Genomics and Proteomics Approaches for Biomarker Discovery in Sporadic Colorectal Cancer with Metastasis

MARÍA GONZÁLEZ-GONZÁLEZ¹, JACINTO GARCIA GARCIA², JOSÉ ANTONIO ALCAZAR MONTERO², LUIS MIGUEL GONZÁLEZ FERNANDEZ², OSCAR BENGOECHEA³, OSCAR BLANCO MUÑEZ³, ALBERTO ORFAO¹, J.M. SAYAGUES¹ and MANUEL FUENTES¹

¹Servicio General de Citometría, Departamento de Medicina and Centro de Investigación del Cáncer and Units of ²Cirugía and ³Unidad de Patología, Hospital Universitario de Salamanca-IIBSAL, Salamanca, Spain

Abstract. Sporadic colorectal cancer (sCRC) is one of the most frequent types of cancer in Europe. Despite understanding of its tumorigenesis, the metastatic process is not yet clear. In this article, we review the most significant genetic and proteomic advances that have been made in regards to research of the sCRC metastatic process, and the new biomarkers that are able to predict disease prognosis or response to treatments, in order to define personalized medicine.

Sporadic colorectal cancer (sCRC) is one of the most frequent types of cancer in Europe; the number of newly-diagnosed cases per year is 450,000. This pathological entity presents with a high mortality rate, the principal cause being the existence of metastasis, especially in the liver. For these reasons, sCRC is also one of the most extensively studied types of cancer; since the 1980s knowledge on this pathology has increased massively, with much research taking place into the genesis of this disease. Despite understanding of tumor genesis, the metastatic process is still yet not clear. In this article, we review the most significant genetic and proteomic advances that have been made in researching the metastatic process of sCRC, and the new biomarkers that are able to predict the prognosis of this disease or response to treatments, in order to define personalized medicine.

Genomic Alterations Involved in Liver Metastatic sCRC

In 1990, Fearon and Vogelstein presented their hypothesis of colorectal cancer tumorigenesis, which defines the genetic alterations involved in transformation from normal intestinal mucosa to colorectal carcinoma (1). This aberrant transformation is a multi-step process that includes three essential genetic alterations in the genesis of the pathology: i) mutation of the adenomatous polyposis coli (APC) gene, coded in chromosomal region 5q, which is thought to occur early on during the development of adenomatous polyps; ii) activation of v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS) gene, coded in the chromosomal region 12p, during the adenomatous stage and; iii) loss of chromosomal regions 17p and 18q that contained tumoral suppressor genes as tumor protein p53 (TP53) and deleted in colorectal carcinoma (DCC), respectively in the transition to carcinoma in situ.

