Molecular evolution of colorectal cancer: from multistep carcinogenesis to the big bang

Adriana Amaro1 · Silvana Chiara2 · Ulrich Pfeffer1

Published online: 7 March 2016

Abstract Colorectal cancer is characterized by exquisite genomic instability either in the form of microsatellite instability or chromosomal instability. Microsatellite instability is the result of mutation of mismatch repair genes or their silencing through promoter methylation as a consequence of the CpG island methylator phenotype. The molecular causes of chromosomal instability are less well characterized. Genomic instability and field cancerization lead to a high degree of intratumoral heterogeneity and determine the formation of cancer stem cells and epithelial–mesenchymal transition mediated by the TGF-β and APC pathways. Recent analyses using integrated genomics reveal different phases of colorectal cancer evolution. An initial phase of genomic instability that yields many clones with different mutations (big bang) is followed by an important, previously not detected phase of cancer evolution that consists in the stabilization of several clones and a relatively flat outgrowth. The big bang model can best explain the coexistence of several stable clones and is compatible with the fact that the analysis of the bulk of the primary tumor yields prognostic information.

Keywords Colorectal cancer · Carcinogenesis · Big bang · Tumor evolution · Tumor heterogeneity · Microsatellite instability · Chromosome instability

1 Introduction

With the present review, we intend to make a contribution to the ongoing discussion on molecular evolution of human cancers with reference to colorectal cancer (CRC). An exhaustive review of the literature on molecular genetics of CRC is therefore beyond the scope of the present article, and we refer to recent reviews [1–8].

CRC is the third most common cancer and the third leading cause of cancer death in men and women in the USA. For 2014, 71,830 men and 65,000 women were expected to be diagnosed with colorectal cancer and 26,270 men and 24,040 women were expected to die of the disease [9]. Incidence and mortality show a decline, although there is an increase in distal colon cancer in young adults [9]. Incidence and mortality can be reduced through preventative interventions such as colonoscopy [10, 11]. In fact, the decline of CRC incidence observed for Western countries is likely to be caused by increasing prevention including colonoscopy that can detect polyps and early adenomas before they develop into CRC [9]. The decline in the USA is contrasted by an increase in many economically transitioning countries especially in Eastern Europe and Asia [12].

The presenting symptoms of patients with colorectal cancer are variable, and tumors may be asymptomatic for long periods. Symptoms, which may result from obstruction, perforation, and bleeding, include abdominal pain, hematochezia, weight loss, anemia, change in bowel habit, and abdominal mass. Right-sided colonic lesions often cause vague epigastric pain and chronic blood loss leading to fatigue related to anemia, while patients with tumors in left colon often present with symptoms of obstruction, constipation, or a change in the caliber of stools. Cancer arising in the recto sigmoid region is associated with tenesmus and hematochezia [13].
According to the International Documentation System of Colorectal Cancer, carcinomas that have a lower border of the tumor 16 cm or less from the anal verge are classified as rectal carcinoma. Using this definition, about 50% of all colorectal carcinoma are located in the rectum, 25% in the sigmoid colon, and 25% in the remaining part of the colon. However, the incidence of right-sided or proximal colon cancer has been increasing in North America and Europe [14]. This anatomic shift may be likely due to the following: increased longevity, variation of response to luminal carcinogens between different sites of the colon and the rectum, genetic factors resulting in microsatellite instability (MSI, see below) in proximal colon cancer, and chromosomal instability (CIN, see below) pathway predominant in left-sided colon and rectum. The location of CRC has an impact on response to chemotherapy and disease-specific survival [15, 16]. No molecular differences in terms of somatic mutation frequencies could be detected among microsatellite-stable (non-MSI) CRC in relation to the location in the colon or the rectum [17].

