen, this protein signature can be evaluated noninvasively in the plasma. Nevertheless, additional assessment of the findings on a larger cohort is needed because only 202 CRC samples were used. Moreover, there may be a need to evaluate whether this protein signature is detectable by ELISA since this technology is generally easily implemented in laboratories where quality assurances and best practice guidelines are in place. Other studies which have also utilised proteomics analysis of plasma for prognostic marker identification are described in Table S1.

The analysis of preclinical models (cell lines and xenograft tumours) using 2-D difference gel electrophoresis (DIGE) and MALDI-TOF/TOF MS followed by validation with human samples provides another approach to the identification of prognostic markers [53]. Comparative proteome analysis found stathmin 1 (STMN1) levels to be lower in colon cancer cell line (HCT-116) compared to its metastatic derivative E1. Both knockdown and overexpression of STMN1 in HCT-116 and E1 showed it was associated with significant changes in cell migration, invasion, adhesion, and colony formation. This study also performed IHC staining on a tissue microarray containing 324 primary CRC. The expression of STMN1 was higher in CRC compared to the adjacent normal mucosa, and increasing intensity of expression was associated with poorer CRC specific survival. These results are consistent with a recent study which showed that the silencing of STMN1 inhibited metastasis in (E1) and (HCT116) colon cancer cell lines [54]. However, in contrast to the finding of that study, the expression of STMN1 assessed by IHC using 546 CRC cases from two independent cohorts showed that the overexpression of STMN1 was associated with improved survival [55]. Therefore, although STMN1 has shown a promising potential as a prognostic marker for CRC, there is a need for further research of STMN1 in CRC.

Although some studies have used large and well-characterised cohorts for the validation of their proteomics results [56, 57], the majority of studies have used relatively

small cohorts. Another limitation is the insufficient reporting in a variety of aspects including the collection and processing procedure of specimens, inclusion criteria of patients, clinicopathological characteristics of the cohort and clearly defined endpoints. Many biomarker studies still suffer from lack of adherence to the Reporting Recommendations for Tumour Marker (REMARK) guidelines [58]. Compliance with the REMARK guidelines should help to standardise and improve the quality of biomarker studies [59].

4. Biomarkers for predicting the outcome of CRC therapy

4.1. The need for predictive markers in the clinical practice

A predictive biomarker is defined as a variable that indicates the outcome of a specific type of therapy and therefore aids in making treatment decisions [60]. Predictive markers are needed in CRC management because the benefit of neoadjuvant and/or adjuvant therapy is not clear for a significant proportion of patients [10, 11]. The increasing range of therapeutic options have further highlighted the need for predictive biomarkers. One of the few predictive markers to be in current clinical practice is the identification of KRAS mutations as KRAS mutant tumours do not respond to anti-epidermal growth factor receptor drugs [61]. In addition, the assessment of mismatch repair proteins seems to offer valuable information on the potential benefit of fluorouracil based adjuvant therapy and immune checkpoint inhibitors [62, 63].

4.2. Recent proteomics studies

Several research groups have used proteomics analysis to identify putative predictive markers for CRC (Table 2). Using a combination of 2D-DIGE and LC-MS/MS to assess

serum samples from a group of patients with CRC who had received chemotherapy and bevacizumab, sixty-four differentially expressed proteins were identified between responders and non-responders [64]. The study also used ELISA and IHC to validate three proteins (apolipoprotein E, angiotensinogen and vitamin D binding protein) which were significantly associated with overall survival and or progression free survival in metastatic CRC patients treated with chemotherapy and bevacizumab. This could be useful as less than 50% of patients showed a response to this therapy [69]. Nevertheless, the number of CRC samples in the validation cohort was relatively small (ELISA: n = 68 and IHC: n = 95), therefore further validation of the results is still required.

A panel of 32 proteins associated with CRC was identified using isobaric tags for relative and absolute quantitation (iTRAQ –LC-MS) analysis of cancer-associated fibroblasts obtained from colon cancer and normal tissue [65]. This study presented strong and well-designed discovery model whereby proteins derived from colon-associated fibroblast can be assessed for biomarker discoveries. High expression of lysyl oxidase-like 2 (LOXL2) was associated with poor overall survival and high recurrence, and demonstrated predictive value for adjuvant therapy in stage II colon cancer. The results were validated on a number of independent cohorts using different methods (IHC, gene expression profiling and polymerase chain reaction (PCR)). Still, the number of colon cancer cases in the validation cohorts (IHC: n = 121 and PCR: n= 70) was relatively small and hence further validation is required. A previous study found LOX, a family member and paralog of LOXL2, to play an important role in promoting CRC angiogenesis using in *in vitro* (SW480 and SW620 cell lines) and mouse models (LS174T human CRC cell lines grown as subcutaneous tumours in nude mice) [70]. The results of the preclinical models were further validated by IHC (on a CRC tissues microarray (n=515)) which showed the expression of LOX correlated with VEGF expression and blood vessel formation in patients [70].

Dasatinib, an inhibitor of Src tyrosine kinases, is currently being evaluated for use in CRC. However, this therapy is expensive and can cause side effects. Therefore, there is a need for predictive biomarkers that can accurately select patients based on the potential response to this drug. In a recent study, delta-type protein kinase C (PKCδ) was identified as predictive marker for dasatinib in CRC [67]. Shotgun phosphotyrosine proteomics was used to obtain a global view of tyrosine phosphorylation in HCT-116 colon cancer cell lines and HCT-116 xenograft tumour. The results showed that the measurement of PKCδ pY313 as a promising method for assessing the response to dasatinib.

However, there is still an apparent lack of research focused on the identification of predictive markers. This can be attributed in part to the lack of readily available, large, well-characterised cohorts.

5. Biomarkers for CRC screening

5.1. The need for screening markers in clinical practice

The five-year survival rate of CRC patients is significantly worse for those patients diagnosed with metastatic disease compared with early stage disease [1]. Considering CRC is often asymptomatic at early stages of development, sensitive screening methods may reduce CRC associated mortality through early diagnosis when treatment is more effective. Flexible sigmoidoscopy and faecal occult blood test based screening programmes have shown some success in reducing the mortality rate of CRC [6, 7]. However, faecal occult blood tests generally suffer from a lack of sensitivity and a significant false positive rate, while endoscopic examination of the colon is invasive, expensive and involves a degree of risk [8]. Therefore, there is a need for accurate, cost effective, reliable, and non-invasive

biomarkers. Blood-based markers are ideal for screening because samples can be obtained with minimal trauma.

The plasma levels of adiponectin measured using ELISA seem to be inversely associated with CRC risk in men in a large prospective cohort (n=616 CRC and n=1205 controls) [71]. Further validation studies are needed before adiponectin can be used as screening marker. MicroRNAs, cell-free DNA and circulating tumour cells are available in the peripheral blood and they have shown encouraging results as useful screening markers for CRC, however further optimisation is still required [3, 72].

5.2. Recent proteomics studies

Proteomics studies enabled the identification of large number of proteins that may potentially be used as screening biomarkers for CRC (supplementary material Table S2). The use of a multiple autoantibody-based assay as a screening tool for CRC has shown promising results [73, 74]. In a recent study, 64 autoantibodies were assessed using multiplex serology assay, and a panel of four autoantibodies showed combined strong diagnostic ability in detecting early CRC [75]. The study conducted following a robust approach using technology that allows simultaneous assessment of a large number of potential markers in blood samples selected from the target screening population. Nonetheless, the relatively small cohort used in validation (49 CRC, 29 non-advanced adenoma, and 99 advanced adenoma) and the dissimilarities in the clinico-pathological characteristics between cohorts used for training and validation are limitations which could influence the outcome of this study. Moreover, there are concerns regarding the determination of the appropriate cut-off values.

A panel of proteins including APC-binding protein EB1 (MAPRE1) were elevated in CRC compared to healthy controls in a study that used LC-MS analysis of plasma samples

and CRC cell lines [76]. The study used one cohort (90 CRC) for discovery phase and two independent cohorts for ELISA validation. The samples were collected 17 months prior to CRC diagnosis in the discovery cohort, 7 months prior to diagnosis in one validation cohort and from newly diagnosed patients in the second validation cohort. The inclusion of prior-diagnosis samples can help in detecting important protein changes that happen very early in the process of CRC development. However, one validation cohort included only 32 CRC (with only four stage I cases) and the second cohort included 58 CRC without providing further characteristics of the samples. Furthermore, the study used two cohorts that consisted entirely of women while the third cohort was composed of both women and men. Therefore, further assessment of the results with a large cohort in a well-designed study is imperative.