Several authors have supported this hypothesis by study of primary colon tumors by using various techniques, including conventional cytogenetics, fluorescence in situ hybridization (FISH) (2), comparative genome hybridation (CGH) arrays (3, 4), and singles nucleotide polymorphism (SNP) arrays (5, 6). All of them identified the three essential common alterations described by Fearon and Vogelstein: APC and KRAS mutations and loss of chromosomal regions 17p and 18q; in addition others changes such as gain of chromosomes 7, 8q, 13q and 20q, together with loss of the 1p, 4,8p and 22q chromosomal regions (Figure 1) were also identified (4, 7-9).Despite knowledge regarding the genetic basis of sCRC genesis; currently, the genetic alterations involved in the metastatic process are not clear. In the last few years, there have been several studies focused on shedding light on the genetic aberrations that are involved in the metastasis of sCRC, especially regarding liver metastasis. For this purpose,
primary tumors and their paired liver metastases were simultaneously analyzed. Additionally, DNA microarrays, such as CGH or SNPs, allowed for massive screening of the whole genome; these arrays identify genetic alterations present in metastatic sCRC tumors (5, 9). First of all, it is important to note that the primary colonic tumors (of metastatic sCRC) presented characteristic copy number alterations; namely loss of 1p, 5q, 8p, 17p, 18q and 22q, and gain of chromosomal regions 8q, 13q and 20q (5, 6, 8). Bearing in mind these alterations, it should be noted that loss of 17p and 18q are believed to play an important role in the pathogenesis of sCRC since these two chromosomes carry genes relevant to the malignant transformation of the gut epithelium and also probably play an important role in the metastatic process. Studies have already suggested that abnormalities of chromosomes 17 and 18, especially loss of heterozygosity (LOH) of 17p and 18q (6, 10, 11), could be associated with more advanced stages of the disease; in this regard, recent findings showed that breakpoints in the 17p11.2 chromosomal region were preferentially found in primary colonic tumors in sCRC patients with liver metastases (6, 12). In addition, the study of liver metastases of sCRC, using the same SNP arrays, showed the presence of the same genetic alterations that were detected in primary tumors; however, the differences between primary tumors and their liver metastases were in the frequency of alterations. In this way, liver metastases exhibited a higher frequency of genetic alterations, as well as, newly-acquired alterations such as loss of 3p, 4 and 10q, and gain of chromosomal regions 5p, 6p or 7p (Figure 1) (2, 3, 6). These results support the hypothesis that these affected chromosomal regions codify tumoral suppressor genes that are involved in the metastatic process.
Regardless of the progress in the identification of the chromosomal aberrations involved in the metastatic process by SNP arrays, this technology presents with a significant limitation, since it only provides information about the alterations displayed in more than 60% of tumoral cells (13, 14), accordingly the information about the tumoral heterogeneity is lost. In line with these findings, our group was able to characterize different clonal evolution patterns in primary colonic tumor and their paired liver metastases using interphase FISH. It was established that secondary clones (alterations present in a low percentage of tumoral cells) found in the primary tumor were present in the liver metastasis at a higher frequency (2). It is suggest that the primary tumor may harbor a cell population with metastatic potential.

Currently, novel approaches based on DNA microarrays allow for identification of chromosomal alterations, especially copy number changes, in metastatic sCRC, however these genes that are responsible for this metastasis process have not been fully-characterized.

### Proteomic Alterations Involved in Liver Metastatic sCRC

The immense progress that has been made in the identification of genomic alterations involved in metastatic sCRC is not as advanced as that of genomics and molecular biology. To an extent, this is due to proteins stability and a high diversity in structure and abundance of proteins. In addition, post-translational modifications of protein (e.g. phosphorylation and acetylation) which affect protein structure as well as its functional capacity, further complicate the study and identification of proteomic alterations.

Most research has been focused on shedding light on the sequential protein alterations involved in the genesis of sCRC and in the metastatic process employing different proteomic techniques, like 2D-gel electrophoresis (15, 16), (DIGE) and chromatography (16), as well as labelling techniques (SILAC and iTRAQ) (17, 18) aiming to determine the protein fingerprint using mass spectrometry (MS). Overall, these studies identify proteins and quantify the expression levels compared with normal intestinal mucosa.

In this way, one of the first events found to be involved in the transformation from normal intestinal cells to cancer cells were described by Yu et al. (19) who defined a decrease in protein expression; mainly proteins involved in the cyto-protection process, RNA and polypeptide synthesis (such as ribonuclease 6), proteins of calmodulin complex. In the case of calmodulin complex, whose function is controlled by intracellular Ca²⁺ and plays a critical role in differentiation and/or apoptosis.

