Abstract. Colorectal cancer (CRC) is the second leading cause of cancer-related death. Despite the progress that has been made towards the identification of the molecular mechanisms involved in CRC, currently there are many unclear points. The current opinion is that microsatellite instability (MSI), CpG island methylator phenotype (CIMP) and chromosomal instability (CIN) seem to play a significant role. MSI is related to point mutations in defective mismatch repair system of DNA. There are two well-established MSI phenotypes: MSI-high (MSI-H) and MSI-low (MSI-L or MSS). CIN refers to a different cellular event which originates from the presence of an abnormal chromosome complement or number. CIMP is the third most commonly involved event, and is defined by widespread methylation of CpG islands of suppressor promoters, with two phenotypes: CIMP-high and CIMP-low which interact with MSI or CIN status V-raf murine sarcoma viral oncogene homolog B (BRAF) is a serine-threonine protein kinase that acts as a downstream effector of the Kirsten rat sarcoma viral oncogene (KRAS) pathway. Various studies have revealed that BRAF V600E mutations appear to be a valid indicator of poor prognosis. KRAS is a proto-oncogene which encodes a GTP-ase involved in cellular response to extracellular stimuli. Its prognostic value is still controversial. However, wild-type KRAS is associated with better response to Endothelial Growth Factor Receptor inhibitors combined with standard chemotherapy. Loss of Heterozygosity, especially involving 18q, is a well-known potential mechanism for tumorigenesis that has been studied in CRC. Vascular endothelial growth factor is a pro-angiogenic factor linked with the aggressiveness of CRC. Emerging data show that cyclooxygenase 2 overexpression is significantly associated with worse outcomes in CRC. Recent studies highlight mi-croRNAs as promising prognostic biomarkers. More specifically, the down-regulation of miR-451, miR-625, miR-29c, miR-126, miR-129 and miR133 is purported to be a prognostic factor, while miR-224 was overexpressed in CRC.
Currently, treatment strategies and clinical outcomes in colorectal cancer (CRC) are determined by cancer stage as defined by TNM or relevant staging protocols (1), based on local tumor penetration, and spread to lymph nodes or other organs. The cornerstone of non-metastatic CRC treatment remains surgical resection. In patients with higher stage disease (III or high risk II) however, there is a significant risk of recurrence. Although certain histological factors such as tumor differentiation, grade, and lymphovascular invasion, have been identified as being to higher risk, there is still a lack of understanding of the molecular factors which may affect the risk of metastasis and recurrence.

In recent years, molecular biomarkers have generated interest as prognostic markers, or as in the case of Kirsten rat sarcoma viral oncogene homolog (KRAS), as factors involved in treatment choice. Generally, a biomarker is defined as a characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological response to a specified therapeutic intervention. Biomarkers can be either prognostic or predictive (2, 3). Prognostic values are defined by the estimated life expectancy post-diagnosis and treatment, whereas predictive biomarkers are related to the response of a patient to a relevant treatment strategy.

The purpose of a molecular classification is to identify similar characteristics among individual tumors, and then empirically predict the pathogenesis and biological behavior of a particular tumor (4, 5). The most accepted way of creating a classification model is to identify and correlate single cellular events that have been statistically proven to play a role in tumorigenesis.

The starting point of these efforts was the identification of significant stages in the progression from healthy epithelium to cancer (6, 7). Currently, some theories exist which explain the transition from adenoma, the first precancerous type of lesion, to carcinoma and then invasion to other tissues. In this way, commonly involved events in CRC were identified and highlighted as the guiding points for a classification model. Subsequently, the first efforts to summarize and correlate these events were published, but none was actually adequately evidence-based to achieve translation to clinical practice.

Most of the current molecular classification models are based on Microsatellite Instability (MSI), Chromosomal Instability (CIN) and CpG island Methylator Phenotype (CIMP) and they correlate these events with other significant mutations i.e. KRAS, and V-raf murine sarcoma viral oncogene homolog B (BRAF) (7, 8).