In another study, the diagnostic potential of MAPRE1 in CRC was assessed using a combination of LC-MS and antibody array analysis of plasma samples, followed by IHC validation on fixed tissue samples [77]. The level of MARPE1 was higher in adenoma and CRC compared to normal healthy samples, which is consistent with the previous study [76]. In this study, the combination of MAPRE1 with carcinoembryonic antigen and adenylate kinase 1 has shown promising results in differentiating adenoma and early CRC, respectively, from healthy controls. Nevertheless, the relatively small number of samples (antibody array: 60 adenomas and 60 CRC/ IHC: 10 adenomas and 66 CRC) makes it difficult to determine the clinical usefulness of this marker combination. MAPRE1 knockdown in APC mutant (HT-29) and APC wildtype (HCT-116) showed an anti-proliferative effect which maybe dependent on APC status [78]. This indicates the importance of further research of the role of MAPRE1 in CRC and the need for further validation.

A diagnostic protein signature (ceruloplasmin, serum paraoxonase/arylesterase 1, serpin peptidase inhibitor, clade A, leucine-rich alpha-2-glycoprotein, and tissue inhibitor of metalloproteinases 1) was identified using proteomics and computational analysis [79]. The

study used robust methodology utilising LC-MS for discovery and targeted LC-MS for validation on a cohort that reflected different stages of CRC development and included other relevant diseases. Comprehensive analysis of the data revealed a five-protein panel with 72% diagnostic accuracy. The protein signature could be utilised in clinical practice as an independent screening test or in combination with existing diagnostic tests. However, there is a need for further validation using larger cohorts and it may be necessary to investigate the finding using alternative technique such as ELISA.

Proteomics analysis of blood proteome is not the only approach exploited to identify screening markers. Other studies have shown promising results using proteomics to assess urine, faecal, tissues and cell lines (Table S2). Follow up validation is required in larger cohorts for all these biomarkers.

6. The role of proteomics in molecular pathological epidemiology (MPE)

The rise in the incidence of CRC in developed countries have been linked to classic epidemiological factors such as diet, physical exercise, smoking, alcohol intake and an ageing population and these factors can be associated with specific molecular abnormalities [80]. Such epidemiological factors can be integrated by MPE to provide a more comprehensive understanding of CRC [81]. The phenotype of a disease can be better defined within the paradigm of MPE, which interpret specific molecular signatures within the context of recognised aetiological factors [82, 83]. For example, both CpG island methylator phenotype and microsatellite status in CRC can be linked to a variety of aetiological factors [84, 85]. Therefore, proteomics can provide another dimension for MPE. The challenges, opportunities, and recommendations of this multidisciplinary approach have been recently discussed at the second international MPE meeting [86].

7. Progress and difficulties

Potentially significant limitations observed in many of proteomics studies are the sample size, annotation and the composition of the cohorts. For example, there is a lack of early stage CRC in samples used for investigating potential screening markers. The presence of advanced CRC instead may distort the findings because changes in the profile of plasma proteome are likely to be greater than in early invasive lesions. Moreover, there is a lack of detailed clinicopathological characteristics of many cohorts (e.g., tumour differentiation, presence of extramural venous invasion, stage).

Another potential problem is the low levels of individual protein markers in early tumours which might not be reliably detected in serum with current technologies since it can be difficult to detect low abundance proteins especially in complex protein mixtures [87]. New strategies such as enrichment technologies (e.g. enrich for N- or C-terminal peptides), labelling approaches (e.g. neutron encoding (NeuCode)) may help overcome the inconsistency and lack of sensitivity of unlabelled MS, particularly when dealing with post-translation modifications [88]. Targeted MS is also gaining popularity because it is highly specific, accurate, and even applicable when there are problems with the antibodies [89]. However, there are difficulties when targeting several biomarkers simultaneously in multiple samples, and the method requires exhaustive and challenging optimisation process. Some of the challenges might potentially be addressed using a wider MS/MS window termed Sequential Window Acquisition of all THeoretical Mass Spectra (SWATH) strategy [90]. Finally, new approaches that are increasingly used include immunocapture strategies such as reverse phase protein microarray and immunocapture coupled to mass spectrometry [91]. Some of the advantages using such methods include the ability to measure multiple targets with highly sensitivity, requiring only small volume of serum or plasma, and with no need for

albumin depletion. However, the main weakness of this method is the dependency on the quality of the antibodies.

8. Expert commentary

"Precision medicine refers to the tailoring of medical treatment to the individual characteristics of each patient" [92]. The ongoing improvements in proteomic technologies should enable a comprehensive profiling of the proteome to provide a platform from which specific biomarkers necessary for precision medicine can be identified. In colorectal cancer there is still an urgent need for sensitive, reliable, and cost-effective biomarkers to complement the current methods of diagnosis, prognosis, therapy determination, and screening. Proteomics studies have generated a large number of potentially useful biomarkers. However, no protein biomarkers appear to have been successfully translated into clinical practice. This is attributed in part to the lack of reproducibility of results and the limitations of the validation studies. The reproducibility of proteomics studies is compromised by deficiencies in the studies design; small sample size, variations in the sample preparation and storage protocols, and complexities of data analysis and interpretation [93]. The lack of standardisation between different laboratories regarding quality assurance in the analytical techniques makes the results difficult to replicate and interpret [94]. The lack of reproducibility of the results could be minimised if the studies adopted strong experimental design and adhered to best practice guidelines. In addition, the introduction of automated quality control might significantly improve the reproducibility of proteomics results [95].

Lack of follow-up validation studies and deficiencies in validation assays also contribute to the lack of biomarkers success [96]. The assessment of biomarkers studies and their subsequent validation can seriously be hindered by the lack of large collaborative

projects, shortage of well-characterised samples, inconsistency of proteomics results, and insufficient reporting. Consequently, it makes it difficult to interpret, analyse, and validate the findings. The characteristics of samples should be carefully selected, justified, and clearly stated because they effect the results of biomarker studies. Finally, the validation methods (often antibody based) also suffer from absence of standardisation, absence of reliable antibodies, lack of best practice and quality controls [97, 98]. Rigorous and standardised characterisation process is needed to validate the antibodies used for immunoassays [99].

9. Five-year view

Further advancement in proteomics technologies will result in a more accurate assessment of plasma, serum, tissues, urine, saliva, and faeces proteome. Moreover, the introduction of sophisticated computational software should lead to improved and consistent data generation and analysis. There will also be significant improvement in study design, quality of samples, quality of antibodies, and adherence to best practice guidelines. It is also expected that there will be more collaborative projects with pooling of a wide range of expertise and resources. Thus, more biomarker targets will be identified using proteomics but their potential impact on the clinical practice is largely determined by the amount of progress made in addressing current limitations.

10. Key issues

- CRC is a common cancer with significant mortality.
- Biomarkers offer a solution to some of the current problems in CRC clinical management.
- Major advancements in proteomics coupled with innovative computational analysis have allowed the analysis of complex protein samples.
- Large numbers of potential CRC biomarkers have already been identified for use in diagnosis, prognosis, therapy determination, and screening.

- The majority of proteomics studies have focused on prognosis and screening.
- The failure of biomarkers is the consequence of three main factors; lack of validation, limitations in proteomics technologies and methodology deficiencies.
- More biomarkers will be identified considering the ongoing advancement in proteomics strategies and computational analysis.
- Addressing the current limitations will give future biomarkers discoveries a higher chance of success.

References

Papers of special note have been highlighted as:

* of interest

** of considerable interest

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin. 2015;65:5-29.

2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013;63:11-30.

3. Coghlin C, Murray GI. Biomarkers of colorectal cancer: recent advances and future challenges. Proteomics Clin Appl. 2015;9:64-71.

4. Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. Lancet. 2012;379:1428-1435.

5. Williams J, Pullan R, Hill J, et al. Management of the malignant colorectal polyp: ACPGBI position statement. Colorectal Dis. 2013;15:1-38.

6. Elmunzer BJ, Hayward RA, Schoenfeld PS, et al. Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: a systematic review and meta-analysis of randomized controlled trials. PLoS Med. 2012;9:e1001352.

7. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol. 2008;103:1541-1549.

8. Creeden J, Junker F, Vogel-Ziebolz S, et al. Serum tests for colorectal cancer screening. Mol Diagn Ther. 2011;15:129-141.

9. Blanco-Calvo M, Concha A, Figueroa A, et al. Colorectal cancer classification and cell heterogeneity: a systems oncology approach. Int J Mol Sci. 2015;16:13610-13632.

10. Van Gijn W, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. Lancet Oncol. 2011;12:575-582.

11. O'Connor ES, Greenblatt DY, LoConte NK, et al. Adjuvant chemotherapy for stage II colon cancer with poor prognostic features. J Clin Oncol. 2011;29:3381-3388.

12. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69:89-95.

13. Cheasley D, Jorissen RN, Liu S, et al. Genomic approach to translational studies in colorectal cancer. Transl Cancer Res. 2015;4:235-255.