Whenever different researchers studied differential protein expression in metastatic sCRC, they concluded that proteins with aberrant expression are implicated in characteristic processes related to cancer such as cell proliferation and

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Name</th>
<th>Functions</th>
<th>Approaches</th>
<th>Biomarker value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breakpoints of 17p11.2</td>
<td>-</td>
<td>Unknown</td>
<td>iFISH, SNPs</td>
<td>Prognostic</td>
</tr>
<tr>
<td>BRAF</td>
<td>B-Raf proto-oncogene serine</td>
<td>DNA replication, recombination</td>
<td>PCR, iFISH, SISH, PCR</td>
<td>Predictive</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatidylinositol-3-kinase</td>
<td>Cell Cycle, cell death, cellular growth and proliferation</td>
<td>iFISH, immunohistochemistry</td>
<td>Prognostic, predictive</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
<td>Cell growth and proliferation</td>
<td>iFISH, PCR</td>
<td>Prognostic, predictive</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Phosphatidylinositol-4,5-bisphosphate 3-kinase</td>
<td>Cellular movement, cell death, apoptosis</td>
<td>PCR, immunohistochemistry</td>
<td>Prognostic, predictive</td>
</tr>
<tr>
<td>PBCA</td>
<td>Pancreatic beta cell, agenesis</td>
<td>Unknown</td>
<td>PCR</td>
<td>Predictive</td>
</tr>
<tr>
<td>ERCC1</td>
<td>DNA excision repair protein</td>
<td>DNA repair</td>
<td>Protein array, ELISA, western blot</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>MAPKAPK3</td>
<td>MAPK-activated protein kinase-3</td>
<td>Cell differentiation, gene expression</td>
<td>Protein array, ELISA, western blot</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>ACVR2B</td>
<td>Activin receptor type IIB</td>
<td>Cell differentiation</td>
<td>Protein array, ELISA, western blot</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>PIM-1</td>
<td>Proto-oncogene serine/threonine-protein kinase</td>
<td>Apoptosis, proliferation</td>
<td>Protein array, ELISA</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>FGFR4</td>
<td>Tyrosine kinase related to fibroblast growth factor receptor</td>
<td>Apoptosis, tumorogenesis</td>
<td>Protein array, ELISA</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>TFF3</td>
<td>Trefoil factor-3</td>
<td>Invasion of cells</td>
<td>Protein array, ELISA</td>
<td>Diagnostically</td>
</tr>
<tr>
<td>AGR2</td>
<td>Anterior gradient-2 homolog</td>
<td>Cell migration and proliferation</td>
<td>Protein array, ELISA</td>
<td>Diagnostically</td>
</tr>
</tbody>
</table>

---

**Table I. List of candidate biomarkers in liver metastatic sporadic colorectal cancer: gene or protein identification, functions, approaches used in determinations and biomarker value.**

- Yu et al.: Biomarkers in Metastatic sCRC (Review)

---

21
migration. These processes can be summarized as proliferation pathways, metabolic pathways, heat-shock responses, structural proteins and adhesion pathways (Figure 2).

As would be expected in metastatic sCRC, gains in the expression of proteins involved in proliferation pathways were identified; in this way, cell surface proteins acting as receptors for extracellular growth factors, such as syndecan-1 (CD138), with cell migration functions, were identified (20). In general, proteins involved in metabolic pathways were found to be over-expressed in metastases, such as fructose-1,6-biphosphate, being an indicator of the extent of cell proliferation. In contrast, the proteins that belong to cyto-protection metabolic pathways are down-regulated, for example phosphogluconate dehydrogenase (which is related to oxidative reactions). In addition, heat-shock proteins (HSPs) were found to be over-expressed in metastatic sCRC; proteins such as HSP70, HSP27 and HSP90 (20, 22, 23) were identified in primary colonic tumors, as well as liver metastases. Similar expression was found for multiple structural proteins, especially proteins related to the cytoskeleton, such as cytokeratins (CK7, CK9, CK17, CK18 and CK19) (19, 22, 23).

Although, several proteins related to adhesion pathways have been identified in metastatic sCRC, most of them are over-expressed. Proteins detected as being involved in adhesion processes using different pathways, including catenins, such as CDH17 (20), which plays a role in adhesion pathways by Ca\textsuperscript{2+} activation, whereas integrins such as ITGA2 or ITGA6 (20) have the same adhesion function but in this case, adhesion is mediated by collagen molecules. In summary, all the above protein alterations lead to the invasive process of sCRC and the development of liver metastases.