In this review article, we summarize what is currently believed in terms of the popular potential biomarkers in CRC and set the parameters for further discussion and points of interest for further research.

MSI

Microsatellite disorders are established as frequent events in CRC, occurring in around 22% of cases (9). There are many reports supporting the effect of MSI on CRC prognosis (3, 10-17).

Microsatellites are short repetitive DNA nucleotide sequences which are prone to frame-shift mutations and base-pair substitutions during DNA replication (3). They are involved in the DNA repair system and their mutations were initially associated with Lynch syndrome, which is an autosomal dominant disorder, highly associated with 3% of multiple CRC types. Mismatch Repair (MMR) genes involved in Lynch syndrome are mutS homolog 1 (MLH1), mutS homolog 2 (MSH2), mutS homolog 6 (MSH6) and post-meiotic segregation increased 2 (PMS2) (18). Loss of these genes results in defective MMR, which in turn results in MSI.

Sporadic CRC is more likely to be associated with MLH1 mutation. In terms of MSI status classification, most studies divide it into three categories: MSI-high (MSI-H) at ≥30%, MSI-low (MSI-L) at 10-30% and microsatellite stable (MSS) (3, 18). Some other studies aim to simplify MSI status as being positive or negative (10, 19, 20). It has been questioned whether MSI-L exists or falls into the same category as MSS (21).

Most studies aiming to establish a classification model in CRC consider MSI status as one of the primary criteria for categorization (10, 20) (Table I). A recent multivariate analysis identified MSI alterations in 119/892 CRC samples and highlights MSI status as one of the major prognostic biomarkers (10). Another interesting study suggested a model which is mainly based on MSI status and its relationship with CpG island Methylator Phenotype (CIMP) and Chromosomal Instability (CIN). In the same study, it was noted that Adenomatous polyposis coli (APC) and KRAS mutations were less frequent in MSI tumors (13.3% and 10.9%, respectively) whereas BRAF V600E mutation is most often seen in MSI tumors (58.7%) (19).

There are many clinicopathological variables associated with MSI status. For instance, patients with MSI-H tumors in general have greater 5-year overall survival irrespective of stage compared to those with MSI-L and MSS tumors. The Pan-European Trials in Adjuvant Colon Cancer III (PETACC III) trial confirmed these retrospective findings and suggested that MSI-H is a strong prognostic factor for relapse-free and overall survival in patients with stage II and III CRC (3). Moreover, MSI-H tumors have been classified as proximally located, poorly-differentiated and with a higher incidence in female gender. They are also characterized by mucinous differentiation, increased age at onset, and round and vesicular nuclei with a prominent nucleolus (2, 10, 18-20).

However, contrary to the promising results in terms of the prognostic value of MSI status, no clear relationship between
MSI status and response to neoadjuvant chemotherapy has yet been proven. However, Ribic et al. concluded that MSI-H status can be associated with poor response to 5-Flurouracil-based chemotherapy compared to MSI-L and MSS (2). Nevertheless, there are multiple studies with conflicting results in terms of the predictive value of MSI status and thus this is still under discussion (20).

**CIMP**

Most studies develop the CRC classification model based on CIMP status and its correlation to CIN and MSI (4, 18-20, 22). CIMP is defined as hypermethylation of CpG island promoter. CIMP results in transcriptional silencing of specific tumor-suppressor and DNA repair genes, including MLH1 (19). Some studies describe CIMP status as CIMP-high, CIMP-low and CIMP-negative (18). There is a discussion whether CIMP low and CIMP negative fall into the same category, resulting in two categories, CIMP-positive and CIMP-negative (22); we use this for simplicity here.