14. Feero WG, Guttmacher AE, McDermott U, et al. Genomics and the continuum of cancer care. N Engl J Med. 2011;364:340-350.

15. Kocevar N, Hudler P, Komel R. The progress of proteomic approaches in searching for cancer biomarkers. N Biotechnol. 2013;30:319-326.

16. Bardhan K, Liu K. Epigenetics and colorectal cancer pathogenesis. Cancers. 2013;5:676-713.

17. Marx V. Biology: The big challenges of big data. Nature. 2013;498:255-260.

18. Fan J, Han F, Liu H. Challenges of big data analysis. Nat Sci Rev. 2014;1:293-314.

19. Ross MG, Russ C, Costello M, et al. Characterizing and measuring bias in sequence data. Genome Biol. 2013;14:R51.

20. Xuan J, Yu Y, Qing T, et al. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett. 2013;340:284-295.

21. Dundas SR, Murray GI. Proteomics of bone and soft tissue sarcomas. Current Proteomics. 2012;9:94-102.

22. Zhou L, Wang K, Li Q, et al. Clinical proteomics-driven precision medicine for targeted cancer therapy: current overview and future perspectives. Expert Rev Proteomics. 2016;13:367-381.

23. Bonaldi T, Noberini R. Recent advances in mass spectrometry analysis of histone post-translational modifications: potential clinical impact of the PAT-H-MS approach. Expert Rev Proteomics. 2016;13:245-250.

24. Brown KJ, Formolo CA, Seol H, et al. Advances in the proteomic investigation of the cell secretome. Expert Rev Proteomics. 2012;9:337-345.

25. Janssen KP, Knez K, Spasic D, et al. Nucleic acids for ultra-sensitive protein detection. Sensors. 2013;13:1353-1384.

26. Vincenti DC, Murray GI. The proteomics of formalin-fixed wax-embedded tissue. Clin Biochem. 2013;46:546-551.

27. Alvarez-Chaver P, Otero-Estevez O, Paez de la Cadena M, et al. Proteomics for discovery of candidate colorectal cancer biomarkers. World J Gastroenterol. 2014;20:3804-3824.

28. Gibson JA, Odze RD. Pathology of premalignant colorectal neoplasia. Dig Endosc. 2016;28:312-323.

29. Cairns SR, Scholefield JH, Steele RJ, et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). Gut. 2010;59:666-689.

30. Wang J, Wang X, Lin S, et al. Identification of kininogen-1 as a serum biomarker for the early detection of advanced colorectal adenoma and colorectal cancer. PloS One. 2013;8:e70519.

31. Agatea L, Crotti S, Ragazzi E, et al. Peptide patterns as discriminating biomarkers in plasma of patients with familial adenomatous polyposis. Clin Colorectal Cancer. 2015; published online 19 December 2015, doi:10.1016/j.clcc.2015.12.002.

32. Choi J, Liu H, Shin DH, et al. Proteomic and cytokine plasma biomarkers for predicting progression from colorectal adenoma to carcinoma in human patients. Proteomics. 2013;13:2361-2374.

33. Bezawada N, Song M, Wu K, et al. Urinary PGE-M levels are associated with risk of colorectal adenomas and chemopreventive response to anti-inflammatory drugs. Cancer Prev Res. 2014;7:758-765.

*Excellent paper which compared large number of samples collected from prospective cohorts. The paper presented promising urinary- based marker and confirmed the importance of prostaglandin pathway in CRC.

34. Eisner R, Greiner R, Tso V, et al. A machine-learned predictor of colonic polyps based on urinary metabolomics. Biomed Res Int. 2013;2013:303982.

35. Besson D, Pavageau AH, Valo I, Bourreau A, et al. A quantitative proteomic approach of the different stages of colorectal cancer establishes OLFM4 as a new nonmetastatic tumour marker. Mole Cell Proteomics. 2011;10:M111.009712.

36. Qiu Y, Patwa TH, Xu L, et al. Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. J Proteome Res. 2008;7:1693-1703.

37. Roeise O, Sivertsen S, Ruud TE, et al. Studies on components of the contact phase system in patients with advanced gastrointestinal cancer. Cancer. 1990;65:1355-1359.

38. Duncan ME, McAleese SM, Booth NA, et al. A simple enzyme-linked immunosorbent assay (ELISA) for the neuron-specific gamma isozyme of human enolase (NSE) using monoclonal antibodies raised against synthetic peptides corresponding to isozyme sequence differences. J Immunol Methods. 1992;151:227-236.

39. Coghlin C, Murray GI. Progress in the identification of plasma biomarkers of colorectal cancer. Proteomics. 2013;13:2227-2228.

40. Shrubsole MJ, Cai Q, Wen W, et al. Urinary prostaglandin E2 metabolite and risk for colorectal adenoma. Cancer Prev Res. 2012;5:336-342.

41. Johnson JC, Schmidt CR, Shrubsole MJ, et al. Urine PGE-M: a metabolite of prostaglandin E2 as a potential biomarker of advanced colorectal neoplasia. Clin Gastroenterol Hepatol. 2006;4:1358-1365.

42. Nakanishi M, Montrose DC, Clark P, et al. Genetic deletion of mPGES-1 suppresses intestinal tumorigenesis. Cancer Res. 2008;68:3251-3259.

43. Colbert Maresso K, Vilar E, Hawk ET. Urinary PGE-M in colorectal cancer: predicting more than risk? Cancer Prev Res. 2014;7:969-972.

44. Kim KW, Krajewski KM, Jagannathan JP, et al. Cancer of unknown primary sites: what radiologists need to know and what oncologists want to know. AJR Am J Roentgenol. 2013;200:484-492.

45. Pavlidis N, Khaled H, Gaafar R. A mini review on cancer of unknown primary site: a clinical puzzle for the oncologists. J Adv Res. 2015;6:375-382.

46. Dragomir A, de Wit M, Johansson C, et al. The role of SATB2 as a diagnostic marker for tumours of colorectal origin: results of a pathology-based clinical prospective study. Am J Clin Pathol. 2014;141:630-638.

47. Magnusson K, de Wit M, Brennan DJ, et al. SATB2 in combination with cytokeratin 20 identifies over 95% of all colorectal carcinomas. Am J Surg Pathol. 2011;35:937-948.

48. McLeod HL, Murray GI. Tumour markers of prognosis in colorectal cancer. Br J Cancer. 1999;79:191-203.

49. Snoeren N, Emmink BL, Koerkamp MJ, et al. Maspin is a marker for early recurrence in primary stage III and IV colorectal cancer. Br J Cancer. 2013;109:1636-1647.

* Excellent paper which used different techniques and followed well-designed approach to identify possible stage specific markers.

50. Payne CM, Holubec H, Crowley-Skillicorn C, et al. Maspin is a deoxycholate-inducible, anti-apoptotic stress-response protein differentially expressed during colon carcinogenesis. Clin Exp Gastroenterol. 2011;4:239-253.

51. Han CL, Chen JS, Chan EC, et al. An informatics-assisted label-free approach for personalized tissue membrane proteomics: case study on colorectal cancer. Mol Cell Proteomics. 2011;10:M110.003087.

* Good paper illustrating interesting method whereby membrane proteome can be analysed for biomarkers discovery.

52. Surinova S, Radova L, Choi M, et al. Non-invasive prognostic protein biomarker signatures associated with colorectal cancer. EMBO Mol Med. 2015;7:1153-1165.

** Excellent paper illustrating the benefits of using targeted proteomic in plasma proteome analysis. Strong study design was used for the identification and validation of proteins signature

53. Tan HT, Wu W, Ng YZ, et al. Proteomic analysis of colorectal cancer metastasis: stathmin-1 revealed as a player in cancer cell migration and prognostic marker. J Proteome Res. 2012;11:1433-1445.

54. Wu W, Tan XF, Tan HT, et al. Unbiased proteomic and transcript analyses reveal that stathmin-1 silencing inhibits colorectal cancer metastasis and sensitizes to 5-fluorouracil treatment. Mol Cancer Res. 2014;12:1717-1728.

55. Ogino S, Nosho K, Baba Y,et al. A cohort study of STMN1 expression in colorectal cancer: body mass index and prognosis. Am J Gastroenterol. 2009;104:2047-2056.

56. O'Dwyer D, Ralton LD, O'Shea A, et al. The proteomics of colorectal cancer: identification of a protein signature associated with prognosis. PloS One. 2011;6:e27718.

57. Mehta RS, Chong DQ, Song M, et al. Association between plasma levels of macrophage inhibitory cytokine-1 before diagnosis of colorectal cancer and mortality. Gastroenterology. 2015;149:614-622.

58. Jankova L, Dent OF, Molloy MP, et al. Reporting in studies of protein biomarkers of prognosis in colorectal cancer in relation to the REMARK guidelines. Proteomics Clin Appl. 2015;9:1078-1086

59. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer. 2005;93:387-391.