In approximately 80% of cases, colorectal carcinoma arises from adenomatous polyps. There is a typical transition from hyper-proliferative epithelium to focally dysplastic crypts, to macroscopically evident tubular adenoma, to progressively dysplastic and or villous adenoma, to invasive cancer [18]; however, this transition is slow and less than one in ten adenomas progresses to carcinoma [19]. Based on extrapolations of the mutation loads, progression from adenoma to carcinoma has been estimated to take 17 years [20]. In 1988, Vogelstein et al. presented evidence for a sequential acquisition of genomic alterations, i.e., somatic mutations and chromosomal aberrations, in CRC [21] in accordance to the multistep model of carcinogenesis [22, 23]. The majority of CRCs are characterized by aneuploidy or CIN, and approximately 15% of CRCs arise due to mutations in DNA repair genes or to widespread hypermethylation in gene promoter regions (CpG island methylator phenotype, CIMP), both leading to extensive instability in simple repeated nucleotide sequences (microsatellite instability, MSI). Convincing evidence shows that serrated polyps, in particular sessile serrated adenoma, are the precursor to MSI and CIMP cancers. These precursor serrated lesions are usually large and show high levels of CIMP and BRAF mutations and occur preferentially in the proximal colon [1].

Epidemiologic studies have implicated a number of environmental cofactors in the development of colorectal cancer, including advanced age, a diet high in red meat, a diet rich in fat, smoking, alcohol consumption, and obesity [24]. Risk factors have been identified in approximately 30% of patients: personal or family history of colorectal adenoma or colorectal cancer, ulcerative colitis, and Crohn’s colitis. Patients with the highest risk for colorectal cancer are those who have hereditary colorectal cancer syndromes: familiar adenomatous polyposis (FAP), Lynch syndrome (hereditary nonpolyposis colorectal cancer, HNPCC) and MYH-associated polyposis (MAP) [25, 26].

After the initial mucosal growth, the tumor protrudes first into the lumen and progresses in lateral transverse direction, leading to circumferential growth. Mural penetration may result in local failure, peritoneal seeding, or both. The probability of regional lymph node metastasis depends primarily on histological features as depth of invasion, histological grade, and absence or presence of lymphatic invasion. The liver is the primary site of hematogenous metastases, followed by the lung. The major venous drainage of the lower rectum occurs by a dual system: drainage from the superior hemorrhoidal veins enters the portal system to the liver, whereas drainage from the middle and inferior hemorrhoidal veins eventually reaches the vena cava to get to the lung. Bone metastases in the sacrum and the vertebral bodies occur through the vertebral plexus. Peritoneal seeding is a frequent result of transmural penetration and tumor shedding [27, 28].

Several factors have been identified to have an impact on recurrence and survival, but TNM staging according to the American Joint Committee on Cancer (AJCC)/Union Internationale Contre le Cancer (UICC) staging system [29] still has the greatest prognostic significance. The TNM system classifies tumors on the basis of the invasiveness of the primary tumor (T stage), the number of loco-regional lymph nodes containing metastatic cancer (N stage), and the presence or absence of distant metastatic disease (M stage). The TNM categories reflect similar survival outcomes for rectal and colon cancer, and therefore these diseases share the same staging system [30].

2 CRC heterogeneity

Intratumoral heterogeneity (IHT) is well documented for CRC [31, 32]. Tumor heterogeneity has important consequences for targeted therapies [32–34] and is expected to drastically affect the potential to prognosticate tumor progression and metastasis and to predict therapy response [35]. IHT consists in generation and establishment of molecularly distinct subclones that carry genomic or epigenomic alterations not present in the bulk of the tumor or other subclones. Intuitively, genomic instability, a hallmark of many cancers and especially of CRC, is expected to generate pronounced tumor heterogeneity yet not all clones might survive and eventually one or a few particularly “fit” subclones might outcompete all the others. However, the differences between subclones can be limited to molecular alterations that do not confer any selective advantage. It is therefore important to distinguish between “driver mutations” that drive tumorigenesis since they confer a growth or survival advantage and “passenger mutations” that do not affect the fitness of the clone. Some passenger mutations may influence tumor development or even have
deleterious effects on cancer growth [36]. The distinction between the passenger and driver mutations is not straightforward. In general, the presence in many tumors of the same type (incidence) and the presence in a large portion of a single tumor (penetrance) are considered characteristics of driver mutations. Evidence of selection in terms of maintenance of the mutation in subsequent steps of tumorigenesis as for example in metasases could be considered an additional criteria, but selection may occur as a consequence of linkage to the driver mutation rather than fitness effects [37]. A typical cancer has approximately eight driver mutations that act through cellular signaling pathways to influence cell fate, cell survival, and genome maintenance [38].