There have been some reasonable studies which associate MSI and CIMP status with CIN, resulting in a classification model which attempts to identify and link macroscopic variables to molecular features (19). In a recent study, Simons et al. suggest a classification based on MSI, CIMP and CIN. According to that, MSI and CIMP-positive tumors were more proximally located (85.7% and 51.9%, respectively), whereas CIMP-positive and CIN or CIN-only tumors were distally located (distal colon to rectum 52.3% and 82.1%, respectively). Triple-negative tumours were located in both proximal and distal colon and in 89% of them BRAF V600E mutation was identified. Regarding CIMP-only tumors, the average age at diagnosis was 67.6 years, 51.9% were located proximally, and the differentiation grade was II at the stage of data analysis. BRAF V600E mutation was present only in 18.5% and p53 overexpression was noted in 66.7% (19).

**CIN**

Tumors with CIN have chromosomal gains and losses, with or without structural rearrangements, possibly reflecting an increased mutation rate; this is called aneuploidy and polyploidy, respectively (10, 23, 24), and occurs in around 60% of CRC cases (2). The mechanism underlying CIN is not yet understood. Most studies tend by definition to differentiate MSI-positive tumors from CIN, as CIN reflects a separate mechanism of carcinogenesis (19, 20). MSI-positive and CIN-positive tumors have been described but they tend to macroscopically appear more like MSI-positive (10). The overlap between CIN and MSI is not clear (19). CIN-positive tumors are generally associated with poor prognosis and tend to be well- or moderately differentiated (10). Simons et al. hold that CIN-only tumors are more frequent in men (54.7%), 76.5% at the stage of diagnosis are grade 2 and there is subsequent p53 overexpression (19). However, BRAF V600E mutation tends not to be involved in such a mechanism of tumorigenesis (10, 19), and some studies do link KRAS mutations with this molecular phenotype (19) (20). CIN tumors are located more often distally. CIMP- and CIN-positive tumors have been described (19).

Domingo et al. have described 3% of MSI-H and CIN-positive tumors tending to have characteristics closer to MSI-H CIN-negative tumors i.e. right-sided location and increased BRAF mutation, while p53 mutation and 17p LOH was not increased significantly, and thus there was no association (10).

**KRAS Mutation**

KRAS is a proto-oncogene involved in cellular response to extracellular stimuli (3, 25). In cases of KRAS mutation, there is a structural activation of downstream signaling pathways i.e. mitogen-activated protein kinase (MAPK) and phosphoinositide-3-kinase/v-akt murine thymoma viral oncogen (PI3K/AKT) pathways (2), and through this mechanism, tumor cells are more
resistant to inhibition of surface receptors of tyrosine kinase such as epidermal growth factor receptor (EGFR). The MAPK pathway was recently found to be one of the fundamental mechanisms involved in tumorigenesis (9).

Point mutations in codons 12, 13 and 61 of KRAS gene are noted in about 30-40% of CRC (2) and in total, around 85 KRAS mutations have been identified (9). These different molecular phenotypes of KRAS may result in different pathways of carcinogenesis, which could alter the macroscopic phenotype.

However, it is still controversial whether KRAS can be considered as a valid biomarker or not, as recent studies suggest there are no clinicopathological feature to support this (3, 9). There are some controversial data emerging that link KRAS mutation with poor prognosis. Phipps et al. found KRAS mutation to be associated with poorer prognosis compared to wild-type KRAS (26). Most studies link KRAS mutation with certain molecular phenotypes i.e. CIN, whereas they are less likely to be noted in MSI-positive tumors (3, 10, 19, 20).

An interesting study in Moroccan patients with advanced CRC showed that 76.09% of patients had wild-type KRAS genotype, whereas 23.91% were KRAS mutants. The majority of KRAS mutations referred to an amino acid substitution of glycine by aspartic acid (68.2%) (27).

On the other hand, despite the debating prognostic value of KRAS, it has been established that wild-type KRAS is associated with better response to EGFR inhibitors in terms of adjuvant chemotherapy setting i.e. cetuximab (1, 3, 9, 28-31). An interesting study by Di Bartolomeo et al. in 2013 concluded that patients with KRAS/neuroblastoma RAS viral (v-ras) oncogene homolog, BRAF and p53 wild-type tumors had the maximum benefit from treatment with cetuximab, oxaliplatin and tegafur-uracil oral chemotherapy (32). Nevertheless, Selcukbikir et al. found that BRAF and KRAS mutations do not seem to be potential biomarkers in the treatment of metastatic CRC with bevacizumab therapy (33).