60. Febbo PG, Ladanyi M, Aldape KD, et al. NCCN task force report: evaluating the clinical utility of tumour markers in oncology. J Natl Compr Canc Netw. 2011; 9:S1-32.

61. Van Cutsem E, Nordlinger B, Cervantes A, et al. Advanced colorectal cancer: ESMO Clinical Practice Guidelines for treatment. Ann Oncol. 2010;21:v93-v97.

62. Sargent DJ, Marsoni S, Monges G, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. J Clin Oncol. 2010;28:3219-3226.

63. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumours with mismatch-repair deficiency. N Engl J Med. 2015;372:2509-2520.

64. Martin P, Noonan S, Mullen MP, et al. Predicting response to vascular endothelial growth factor inhibitor and chemotherapy in metastatic colorectal cancer. BMC Cancer. 2014;14:887.

65. Torres S, Garcia-Palmero I, Herrera M, et al. LOXL2 is highly expressed in cancer-associated fibroblasts and associates to poor colon cancer survival. Clin Cancer Res. 2015;21:4892-4902.

*The study described the purification and analysis of proteome in colon cancer associated fibroblast using multiple techniques.

66. Croner RS, Sevim M, Metodiev MV, et al. Identification of predictive markers for response to neoadjuvant chemoradiation in rectal carcinomas by proteomic isotope coded protein label (ICPL) analysis. Int J Mol Sci. 2016;17:209.

67. McKinley ET, Liu H, McDonald WH, et al. Global phosphotyrosine proteomics identifies PKCδ as a marker of responsiveness to Src inhibition in colorectal cancer. PloS One. 2013;8:e80207.

68. Katsila T, Juliachs M, Gregori J, et al. Circulating pEGFR is a candidate response biomarker of cetuximab therapy in colorectal cancer. Clin Cancer Res. 2014;20:6346-6356.

69. Van Cutsem E, Köhne C, Hitre E, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer N Engl J Med. 2009;360:1408-1417.

70. Baker AM, Bird D, Welti JC, et al. Lysyl oxidase plays a critical role in endothelial cell stimulation to drive tumour angiogenesis. Cancer Res. 2013;73:583-594.

71. Song M, Zhang X, Wu K, et al. Plasma adiponectin and soluble leptin receptor and risk of colorectal cancer: a prospective study. Cancer Prev Res. 2013;6:875-885.

72. Zoratto F, Rossi L, Verrico M, et al. Focus on genetic and epigenetic events of colorectal cancer pathogenesis: implications for molecular diagnosis. Tumour Biol. 2014;35:6195-6206.

73. Chen H, Werner S, Tao S, et al. Blood autoantibodies against tumour-associated antigens as biomarkers in early detection of colorectal cancer. Cancer Lett. 2014;346:178-187.

74. Villar-Vazquez R, Padilla G, Fernandez-Acenero MJ, et al. Development of a novel multiplex beads-based assay for autoantibody detection for colorectal cancer diagnosis. Proteomics. 2016;16:1280-1290.

** This study developed non-invasive diagnostic assay for different antigens in sera samples. The assay showed strong diagnostic performance and was validated on a large size cohort.

75. Chen H, Werner S, Butt J, et al. Prospective evaluation of 64 serum autoantibodies as biomarkers for early detection of colorectal cancer in a true screening setting. Oncotarget. 2016: published online 19 February 2016, doi:10.18632/oncotarget.7500.

76. Ladd JJ, Busald T, Johnson MM, et al. Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women. Cancer Prev Res. 2012;5:655-664.

77. Taguchi A, Rho JH, Yan Q, et al. MAPRE1 as a plasma biomarker for early-stage colorectal cancer and adenomas. Cancer Prev Res. 2015;8:1112-1119.

78. Stypula-Cyrus Y, Mutyal NN, Cruz MD, et al. End-binding protein 1 (EB1) up-regulation is an early event in colorectal carcinogenesis. FEBS Lett. 2014;588:829-835.

79. Surinova S, Choi M, Tao S, et al. Prediction of colorectal cancer diagnosis based on circulating plasma proteins. EMBO Mol Med. 2015;7:1166-1178.

** Excellent paper which described innovative proteomic approach for biomarker discovery and validation. The study also validated their results by transcriptomics validation on external cohorts.

80. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. Nat Rev Disease Primers. 2015: published online 05 November 2015, doi:10.1038/nrdp.2015.65.

81. Ogino S, Chan AT, Fuchs CS, et al. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut. 2011;60:397-411.

82. Colussi D, Brandi G, Bazzoli F, et al. Molecular pathways involved in colorectal cancer: implications for disease behaviour and prevention. Int J Mol Sci. 2013;14:16365–16385.

83. Ogino S, Nishihara R, VanderWeele TJ, et al. The role of molecular pathological epidemiology in the study of neoplastic and non-neoplastic diseases in the era of precision medicine. Epidemiology. 2016;27:602-611.

84. Ogino S, Lochhead P, Chan AT, et al. Molecular pathological epidemiology of epigenetics: emerging integrative science to analyse environment, host, and disease. Mod Pathol. 2013;26:465-484.

85. Kocarnik JM, Shiovitz S, Phipps AI. Molecular phenotypes of colorectal cancer and potential clinical applications. Gastroenterol Rep. 2015;3:269-276.

86. Ogino S, Campbell PT, Nishihara R, et al. Proceedings of the second international molecular pathological epidemiology (MPE) meeting. Cancer Causes Control. 2015;26:959-972.

87. Hori SS, Gambhir SS. Mathematical model identifies blood biomarker-based early cancer detection strategies and limitations. Sci Transl Med. 2011;3:109ra116.

88. Pagel O, Loroch S, Sickmann A, et al. Current strategies and findings in clinically relevant post-translational modification-specific proteomics. Expert Rev Proteomics. 2015;12:235-253.

89. Hathou Y. Proteomic methods for biomarker discovery and validation. Are we there yet? Expert Rev Proteomics. 2015;12:329-331.

90. Selevsek N, Chang CY, Gillet LC, et al. Reproducible and consistent quantification of the saccharomyces cerevisiae proteome by SWATH-mass spectrometry. Mol Cell Proteomics. 2015;14:739-749.

91. Fredolini C, Byström S, Pin E, et al. Immunocapture strategies in translational proteomics. Expert Rev Proteomics. 2016;13:83-98.

92. National Research Council Committee on a Framework for developing a new taxonomy of disease. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease**.** Washington, DC: National Academies Press (US); 2011.

93. Coghlin C, Murray GI. Progress in the development of protein biomarkers of oesophageal and gastric cancers. Proteomics Clin Appl. 2016;10:532-545.

94. Ioannidis JP, Khoury MJ. Improving validation practices in "omics" research. Science. 2011;334:1230-1232.

95. Bielow C, Mastrobuoni G, Kempa S. Proteomics quality control–a quality control software for MaxQuant results. J Proteome Res. 2016;15:777-787.

96. de Wit M, Fijneman RJ, Verheul HM, et al. Proteomics in colorectal cancer translational research: biomarker discovery for clinical applications. Clin Biochem. 2013;46:466-479.

97. Bradbury A, Pluckthun A. Reproducibility: standardize antibodies used in research. Nature. 2015;518:27-29.

98. O'Hurley G, Sjöstedt E, Rahman A, et al. Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. Mol Oncol. 2014;8:783-798.

99. Schuster C, Malinowsky K, Liebmann S, et al. Antibody validation by combining immunohistochemistry and protein extraction from formalin-fixed paraffin-embedded tissues. Histopathology. 2012;60:E37-E50.

Table 1. Summary of recent diagnostic biomarkers discoveries and their potential applications in CRC.