Tumor heterogeneity is particularly pronounced in CRC since almost all of these cancers are caused by genomic instability in the form of either chromosome (CIN) [39] or microsatellite (MSI) instability [40, 41]. In Lynch syndrome (formerly hereditary nonpolyposis CRC) [42], cancer development is driven by the mutation of one of several mismatch repair (MMR) genes (MLH1, MSH2, PMS2, and MSH6) [43]. Mutation of MMR genes or polymerase ε (POLE) is also observed for sporadic cancers [17], and MSI is also observed in sporadic cancers with bi-allelic methylation of MLH1 [44] in part due to CIMP [45] that is largely overlapping with MSI [17, 46]. Given the complexity of mechanisms surveilling genomic integrity, no single molecular cause for CIN has been identified [47], but mutations in mitotic checkpoint genes, microtubule spindle defects, and telomere dysfunction certainly can lead to CIN [48] and conservative replication of chromosomal double strand breaks has recently been added to the potential causes of CIN [49]. It is not clear whether there are triple negative CRC's (CIN-, MSI-, CIMP-) [46]. Based on evidence of somatic CAG repeat shortening in the androgen receptor gene in non-MSI CRC early during cancer development, an independent mechanism for the somatic alteration of CAG repeat length in microsatellite stable colon cancer has also been proposed [50]. Clustered mutations (kataegis, localized hypermutation) caused by upregulation of the expression of DNA cytidine deaminase, APOBEC3B, also occur though with a low frequency in CRC and no apparent correlation to CIN or MSI [51]. Clustered chromosomal rearrangements in localized and confined genomic regions (chromothripsis) have been described for most CRCs [52], but its relation to CIN is unclear [53].

Mismatch repair gene mutations do not themselves cause cancer; rather, they generate a mutator phenotype [54, 55] that causes the accumulation of mutations among those that might contribute to cancer progression as, for example, the mutations of the BCL2-associated X protein (BAX) [56], activin receptor type 2 (ACVR2A) [57] transforming growth factor receptor B2 (TGFBR2) [58], solute carrier family 22 (organic anion transporter) member 9 (SLC22A9) [59], small ArfGAP 1 (SMAP1) [59], and olfactory receptor, family 7, subfamily E, member 24 (OR7E24) [59]. Importantly, many of such slipping frameshift mutations that are typical for MSI may escape detection by several massive parallel sequencing technologies since homopolymeric regions frequently yield erroneous sequencing results and are therefore filtered out [60]. Chromosomal instability leads to highly frequent deletions, amplifications, inversions, and translocations among those that hit cancer-promoting genes such as ERRB2, IGF2 amplifications, and TCF7L1 translocation and fusion to NAV2 [17, 61]. MSI and CIN may also occur in combination [62]. CRC and other cancers that are characterized by MSI or CIN are therefore intrinsically heterogeneous even if the initial cause, the acquisition of the mutator phenotype, is homogeneous. Non-MSI and non-CIN cancers, which are probably rare in CRC, are expected to be driven by conventional driver mutations such as commonly observed mutations in signaling pathways and are expected to show tumor heterogeneity similar to other solid cancers. Aneuploidy is a hallmark of cancer [63] and plays an important role in most cancers by generating the template on which Darwinian selection can act [64]. There is no clear distinction between aneuploidy and CIN with the latter corresponding to the accelerated acquisition of chromosomal alterations leading to aneuploidy. Aneuploidy but not necessarily CIN is involved in most cancers [65]; MSI, for unknown reasons, is not. In other words, the primary mutations in MSI and CIN cancers are not driver mutations in the sense that they directly act on cell growth control mechanisms but they create a mutator phenotype that makes the acquisition of secondary mutations in growth control genes highly likely.

The degree to which colon cancers show genomic instability has prognostic value [66]. CIN confers a poor and MSI a favorable prognosis [67, 68]. Increasing numbers of mutations are expected to increase the likelihood of cancer-promoting mutations, but this might be counterbalanced by also increasing the frequency of mutations in essential genes that limit cell viability, and CIN and MSI differ for the total mutation load they provoke. MSI but not CIN CRCs show hypermutation [17]. Hypermutation also implies increased frequency of mutations that affect viability; therefore, strong MSI might limit tumor progression due to frequent mutations in essential genes [36]. CIN yields amplifications of ERBB2 and IGF2 and fusion of NAV2 and TCF7L1 that are positively selected [17] and apparently compensate for the occurrence of non-viable genome rearrangements in other cells caused by CIN.