Brüera et al. also noted a worse prognosis in patients with metastatic CRC and KRAS (c35 G)/BRAF mutations (34), and a similar hypothesis is supported by Loupakis et al. where EGFR ligand was significantly modulated by cetuximab plus irinotecan therapy (35).

Tian et al. state that a combined signature of KRAS/BRAF and phosphatidyli-nositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PI3KCA) potentially offered an optimized source of information for cetuximab response-siveness of the tumor (36). Saridaki et al. concluded that KRAS and BRAF mutations, and EGFR expression can be used as biomarkers to further select patients undergoing anti-EGFR therapy (37).

**BRAF Mutation**

*BRAF* gene encodes a serine-threonine kinase that acts as an inhibitor of the RAS/MAPK intracellular signaling pathway. Mutations of *BRAF* occur at an early stage of colorectal carcinogenesis (38).

In terms of correlations of *BRAF* with the main cellular events, most studies associate *BRAF* V600E activating mutation with sporadic MSI-H CRCs (2, 3, 9, 10, 19) and it is linked with poor prognosis (39) (Table II). What is more, no association between *BRAF* mutations and CIN has yet been noted (2, 10, 19). According to Doming et al., *BRAF* alterations occur in 10% of CRC cases and there is a negative association with *KRAS* mutation (10), whereas there is primary positive association with MSI and CIMP-high (4, 31).

In a recent study by Lochhead et al., *BRAF* mutation occurred in 10% to 20% of CRC and was associated with MSI-high through its relationship to CIMP-high (13). The same study concluded that *BRAF* mutations are associated with inferior prognosis. Eklof et al. analyzing the mutation status in KRAS and *BRAF*, and PIK3CA and PTEN expression in two separate CRC cohorts, state that *KRAS* and *BRAF* status are important in the establishment of the prognosis in CRC and should always be considered (40). Similarly, Yokota et al. conclude that mutated *BRAF* is one of the most powerful prognostic markers in CRC (41).

Nevertheless, *BRAF* as a predictive marker is still under discussion (3, 29). There are some studies which tend to use a signature of *BRAF/KRAS* as a predictive factor for the response to EGFR inhibitors (32-37). Another interesting meta-analysis (42), taking into consideration 21 trials including 5229 patients, concluded that *BRAF* mutation is a predictive biomarker for poor prognosis in patients with metastatic CRC undergoing therapy with anti-EGFR monoclonal antibody, especially in those with wild-type KRAS.

Finally, Mesteri et al. concluded that *BRAF* V600 mutation can be used as a classification criterion for the evaluation of serrated lesions and progression to sessile/serrated adenoma polyps and CRC.

**Vascular Endothelial Growth Factor (VEGF)**

VEGF is a proangiogenic factor involved in endothelial cell proliferation, migration and vascular permeability (3). Increase in its expression is associated with poor prognosis, low response to preoperative radiotherapy, and greater likelihood of recurrence (43).

Angiogenesis plays a significant role in CRC as neovascularization markers are overexpressed in most of the cancer cells (44). VEGFA is involved in early tumor stages i.e. adenoma formation. According to the same study, VEGFA was associated with advanced cancer stage and metastatic disease via a angiolympathic invasion pathway.