Target(s)	Proteomics		Validation		Findings	Potential utility	Ref
	Method(s)	Sample type	Method(s)	Sample type			
Kininogen-1	Clinprot-MALDI-TOF/TOF-MS	Sera: 35 healthy, 35 advanced colorectal adenoma (ACA) and 40 preoperative CRC	ELISA IHC	Sera: 85 healthy, 80 ACA, 143 preoperative CRC (AJCC stage: I = 14, II = 63, III = 37 and IV = 29) and 58 postoperative CRC Tissues: 75 normal mucosa, 77 ACA and 248 CRC (Dukes' stage: A = 53, B = 101, C = 63 and D = 31)	The sensitivity, specificity, and accuracy are 70.13%, 65.88% and 67.90%, respectively	Diagnosis and screening	[30]
Peptide signature	MALDI-TOF/TOF-MS	Plasma: 38 healthy,13 FAP, 26 adenoma and 58 sporadic CRC (29 early stages and 29 late stages)	ELISA	Plasma: 22 healthy, 10 FAP, 8 adenoma and 36 CRC	Associated with malignant transformation of adenomas in FAP patients	Diagnosis and screening	[31]
Panel of proteins including; IL8, IP-10 and TNF-alpha	MALDI-TOF/TOF-MS	Plasma: 30 adenomas and 30 carcinomas (AJCC Stage: I = 4, II = 15 and III = 11)	ELISA and multiplex array	Plasma: 30 adenomas and 30 carcinomas (AJCC stage: I = 4, II = 15 and III = 11)	A significant increase in the levels of proteins in carcinoma compared to adenomas	Diagnosis and screening	[32]
PGE_2 metabolites (PGE-M)	LC-MS	Urine: 420 control, 130 low risk adenoma and 290 high risk adenoma	None	NA	PGE-M level is indicator of an increased risk for advanced adenoma and identifies patients who might benefit from NSAID chemoprevention	Diagnosis, screening and predictive	[33]
Metabolomics profiles	NMR spectrum and learning algorithm	Urine: 633 healthy, 110 hyperplastic polyps 243 adenoma and 2 CRC	None	NA	Sensitivity of 64% and a specificity of 65%	Diagnosis	[34]
Olfactomedin 4 (OLFM4)	iTRAQ - MALDI-TOF/TOF-MS	Tissues: 4 adenoma and 24 CRC	IHC	Tissues: 30 adenomas, 12 intra-mucosal carcinoma and 84 CRC (AJCC stage: I = 26, II = 14, III = 25 and IV = 19)	OLFM4 increases in adenomas and in early stage CRC before dropping significantly in stages (III-IV)	Diagnosis and prognosis	[35]

Table 2. Summary of recent predictive biomarkers discoveries and their potential applications in CRC.

Target(s)	Proteomics		Validation		Findings	Potential utility	Ref
	Method(s)	Sample type	Method(s)	Sample type			
AGT, APOE and DBP	2D-DIGE and LC-MS/MS	Sera: 23 CRC responders to treatment (AJCC stage: II = 1 and III = 10) and 12 non-responders (stage IV)	ELISA IHC	Sera: 68 CRC (AJCC stage: I = 3, II = 8, III = 14 and IV = 43) Tissues: 95 CRC (AJCC stage: I = 1, II = 14, III = 34 and IV = 46)	Proteins are associated with survival outcomes in metastatic CRC patients treated with chemotherapy and bevacizumab	Predictive	[64]
LOXL2	iTRAQ – LC-MS	Cell lines Tissues: 12 matched colon cancer (AJCC stage: II = 5 and III = 7)	PCR IHC Transcriptomics	Tissues: 70 colon cancer (AJCC stage: I = 8, II = 26, III = 22 and IV = 14) Tissues: 121 colon cancer (AJCC stage: I = 31, II = 53, III = 9 and IV = 28) Tissues: two external cohorts: 232 and 90 colon cancers	LOXL2 identifies a subgroup of patients (stage II and III CRC) who can benefit from adjuvant chemotherapy LOXL2 has a prognostic value in stage II patients	Predictive and prognostic	[65]
Protein panel (n=14), validated (n=4): HADHA, PLEC1, TAGLN and TKT	Isotope coded protein label	Tissues: 20 rectal carcinoma (AJCC stage: II = 10 and III = 10)	IHC	Tissues: 10 good responders and 10 bad responders to neoadjuvant chemoradiation after surgery	This protein panel predicts the response for neoadjuvant chemoradiation in rectal carcinomas	Predictive	[66]
Delta-type protein kinase C (PKCδ)	LC-MS/MS	Cell line and animal xenograft tumour treated with dasatinib	IHC and western blot	Cell lines and animal xenograft tumour treated with dasatinib	PKCδ pY313 assessment can determine the benefit of dasatinib in a subset of CRC patients	Predictive	[67]
Phosphorylated epidermal growth factor receptor (pEGFR)	LC–MS/MS	3D secretomes of CRC isogenic cells treated with cetuximab Sera: 18 metastatic CRC with the KRAS (exon 2) WT status, treated with cetuximab plus FOLFIRI	ELISA	Plasma: 18 metastatic CRC with the KRAS (exon 2) WT status, treated with cetuximab plus FOLFIRI	pEGFR is associated with CRC cells sensitivity to cetuximab and therefore patients' response to this drug	Predictive	[68]

Figure legend

Figure 1.

An overview of colorectal cancer biomarkers types, methods of assessment and potential utilities of biomarkers in clinical practice.

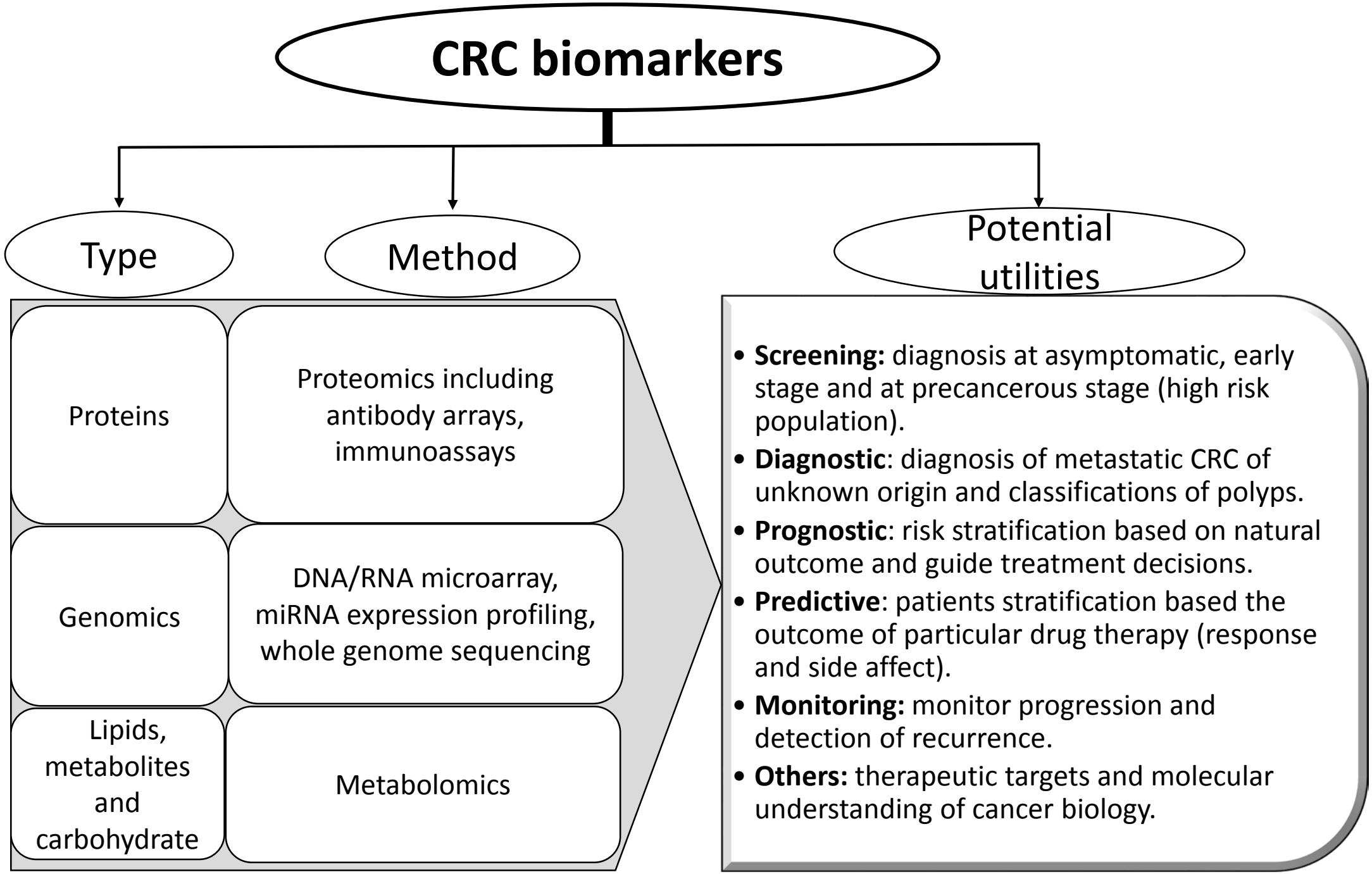
CRC biomarkers
Type
Method
Potential utilities
Proteins
Proteomics including antibody arrays, immunoassays
Genomics
DNA/RNA microarray, miRNA expression profiling, whole genome sequencing
Lipids, metabolites and carbohydrate
Metabolomics
• Screening: diagnosis at asymptomatic, early stage and at precancerous stage (high risk population).
• Diagnostic: diagnosis of metastatic CRC of unknown origin and classifications of polyps.
• Prognostic: risk stratification based on natural outcome and guide treatment decisions.
• Predictive: patients stratification based the outcome of particular drug therapy (response and side affect).
• Monitoring: monitor progression and detection of recurrence.
• Others: therapeutic targets and molecular understanding of cancer biology.