In line with what has been observed for other tumor types, the correlation between somatic mutations and the gene expression phenotype of CRCs is not very strong [17], although gene and protein expression show significant differences when comparing CIN and MSI CRC. Three [69, 70] or five [71] genomic and five proteomic subtypes [72] have been identified. Several authors have attempted to delineate CRC pathways in accordance with the original multistep carcinogenesis concept [73]. These pathways most likely are
oversimplifications, but certain molecular events tend to occur in sequence with preferential combinations thereof. The classical pathway consists in mutations in the adenomatous polyposis coli (APC) detectable in early adenoma followed by KRAS, PIK3CA, SMAD4, and TP53 mutations and loss of heterozygosity of chromosomes 18q and 17p that lead the way to intermediate and late adenoma. This pathway leads to carcinoma with CIN. The serrated pathway starts with BRAF mutations in hyperplastic polyps followed by CIMP in sessile serrated adenoma. CIMP leads to MLH1 methylation driving dysplasia and finally leading to carcinoma with the MSI phenotype [7, 46] (Fig. 1).

Mutations in signaling pathways most likely predetermine the molecular evolution of the cancer since a complete switch to another pathway is unlikely. The tumor remains addicted to the pathway in which the transformation event occurred [74]. In this case, the molecular analysis of the primary tumor yields prognostic information since the tumor follows a route that is predetermined by the pathway initially chosen. Eventually, selective pressure, as it is, for example, exerted by drugs that target the etiological pathway, especially in the presence of huge amounts of cancer cells in late stage cancers, makes the emergence of clones with the activation of alternative pathways more likely [34, 75]. But even in this case, the repertoire of pathways that can be activated is probably not unlimited.

MSI and CIN cancers are expected to be more complex since several pathways can be affected leading to a high heterogeneity in these groups of cancers. Yet, even in the single cancer in a single patient, more than one pathway can be activated contemporaneously or in sequence through subsequent MSI or CIN events. However, the comparison of MSI CRC with MSI endometrial cancer shows different mutation spectra [59, 76] that at least in part rely on the fact that different genes are “available” for mutation likely dependent on the positioning of nucleosome that can protect the DNA [59].

Hence, oncogene addiction is expected to be less pronounced in cancers with an instable genome, but this has not been studied systematically. Oncogene addiction has been proposed for EGFR amplification in CRC [77], yet EGFR overexpression frequently occurs in MSI CRCs and is associated with somatic mutation of an adenosine repeat in the 3′-UTR of the EGFR gene [78].

3 Stem cells in CRC

The intriguing stem cell theory of cancer development predicts that only a small part of the tumor is made up of cells that contain the potential to generate a tumor [79]. Such cancer initiating or cancer stem cells have also been detected in CRC [80–82]. DNA methylation analyses are consistent with the presence of several stem cell lineages within a single CRC [83]. This makes a fundamental difference for cancer therapy since killing of the bulk of tumor cells might be completely