Proteomics tools have witnessed rapid growth over the past decade, with increasing emphasis on development of robust and high-throughput (HT) technologies to determine the dynamic proteomic changes derived from changes in cell state and function (24, 25). For HT studies, protein microarrays have been designed and implemented; hence, most of their applications are focused on the discovery of biomarkers with diagnostic and prognostic values.

Figure 2. Schematic representation displaying the summary of sequential protein alteration from normal intestinal tissue to metastatic colorectal cancer. Pathways with over-expression of proteins are represented with a dotted background and those with low expression of proteins are represented with a lined background.
Identification of Biomarkers in Metastatic sCRC

A biomarker is a biochemical molecule that can be measured objectively and be employed to detect pathology (diagnostic value), determine the disease evolution (prognostic value) or to choose treatment and determine its efficacy (predictive value). The techniques used to determine biomarkers must be reproducible, specific, sensitive, fast, simple and have a tolerable ratio of cost/efficacy. Currently, the clinical protocol in diagnostic metastatic sCRC is based on two techniques: i) image techniques such as ecography, computed tomography, scanner or endoscopy, which allow for the clinical determination of the existence of sCRC and liver metastasis (26); ii) the quantification of tumoral markers such as the well-established carcinoembryonic antigen (CEA) and carbohydrate antigen 19.9 (CA19.9) (27) by ELISA assays and the determination of fecal occult blood test (28). Although the existence of biomarkers that are used for their predictive value such as the detection of mutations of gene v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), by polymerase chain reaction (PCR) approaches (29) or the determinations of the expression of gene epidermal growth factor receptor (EGFR) by immuno-histochemistry assays the tumoral samples (30). Recently, research has focused on the discovery of biomarkers with higher efficacy in diagnostic and prognostic uses than the biomarkers described above, in addition, employing samples collected using non-invasive techniques, such as serum or plasma; for these reasons, DNA and protein microarrays are very useful tools.

Discovery of Biomarkers from Genomic Approaches

Currently, the use of DNA microarrays is not possible as routine analysis in clinical laboratories; for this reason, it is important to develop approaches which can be easily implemented in clinical laboratories, such as PCR and iFISH. Bearing in mind this requirement, breakpoints in the chromosomal region 17p11.2 could be used as a biomarker with prognostic value (12), employing a conventional cytogenetic approach (iFISH) for the detection of these breakpoints in sCR; such breakpoints indicate the presence of liver metastases and are associated with adverse overall survival. Using iFISH, phosphoinostide-3-kinase, catalytic, alpha polypeptide (PIK3CA) proto-oncogene was studied in sCRC (31, 32), the results of this study showed that gains and amplifications of this gene are observed as an early event of carcinogenesis in sCRC and are associated with better survival. Additionally, different mutations in tumoral suppressor gene phosphatase and tensin homolog (PTEN) (33, 34) have been used as prognostic biomarkers.

In addition, new studies are focused on the analysis of different genes that could be involved in the response to chemotherapy of patients with metastatic sCRC, which, therefore, could be used as predictive biomarkers. In this way, different research groups determined that mutations in the v-raf murine sarcoma viral oncogene homolog B1 (BRAF) gene, in cases with wild-type KRAS gene, could be involved in the resistance of treatment of metastatic sCRC (35-37). Other studies analyzed mutations in the PIK3CA gene (proto-oncogene) and their relation with treatment response (38). However, the determination of gene expression levels of pancreatic beta cell agenesis (PBCA), or excision repair cross-complementing rodent repair deficiency complementation group 1 (ERCCI) did not predict the efficacy of the oxaliplatin therapy (35). Recent studies of EGFR using iFISH approaches found that gains of this gene are related to anti-EGFR therapy response (30).

These novel and recently described genetic alterations could become novel biomarkers with clinical value; however, further studies in larger series of patients with metastatic sCRC patients are required.