Table S1. Summary of recent prognostic biomarkers discoveries and their potential applications in CRC

Target(s)	Proteomics		Validation		Findings	Potential utility	Ref
	Methods(s)	Sample type	Method(s)	Sample type			
IGF1-R, IRF2BP1 and MX1	LC-MS	Tissues: 19 CRC (10 lymph node metastatic and 9 non-metastatic)	IHC	Tissues: 40 CRC (UICC stage: II = 20 and III = 20)	Expression of these proteins is associated with lymph node metastasis	Prognosis	[1]
ALDOA, CA1, GRP78 and PPIA	MALDI-TOF-MS and 2D-DIGE	Tissues: 5 CRC for each stage	IHC Western blot (WB)	Tissues: 103 CRC (AJCC stage: I = 3, II = 45, III = 30 and IV = 25) Tissues: 1 control and 1 CRC sample for each stage	Dynamic patterns of proteins expression are associated with CRC prognosis especially for stage III and IV	Prognosis	[2]
Mapsin	Nano LC-MS Gene expression	Tissues: 5 stage IV patients (time to recurrence<6 months) and 5 patients longer time to recurrence Tissues:30 stage IV CRC (divided based on time to recurrence)	IHC WB Transcriptomics	Tissues:419 CRC (AJCC stage: II = 243 and III = 176) Tissues; 5 stage IV patients (time to recurrence<6 months) and 5 patients longer time to recurrence External cohorts	Maspin expression is independent predictor of time to recurrence and is associated with diseases specific survival in stage III CRC	Prognosis	[3]
STOML2	LC-MS/MS	Tissues: 28 pair of normal and CRC (Dukes' stage: A = 4, B = 7, C = 11 and D = 6)	ELISA IHC	Tissues:70 CRC and 70 healthy (29 early stage and 41 advance stage) Tissues: 184 adenoma and 205 CRC matched with normal mucosa (AJCC stage: I + II 33 and III + IV = 172)	Overexpression of STOML2 is associated with poor survival. Plasma concentrations of STOML2 were higher in early-stage CRC compared with healthy individuals	Prognosis and screening	[4]
CDH17, DEFA1, EZR, FN1, ITGB2, MLEC and TNC	LC–MS/MS	Tissues: 8 primary CRC (2 for each stage) and their corresponding adjacent normal mucosa	IHC and WB	Tissues: 8 primary CRC (2 for each stage) and their corresponding adjacent normal mucosa	Proteins signature is associated with CRC stage and epidermal growth factor receptor expression	Prognosis	[5]

ALDH1 and 14-3-3β	LC–MS and 2D-DIGE	Tissues: 28 normal and 28 stage B CRC	IHC	Tissues: 515 CRC (Dukes' stage: A = 90, B = 201 and C = 224) and 50 normal mucosa	ALDH1 and 14-3-3β negative tumours have a better prognosis than tumours showing either 14-3-3β or ALDH1 positivity	Prognosis	[6]
RAI3	LC–MS/MS	Tissues: 4 colon cancer and 4 normal	IHC WB	367 colon cancer (Dukes' stage: A = 49, B = 122, C = 144 and D = 52) and 51 normal mucosa 4 colon cancer and 4 normal	High RAI3 expression is associated with colon cancer recurrence in small subgroup of patients	Prognosis	[7]
Cytokeratin 17 and Moesin	2D- DIGE and MALDI-TOF/ TOF-MS	Tissues: 4 CRC (AJCC stage: II = 1 and III = 3) and 4 normal	IHC WB	Tissues: 166 CRC (AJCC stage: I = 33, II = 59, III = 65 and IV = 19) Tissues: 4 CRC (AJCC stage: II = 1 and III = 3) and 4 normal	Moesin and KRT17 were not expressed in normal mucosa and their expression increased as tumour (pT) stage advanced	Prognosis and diagnosis	[8]
FXYD3, GSTM3 and S100A11	MALDI-TOF MS LC-MS	Tissues: 54 colon cancer (UICC stage: II = 21 and III = 33) Tissues: 6 colon cancer (II = 3 and III = 3)	IHC	Tissues: 168 colon cancer (UICC stage: II = 87 and III = 81)	Protein expressions correlate with the presence of nodal metastases	Prognosis	[9]
HSP27	2D-DIGE and LC-MS/MS	Tissues: 9 colon and 3 rectal cancer (stage III)	IHC	Tissues: 199 colon cancer and 205 rectal cancer (AJCC stage I + II = 108 and III + IV = 97) Independent cohort: 200 colon cancer and 115 rectal cancer	HSP27 expression is associated with poor outcome in rectal cancer	Prognosis	[10]
Metabolomic profile	(H NMR) spectrometry and GC-MS	Sera: 42 stages II and III loco-regional CRC, 45 liver-only metastases and 25 extrahepatic metastases	None	NA	Metabolome profile is different in locoregional CRC, in liver-only metastases and in extrahepatic metastases	Prognosis	[11]
CEA, IL-8 and prolactin	Multiplex immunoassay platforms	Plasma: 75 CRC (15 for each Dukes stage A-D) and 15 healthy	None	NA	Protein signature is associated with increased CRC progression and correlates with Dukes' stage	Prognosis	[12]

Gelsolin	Cy-dye labelled proteins (MALDI-TOF MS, LC-MS)	Plasma: 32 CRC, collected before surgery and one closest to distal metastasis diagnosis (AJCC stage at first diagnosis: I = 2, II= 7 and III = 23)	WB and ELISA ELISA IHC	Plasma: the same cohort used in proteomic Plasma: 149 CRC (I + II = 74 and III+IV=75) and 25 normal Tissues: 148 CRC (I = 10, II = 64, III = 48 and IV = 26) and 133 normal mucosa	Plasma levels of secretory gelsolin are higher in patients with distal metastases (stage IV versus stages I–III CRC before treatment)	Prognosis	[13]
CD44, CDH5, CFH, HLA-A, HP and PTPRJ	LC-MS/MS	Plasma: 202 CRC (AJCC stage: I = 43, II = 58, III = 49 and IV = 52)	Targeted LC-MS (SRM) Transcriptomics	Plasma: 202 CRC (AJCC stage: I = 43, II = 58, III = 49 and IV = 52) Three external cohorts	This panel provides a prognostic information on survival and other prognostic parameters	Prognosis	[14]
MIC1 and PTGS2	ELISA IHC		Plasma: 618 CRC (AJCC stage: I + II + III = 533 and IV = 85) Tissues: 245 CRC (stage not stated)		MIC1 level measured prior diagnosis is associated with disease specific mortality, mainly in PTGS2-positive tumours	Prognosis	[15]
STMN1	2-D DIGE, MALDI-TOF/TOF MS	CRC cell lines (HCT-116 and its metastatic derivative E1)	IHC WB	Tissues: 324 CRC (AJCC stage: I = 22, II = 120, III = 97 and IV = 85) Cell lines	Higher expression of STMN1 correlates with poorer prognosis. STMN1 expression is higher in primary and metastatic CRC compared with normal mucosa	Prognosis	[16]
COL6A3	iTRAQ-LC-MS	Cell lines	IHC ELISA	Tissues: 90 matched CRC (AJCC stage: I = 9, II = 47, III = 31 and IV = 2) Plasma: 42 CRC (16 lymph node positive) and 48 normal	Expression of COL6A3 is higher in CRC and it is associated with Dukes stage, T stage and recurrence	Prognosis and screening	[17]

Table S2. Summary of recent screening biomarkers discoveries and their potential applications in CRC