Moreover, a relationship between VEGFC and collagen triple helix containing 1 has been noted in terms of prognostic value in rectal cancer (45). VEGFC activates the tyrosine kinase-linked receptor of the VEGFR3 pathway.
and subsequently explain the prognostic value of \textit{Enzyme Substrate} \textit{(SMAD4 Homolog 4 (SMAD4))} （Disease Free Survival.）with negativity for \textit{VEGF} was strongly correlated with improved for \textit{RAF kinase inhibitor protein} with \textit{COX2} negativity, or \textit{Disease Free Survival} （\textit{BRAF-}）associated with lower risk of wild-type \textit{CRC} but not of \textit{COX2}-related Pathway. Several prospective studies have shown that \textit{COX2} overexpression could be interpreted as an adverse prognostic factor for \textit{CRC} （46）. There are some studies which support that the use of \textit{COX2} inhibitors, \textit{i.e.} aspirin, can be associated with lower risk of \textit{CRC}, especially in patients who have been diagnosed with \textit{CRC} in which \textit{COX2} is overexpressed （47）.

Another study evaluated the correlation between aspirin intake and \textit{CRC} according to \textit{BRAF} mutation status using questionnaire on aspirin use, and collected data from 1986 to 2006 （48）. The study concluded that regular aspirin use was associated with lower risk of \textit{wild-type BRAF CRC} but not of \textit{BRAF-mutated cancer risk}.

Finally, a recent study noted （49）that positivity for \textit{COX2} and \textit{VEGF} was strongly correlated with decreased \textit{Disease Free Survival} （\textit{p}=0.007）, whereas combinations of positivity for \textit{RAF kinase inhibitor protein} with \textit{COX2 negativity}, or with negativity for \textit{VEGF} was strong correlated with improved \textit{Disease Free Survival}.

\textbf{LOH 18q}

\textit{LOH}, especially in 18q, is a well-known potential mechanism for \textit{tumorigene-sis} that has been studied in \textit{CRC}. \textit{LOH} is an important mechanism of inactivation of tumor-suppressor genes （3）. There have been several studies to inter-pret the potential prognostic effect of \textit{LOH 18q} in \textit{CRC}. \textit{LOH} in 18q has been inversely associated with \textit{MSI} （15）. There are many tumor-suppressor genes in 18q, \textit{e.g. Deleted in Colorectal Carcinoma (DCC)}, \textit{Mothers Against Decapen-taplegic Homolog 4 (SMAD4)}, \textit{SMAD2}, and \textit{CDK5 And ABL1 Enzyme Substrate (CABLES1)} （15）which might interfere with and subsequently explain the prognostic value of \textit{LOH 18q}.

Table II. \textit{KRAS} and \textit{BRAF} gene mutations baseline characteristics. \textit{This table describes the baseline characteristics of \textit{KRAS} and \textit{BRAF} mutations regarding predictive value and common mutation patterns.}

<table>
<thead>
<tr>
<th>Gene wt/mutant</th>
<th>Baseline characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{KRAS} wt</td>
<td>Better response to \textit{EGFR inhibitors} \textit{(1, 3, 9, 28-31)}</td>
</tr>
<tr>
<td>\textit{KRAS} Mutant</td>
<td>Most likely at codons 12, 13, 61 （2）</td>
</tr>
<tr>
<td>\textit{BRAF V600E mutation}</td>
<td>Predictive value is still under discussion （3, 29）</td>
</tr>
</tbody>
</table>

\textit{BRAF V600E: v-raf murine sarcoma viral oncogene homolog B V600E mutation; EGFR: epidermal growth factor.}

Nevertheless, in the same study, in 555 non-\textit{MSI}-high tumor samples with informative 18q LOH data, LOH was present in at least one 18q marker in 362 tumors （65%）. There was an association between \textit{LOH 18q and obesity} （\textit{Body mass index} ≥30 kg/m²; \textit{p}=0.018）, distal colon location （\textit{p}=0.025）, low tumor grade （\textit{p}=0.0060）, low-level long interspersed nucleotide element-1 methylation （\textit{LINE1}）methylation （\textit{p}=0.040）, wild-type \textit{KRAS} （\textit{p}=0.015）, and \textit{ohn Cunningham (JC) virus T antigen (JCVT)} （\textit{p}=0.0004）. The same study concluded that there is no association of \textit{LOH 18q} with prognosis as yet.