Discovery of Biomarkers by Proteomic Approaches

The development of protein microarrays, as well as novel approaches for the identification of protein fingerprinting, has resulted in important progress in the identification of proteins, characteristic of metastatic sCRC present in serum or plasma. These technologies are very useful because the quantity of sample employed for these assays is minimal, in the order of microliters, solvents can be collected without using invasive techniques.

Based on their biomarker role, identified by genomic approaches, different investigations were focused on the detection of expression (using immunohistochemistry essentially) of PTEN and PIK3CA proteins (33, 34). The results of these assays determined low expression of PTEN to be related to a poor prognosis (33) and that a high expression of PIK3CA could be like a predictive biomarker for adjuvant chemo-radiotherapy (34).

Antibody arrays were used by Babel et al., to identify specific and unique proteins in the serum of patients with metastatic sCRC that could be used as diagnostic biomarkers (39). For this purpose, the serum profiles of healthy individuals versus patients with metastatic sCRC were compared and specific proteins were found in metastatic sCRC, including mitogen activated protein kinase activated protein kinase 3 (MAPKAPK3), activin A receptor II B (ACVR2B), Pim-1 oncogene (PIM-1), v-src sarcoma viral oncogene homolog (SRC) and fibroblast growth factor receptor-4 (FGFR4). The expression of all of these proteins was validated and confirmed as belonging to sCRC tumoral cells. Several scientists employing MS were able to identify proteins secreted by tumoral cells (secretome approach) in serum of patients with sCRC, such as trefoil factor-3 (TFF3), anterior gradient 2 homolog (AGR2) and lipocalin-2 (LCN2). As described above, these protein biomarkers need to be validated in larger series of patients with sCRC, before their implementation in clinical practice (40).
Future of Biomarkers for Metastatic sCRC

In summary, the use of approaches for massive screening, at the genomic and proteomic levels, are proved to offer advantages in the discovery of biomarkers (24, 25). In addition to microarray technology, current research projects are focused on determination of inherent biological characteristics of cancer evolution, as well as on the identifications of biomolecules generated by humans in response to tumoral cells, essentially biomolecules from specific activation of the immunological system. Bearing in mind these novel approaches, research is focused on the identification and quantification of circulating tumoral cells in the peripheral blood; such cells would appear to have an important role in the evolution of cancer and would be involved in the metastatic process (41). In this regard, the activation of the immune system towards tumoral cells, towards tumoral proteins, could be explained in the discovery of new biomarkers (42, 43). In line with this hypothesis, the identification of auto-antibodies against tumoral proteins in patients’ serum/plasma, using developing protein microarrays such as nucleoid programmable protein arrays (NAPPA) has demonstrated their efficiency in the screening of autoantibodies in different disorders (e.g. breast cancer) (44).

In summary, the introduction of biomarkers into clinical practice entails three steps: discovery, verification and validation. Personalized medicine will ultimately be based on a collection of specific biomarkers that are diagnostic and prognostic for metastatic sCRC and which can also, predict the response to treatment.

Acknowledgements

This work has been partially supported by grants from the Consejería de Sanidad, Junta de Castilla y León, Valladolid, Spain (SAN673/SA39/08 and SAN/103/2011), Fundación Memoria de Don Samuel Solórzano Barruso, Salamanca, Spain, Caja de Burgos (Obra Social), Burgos, Spain, Grupo Excelencia de Castilla y León (GR37) and the Red Temática de Investigación Cooperativa en Cáncer (RTICC) from the Instituto de Salud Carlos III (ISICIII), Ministerio de Sanidad y Consumo, Madrid, Spain (PI12/02053-FIS, PI11/02114 and RD060020/0035-FEDER). JM Sayagües and M González are supported by grants (CP05/00321 and FI08/00721, respectively) from the ISCIII, Ministerio de Ciencia e Innovación, Madrid, Spain.

Conflicts of Interest

None to declare.

References


Received May 20, 2012
Revised October 23, 2012
Accepted October 29, 2012