Target(s)	Proteomics		Validation		Findings	Potential utility	Ref
	Method(s)	Sample type	Method(s)	Sample type			
CP, LRG1, PON1, SERPINA3 and TIMP1	LC MS/MS	Plasma: 23 non advanced adenoma, 11 hyperplastic polyp, 66 normal and 97 CRC (AJCC stage: I = 32, II = 26, III = 31 and IV = 8)	Targeted LC-MS (SRM)	Plasma: 4 advanced adenoma, 2 benign adenoma, 1 dysplastic polyp, 6 diverticular disease, 4 Crohn, 50 healthy and 202 CRC (AJCC stage: I = 43, II = 58, III = 49 and IV = 52)	This panel detect CRC at 72% accuracy compared with 49% for CEA	Screening	[18]
Autoantibodies: IMPDH2, MAGEA4, MDM2 and TP53	Multiplex serology, a fluorescent bead-based GST capture immunosorbent assay	Sera: 124 normal and 352 CRC (AJCC stage: I = 96, II = 102, III= 105 and IV=49)	The same method	Sera: 49 CRC (AJCC stage: 0 = 4, I = 18, II = 5, III = 19 and IV = 3); 100 normal, 29 non-advanced adenoma, and 99 advanced adenoma	Sensitivity of autoantibodies is 26% for early stage CRC at a specificity of 90%. Detected 20% of advanced adenomas	Screening	[19]
CEA, IGFBP2, LRG1 and MAPRE1	2D-HPLC and LC-MS/MS	Cell lines and plasma: 18 months pre diagnosis, 90 CRC (AJCC stage: I = 8, II = 29, III = 37 and IV = 16) and 90 controls	ELISA	Plasma: 58 newly diagnosed CRC (stage not provided) and 58 age-matched controls Plasma: 7 months prior diagnosis, 32 CRC (AJCC stage: I = 4, II = 13, III = 12 and IV = 3) and 32 controls	Predictive value in pre-diagnostic CRC plasmas (41% sensitivity at 95% specificity)	Screening	[20]
Anti-p53, CEA, CYFRA 21-1, OPN and seprase	ELISA		Sera: 301 CRC (UICC stage: 0 = 6, I = 53, II = 68, III = 76 and IV = 68), 14 hyperplastic polyps, 143 advanced adenoma, 135 healthy, 176 other cancers (prostate, liver, lung, breast, kidney, bladder, ovary, and endometrium) 258 disease and other controls (Diverticulitis, inflammatory bowel disease, infection-related diarrhoea)		Diagnostic power: 69.6% sensitivity at 95% specificity and 58.7% at 98% specificity	Screening	[21]

Anti-p53, CEA, ferritin, osteopontin and seprase	[21]		Electrochemiluminescence cobas e601 assay	Sera: 1,200 controls, 420 advanced adenoma, 4 carcinoma in situ, and 36 CRC (UICC stage: I = 13, II = 5, III=12 = 6 and IV = 2)	Performance is inferior to FIT, but comparable with the faecal occult blood test (FOBT)	Screening	[22]
Complement component 9 (C9)	2DICAL and LC-MS	Plasma: 59 healthy and 31 CRC (AJCC stage: 0 = 5, I = 10, II = 7, III = 6 and IV = 3)	RPPM	Plasma: 115 CRC (0=17, I = 35, II = 28, III = 25 and IV = 10) and 230 healthy Plasma: 109 healthy, 100 CRC, 105 gastric cancer, 14 hepatocellular carcinoma, 10 oesophageal cancer, 14 pancreatic cancer, 18 cholangiocarcinoma and 8 pancreatitis	C9 was elevated in patients with early stages of CRC	Screening	[23]
C3, C9, GSN, HABP2, ORM1 and SAA2	HPLC and MRM LC/MS	Sera: 259 healthy and 172 CRC (AJCC stage: I = 19, II = 53, III = 71 and IV = 27)	None	NA	Diagnostic assay showed promising results in detecting CRC (sensitivity of 93.75%, a specificity of 82.89%)	Screening	[24]
ORM2	iTRAQ coupled with micro Q-TOF/MS.	Plasma: 10 CRC and 10 healthy	ELISA WB	Plasma: 65 control, 59 hyperplastic polyp, 62 inflammatory bowel disease, 53 adenoma and 180 CRC (AJCC stage: I = 49, II = 31, III = 62 and IV = 38) Tissues: 41 pairs of normal and CRC samples	ORM2 level in plasma and tissue was higher in CRC compared with the healthy samples	Screening	[25]
Clusterin	LC-ESI-MS/MS	Plasma: 10 CRC and 10 healthy in each of first two phases	Targeted LC-MS (SRM)	Plasma: 48 CRC and 48 healthy (Stage of CRC not stated)	Plasma levels of clusterin were higher in CRC compared with control and protein was associated with risk of CRC (only in men)	Screening	[26]

Autoantibodies against: EDIL3, GTF2B, HCK, P53, PIM1 and STK4	MALDI-TOF-MS	Proteins expressed in E.coli	Multiplex beads assay and ELISA	Sera: 135 CRC (AJCC stage: I = 35, II = 25, III = 46 and IV = 29), 65 other cancer types, 14 inflammatory bowel disease and 93 healthy	Combination of autoantibodies achieved diagnostic accuracy of 89.7%, with 66% sensitivity at 90.0% fixed specificity	Screening	[27]
Collagen I	LC–MS	Sera: 91 CRC (UICC stage: I = 21, II = 41, III = 22 and IV = 7) and 33 healthy	ELISA WB PCR	Sera: same cohort used in proteomic Tissues: 26 pair of normal and CRC (UICC stage: I = 7, II = 7, III = 7 and IV = 5) Sera: same cohort as used in WB	The expression of collagen I may be an early event in CRC tumorigenesis and could provide prognostic information	Screening and prognosis	[28]
Adipophilin	LC/MS (2DICAL)	Plasma: 21 healthy and 22 CRC (AJCC stage: I = 3, II = 6, III = 8 and IV = 5)	RPPM	Plasma: 109 healthy and 101 CRC (AJCC stage: I = 19, II = 31, III = 32 and IV = 17) Plasma: 87 healthy and 26 CRC (AJCC stage: I = 12, II = 5, III = 8 and IV = 1)	Adipophilin is expressed primarily in the basal sides of CRC cells, while it is absent from adjacent normal mucosa, and the detection power was superior to that of CEA	Screening	[29]
MRC1 and S100A9	LC-MS	Sera: 3 healthy and 3 CRC	WB ELISA	Cell lines and sera: 3 healthy and 3 CRC Sera: 96 healthy and 112 CRC (AJCC stage: I = 21, II = 50 and III = 41)	Proteins were differentially expressed between normal and CRC	Screening	[30]
MAPRE1	LC/MS	Mouse model, cell lines and plasma: 60 adenomas, 60 CRC (AJCC stage: I = 11, II = 19, III = 21 and IV = 9) and 60 healthy	Antibody array IHC	Plasma: 60 adenomas, 60 CRC (AJCC stage: I = 11, II = 19, III = 21 and IV = 9) and 60 healthy Tissues: 20 normal tissues, 10 adenomas, and 66 CRC (stage not provided)	Protein levels were higher in adenoma and early stages of CRC compared with normal mucosa, the diagnostic power is stronger with other markers (CEA and AK1)	Screening and diagnosis	[31]

Proteins (A2M, APOH, IGL@, MACF1 and VDB) and metabolite signature	2DIGE, Finnigan LTQ-MS and GC-MS	Sera: 30 CRC (Dukes' stage: A = 3, B = 13, C = 8 and D = 6) and 30 healthy	ELISA	Sera: same cohort as used in proteomics	Differential expression of proteins in CRC compared with healthy. 93.5% of CRC patients are identified using the 6 metabolites	Screening	[32]
Volatile organic compounds signature	FAIMS and GC-MS	Urine: 83 CRC (AJCC stage: I = 9, II = 24, III = 32, IV = 9 and no stage = 9) and 50 control	None	NA	Sensitivity and specificity for CRC detection were 88% and 60% respectively	Screening	[33]
Angiopoietin-2, calprotectin, FGF-23, IL-13, M2-PK, MMP-10 and TPO	Biotin label-based protein array	Faeces: 20 CRC (AJCC stage: I = 2, II = 7, III = 9 and IV = 2) and 20 healthy	ELISA and multiplex faecal protein biochip	Faeces: same cohort as used in proteomics	Proteins levels are significantly higher in CRC compared with healthy controls	Screening	[34]
A1AT and CTSD	Gel-enhanced LC−MS	Tissues: 37 CRC (AJCC stage: I = 13 and II = 24) and 37 normal	IHC WB	Tissues: 93 CRC (AJCC stage: I = 4, II = 86 and III = 3) Tissues: paired samples from 14 early stage CRC, and sera: 84 samples (42 early CRC and 42 healthy)	Less A1AT and more CTSD in CRC compared with healthy samples. Combination of both proteins identified 96.77% of CRC	Screening	[35]
EFEMP2	Nano LC-MS/MS	Tissues: 9 CRC (AJCC stage: I = 7 and II = 2) paired with normal mucosa	IHC ELISA WB	Tissues: 88 CRC (UICC stage: I = 23, II = 29, III = 26 and IV = 10), 19 adenoma and 16 normal colon Sera: 79 healthy, 14 adenoma, and 122 CRC, stage not stated, but smallest proportion of cases is in stage I (figure 7B) Tissues and sera: 9 pairs of CRC and normal mucosa	The expression level of EFEMP2 increases in early stages CRC. Diagnostic accuracy significantly better than CEA	Screening	[36]

CAMP, ERp29, HSPA8 and TPM3	2D LC-MS/MS	Tissues: 3 CRC and 3 normal	IHC WB	Tissues: 69 CRC matched with normal (AJCC stage: I = 15, II = 21 and III = 33) Tissues: 3 CRC and 3 normal	The protein panel can detect CRC via IHC (accuracy of 73.2%)	Screening	[37]
SORD	iTRAQ 8-plex labelled LC-MS/MS	Cell lines and tissues: 30 adenomas and 30 normal	IHC WB	Cell lines and tissue: normal colon, colorectal adenomas, and adenocarcinomas (numbers not provided) Cell lines and tissues: 4 pairs of adenoma and normal mucosa	Significant increase in SORD expression in adenomas and cancer cell lines	Screening	[38]
TRFM	LC-MS/MS	Cell lines	ELISA WB	Plasma: 77 healthy and 228 CRC (I = 68, II = 68, III = 65 and IV = 27) Plasma: 80 CRC (20 per stage), 10 adenoma, 10 polyps and 30 healthy controls	TRFM expression increases in stages I and II compared with stages III and IV	Screening	[39]

References

1. Croner RS, Stürzl M, Rau TT, et al. Quantitative proteome profiling of lymph node-positive vs.-negative colorectal carcinomas pinpoints mx1 as a marker for lymph node metastasis. Int J Cancer. 2014;135:2878-2886.