\textbf{miRNA – Developing New Promising Biomarkers}

\textit{miRNAs are short, 18-25 nucleotide non-coding single-stranded RNA se-quences which are involved in regulation of gene expression on a post-transcriptional level, through binding to their target protein-encoding mRNA} （50）. They are believed to play a potentially significant role in the \textit{pathogenesis of CRC} （51）. There are several studies which identify the presence of vari-ous \textit{miRNAs} as predictive or prognostic biomarkers （51-57）. In the study of \textit{Svoboda et al.}, \textit{miR-215}, \textit{miR-99a*}, \textit{miR-196b}, \textit{miR-450b-5p} and \textit{LET-7e} were associated with expression of thymidylate synthetase and radioresistance or chemoresistance to its inhibitors （51）.

\textit{miRNAs are also involved in the regulation of the EGFR signaling pathway} （57） and thus they may have essential value in predicting the response to \textit{EGFR inhibitors}. Another study identified that \textit{miR-451} inhibits cell growth through down-regulating the \textit{P13K/AKT pathway} and thus it potentially has a repressive role （52）. \textit{Lou et al.} claimed that \textit{miR-625} may have a prognostic and predictive role in \textit{CRC} as decreased levels are associated with high inci-dence of metastasis and poor prognosis （53）. \textit{Yang et al.} identified \textit{miR-29c} as having an \textit{antitumorigenic role} and, preoperatively, decreased levels of \textit{miR-29c} were associated with \textit{CRC relapse}. \textit{miRNA-126} can have \textit{anticarcinogenic properties} through the inhibition of \textit{neovascularization via blocking the 3- UTR'} of \textit{VEGF} （55）.

\textit{Karaayvaz et al.} highlighted that \textit{miR-129} could potentially have a tumor-suppressor role and might be a novel target for

---

\textit{Sideris and Papagr...Molecular Biomarkers in Colorectal Cancer (Review)}
search with promising results (Table III). -2 genes (58). miRNAs seem to be an interesting field of research based on their prognostic and predictive value will generate a new level of molecular heterogeneity in colorectal cancer. J Pathol 229(3): 441-448, 2013.


anti-CRC therapies (56). Dong et al. claim that miR-133a serves as a functional tumor suppressor in CRC via enhancing apoptosis and inhibiting cell proliferation (57). According to the same study, miRNA 133a increased p53 protein and induced p21 transcription, and can serve as a sensitizer to doxorubicin and oxaliplatin. One potential mechanism of the anticancer properties of miRNA-133a is the fact that it can repress the ring finger and FYVE-like domain containing E3 ubiquitin protein ligase 3’UTR reporter activity and reduce its protein level. Finally, Liao et al. noted that miRNA-224 can induce tumor expansion and cell proliferation via repressing PH domain and leucine rich repeat protein phosphatase 1 and -2 genes (58). miRNAs seem to be an interesting field of research with promising results (Table III).

Conclusion

CRC is the second most frequent cancer in Europe and despite progress in the understanding of the molecular background of CRC, effective molecular classification remains a challenge. Facing these difficulties in the establishment of a key model, researchers have started to try different, more innovative approaches, in which gene signatures dominate. In fact, the idea of gene signatures is very promising (59, 60), but there are still several issues of reproducibility and possible overlap (61, 62). Metabolomics is a new and promising field which tends to utilize simple products from cellular metabolism in order to identify differential and pathognomonic features of cancer (63-66). The main advantage of this approach is the easy access; however, the low specificity for CRC remains a great challenge. Nevertheless, there is much to be looked at in this field, as metabolomics could potentially be a new concept in the investigation and classification of CRC.

Thus, molecular classification of CRC still remains a challenge that researchers have to overcome in the coming years. Creating an effective classification of CRC biomarkers based on their prognostic and predictive value will generate an effective decision-making tool in everyday’s clinical practice.