2. Peng Y, Li X, Wu M, et al. New prognosis biomarkers identified by dynamic proteomic analysis of colorectal cancer. Mol Biosyst. 2012;8:3077-3088.

3. Snoeren N, Emmink BL, Koerkamp MJ, et al. Maspin is a marker for early recurrence in primary stage III and IV colorectal cancer. Br J Cancer. 2013;109:1636-1647.

4. Han CL, Chen JS, Chan EC, et al. An informatics-assisted label-free approach for personalized tissue membrane proteomics: case study on colorectal cancer. Mol Cell Proteomics. 2011;10:M110.003087.

5. Sethi MK, Thaysen-Andersen M, Kim H, et al. Quantitative proteomic analysis of paired colorectal cancer and non-tumorigenic tissues reveals signature proteins and perturbed pathways involved in CRC progression and metastasis. J Proteomics. 2015;126:54-67.

6. O'Dwyer D, Ralton LD, O'Shea A, et al. The proteomics of colorectal cancer: identification of a protein signature associated with prognosis. PloS One. 2011;6:e27718.

7. Zougman A, Hutchins GG, Cairns DA, et al. Retinoic acid-induced protein 3: identification and characterisation of a novel prognostic colon cancer biomarker. Eur J Cancer. 2013;49:531-539.

8. Kim CY, Jung WY, Lee HJ, et al. Proteomic analysis reveals overexpression of moesin and cytokeratin 17 proteins in colorectal carcinoma. Oncol Rep. 2012;27:608-620.

9. Meding S, Balluff B, Elsner M, et al. Tissue-based proteomics reveals FXYD3, S100A11 and GSTM3 as novel markers for regional lymph node metastasis in colon cancer. J Pathol. 2012;228:459-470.

10. Tweedle EM, Khattak I, Ang CW, et al. Low molecular weight heat shock protein HSP27 is a prognostic indicator in rectal cancer but not colon cancer. Gut. 2010;59:1501-1510.

11. Farshidfar F, Weljie AM, Kopciuk K, et al. Serum metabolomic profile as a means to distinguish stage of colorectal cancer. Genome Med. 2012;4:42.

12. Mahboob S, Ahn SB, Cheruku HR, et al. A novel multiplexed immunoassay identifies CEA, IL-8 and prolactin as prospective markers for Dukes' stages A-D colorectal cancers. Clin Proteomics. 2015;12:10.

13. Tsai MH, Wu CC, Peng PH, et al. Identification of secretory gelsolin as a plasma biomarker associated with distant organ metastasis of colorectal cancer. J Mol Med. 2012;90:187-200.

14. Surinova S, Radova L, Choi M, et al. Non-invasive prognostic protein biomarker signatures associated with colorectal cancer. EMBO Mol Med. 2015;7:1153-1165.

15. Mehta RS, Chong DQ, Song M, et al. Association between plasma levels of macrophage inhibitory cytokine-1 before diagnosis of colorectal cancer and mortality. Gastroenterology. 2015;149:614-622.

16. Tan HT, Wu W, Ng YZ, et al. Proteomic analysis of colorectal cancer metastasis: stathmin-1 revealed as a player in cancer cell migration and prognostic marker. J Proteome Res. 2012;11:1433-1445.

17. Qiao J, Fang CY, Chen SX, et al. Stroma derived COL6A3 is a potential prognosis marker of colorectal carcinoma revealed by quantitative proteomics. Oncotarget. 2015;6:29929-29946.

18. Surinova S, Choi M, Tao S, et al. Prediction of colorectal cancer diagnosis based on circulating plasma proteins. EMBO Mol Med. 2015;7:1166-1178.

19. Chen H, Werner S, Butt J, et al. Prospective evaluation of 64 serum autoantibodies as biomarkers for early detection of colorectal cancer in a true screening setting. Oncotarget. 2016: published on 19 February 2016, doi:10.18632/oncotarget.7500.

20. Ladd JJ, Busald T, Johnson MM, et al. Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women. Cancer Prev Res. 2012;5:655-664.

21. Wild N, Andres H, Rollinger W, et al. A combination of serum markers for the early detection of colorectal cancer. Clin Cancer Res. 2010;16:6111-6121.

22. Werner S, Krause F, Rolny V, et al. Evaluation of a 5-marker blood test for colorectal cancer early detection in a colorectal cancer screening setting. Clin Cancer Res. 2016;22:1725-1733.

23. Murakoshi Y, Honda K, Sasazuki S, et al. Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray. Cancer Sci. 2011;102:630-638.

24. Brock R, Xiong B, Li L, et al. A multiplex serum protein assay for determining the probability of colorectal cancer. Am J Cancer Res. 2012;2:598-605.

25. Zhang X, Xiao Z, Liu X, et al. The potential role of ORM2 in the development of colorectal cancer. PLoS One. 2012;7:e31868.

26. Bertuzzi M, Marelli C, Bagnati R, et al. Plasma clusterin as a candidate pre-diagnosis marker of colorectal cancer risk in the Florence cohort of the European prospective investigation into cancer and nutrition: a pilot study. BMC Cancer. 2015;15:56.

27. Villar-Vazquez R, Padilla G, Fernandez-Acenero MJ, et al. Development of a novel multiplex beads-based assay for autoantibody detection for colorectal cancer diagnosis. Proteomics. 2016;16:1280-1290.

28. Zou X, Feng B, Dong T, et al. Up-regulation of type I collagen during tumorigenesis of colorectal cancer revealed by quantitative proteomic analysis. J Proteomics. 2013;94:473-485.

29. Matsubara J, Honda K, Ono M, et al. Identification of adipophilin as a potential plasma biomarker for colorectal cancer using label-free quantitative mass spectrometry and protein microarray. Cancer Epidemiol Biomarkers Prev. 2011;20:2195-2203.

30. Fan N, Chen H, Song W, et al. Macrophage mannose receptor 1 and S100A9 were identified as serum diagnostic biomarkers for colorectal cancer through a label-free quantitative proteomic analysis. Cancer Biomark. 2016;16:235-243.

31. Taguchi A, Rho JH, Yan Q, et al. MAPRE1 as a plasma biomarker for early-stage colorectal cancer and adenomas. Cancer Prev Res. 2015;8:1112-1119.

32. Ma Y, Zhang P, Wang F, et al. An integrated proteomics and metabolomics approach for defining oncofetal biomarkers in the colorectal cancer. Ann Surg. 2012;255:720-730.

33. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, et al. Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. PLoS One. 2014;9:e108750.

34. Wang HP, Wang YY, Pan J, et al. Evaluation of specific fecal protein biochips for the diagnosis of colorectal cancer. World J Gastroenterol. 2014;20:1332-1339.

35. Xie L, Zhao C, Cai S, et al. Novel proteomic strategy reveal combined α1 antitrypsin and cathepsin D as biomarkers for colorectal cancer early screening. J Proteome Res. 2010;9:4701-4709.

36. Yao L, Lao W, Zhang Y, et al. Identification of EFEMP2 as a serum biomarker for the early detection of colorectal cancer with lectin affinity capture assisted secretome analysis of cultured fresh tissues. J Proteome Res. 2012;11:3281-3294.

37. Fan N, Gao J, Liu Y, et al. Label-free quantitative mass spectrometry reveals a panel of differentially expressed proteins in colorectal cancer. Biomed Res Int. 2015;2015:365068.

38. Uzozie A, Nanni P, Staiano T, et al. Sorbitol dehydrogenase overexpression and other aspects of dysregulated protein expression in human precancerous colorectal neoplasms: a quantitative proteomics study. Mol Cell Proteomics. 2014;13:1198-1218.

39. Shin J, Kim H, Kim G, et al. Discovery of melanotransferrin as a serological marker of colorectal cancer by secretome analysis and quantitative proteomics. J Proteome Res. 2014;13:4919-4931.