Fig. 1 Main routes of colorectal cancer evolution. Two main routes of colorectal cancer evolution leading respectively to chromosomal instability (CIN, left) and microsatellite instability (MSI, right) can be distinguished. The classical pathway of adenoma evolving into carcinoma is characterized by prevalent mutations in APC and KRAS and a series of less frequent mutations and chromosomal rearrangements (CIN). The serrated pathway shows prevalent mutation of BRAF and mismatch repair-associated mutations or gene silencing typical of the CpG island methylator phenotype (CIMP). CIN is positively and MSI inversely associated with the risk of developing metastases. Tumor heterogeneity increases during cancer evolution (but see also Fig. 4)
inefficacious if the tumor stem cells are left unhurt [84]. Yet, the stem cell theory most likely has only minor effects on prognostication since it is expected that a tumor stem cell with specific molecular alterations will yield bulk tumor cells that are molecularly distinct from those generated by a stem cell with different molecular lesions. The bulk tumor cells can therefore be taken as a marker for the stem cell that has generated them. The driving forces, MSI, CIN, and pathway mutations, are expected to occur in stem cells and to be propagated in bulk tumor cells [85]. Mutator mutations are expected to raise the probability of generation of stem-like cancer cells through mutations that establish a self-propagating potential. Frequent mutations of the TGFBR2 [58] and the alternative type 2 TGFβ receptor, ACVR2A [86], TGFβ expression [87, 88] in MSI CRC, and the association of genetic variants of TGFβ related genes with CRC risk [89] indicate the TGFβ pathway as an important contributor to CRC carcinogenesis, most likely through its role in epithelial–mesenchymal transition (EMT) [90–92], although expression of functional TGFBR2 in cancers with biallelic microsatellite mutation in exon 4 of the TGFBR2 gene, likely as a result of transcriptional slippage, has been described [93]. EMT also occurs for CIN CRC where TGFβ signaling is disrupted by genomic deletion of the chromosome region 18q21.1 containing the downstream targets of TGFβ receptors, SMAD2 and SMAD4 [94, 95], or by mutations of SMAD 2, 3, and 4 [96]. Activation of the β-catenin/Wnt pathway as a consequence of APC mutations can also lead to EMT [97]. APC mutations have been shown to occur in CRC stem cells [98]. It appears that the two main CRC evolution branches, CIN and MSI, with APC and TGFBR2 mutations are linked to EMT.

**Fig. 2** Cancer stem cells and epithelial–mesenchymal transition. Cancer stem cells are most likely characterized by mutations of TGF-β receptor 2 or ACVR2A in the microsatellite instable cancers and by loss of heterozygosity of chromosome 18q containing SMAD genes or by mutation of SMAD genes. APC mutations are more frequent in CIN than in MSI cancers. These mutations determine disruption of the TGF-β and β-catenin pathways, and the altered intracellular signaling causes epithelial to mesenchymal transition (EMT). Together, the stem cell and EMT phenotypes determine tumor progression and metastasis. EMT and cancer stem cells are likely to be interrelated.

### 4 CRC evolution models

The molecular evolution of cancer during the first steps cannot be observed since it occurs in apparently healthy subjects long before a cancer diagnosis. Genomics can, however, deliver information on cancer evolution through the analysis of primary tumors and metastases obtained from the same patient. In the evolution of the tumor from the primary lesion to a distant metastasis, many molecular alterations, mutations of any kind, and gene expression events are expected to occur, especially in the case of a genetically unstable tumor. The central question is which, if any, of these alterations are needed for the tumor to become metastatic and when it has been acquired.

Several different scenarios of tumor progression towards metastasis are imaginable: (i) tumor progression follows a route that is predetermined by molecular characteristics such as driver mutations, the profile of transcriptional activity, the cell type of origin of the primary tumor, or a combination of these; (ii) after transformation, the initial tumor undergoes a limited phase of high genomic instability followed by stabilization of some clones that evolve linearly (big bang model, see below); (iii) cooperative evolution where clones with...
Cancer evolution models. Cancer evolves from atypia to overt primary tumors that invade the surrounding tissue and eventually metastasize to lymph nodes and distant tissues where micrometastases can grow to life-threatening metastases. Metastasis can be limited by the inability to form metastasis or by the inability of disseminated cells to grow in the target tissues. The models make different statements on the possibility to predict distant metastasis by analyzing the primary tumor. 

a) Linear evolution: The initial driver mutation determines the metastasis risk, and additional mutations have limited effects on the metastasis risk. The tumor fate is determined from the very beginning and can be predicted.

b) Big bang model: After transformation, the tumor has a phase of genomic instability yielding clones with various molecular lesions. Some clones stabilize and grow to form a highly heterogeneous primary cancer with a linear evolution towards metastasis (flat subclone growth).

c) Cooperative evolution: Different subpopulations of the primary cancer with various molecular lesions are needed to accomplish metastasization.

d) Clonal selection: The classical tumor evolution model predicts the selection of subclones present in the primary tumor that have acquired additional mutation(s). Successive steps of carcinogenesis require additional mutations that confer selective advantages. This model would allow for prediction of the metastasis risk only through analyses at the single cell level.

e) Metastasis-specific mutations: Initial dissemination is determined by the mutation(s) present in the primary cancer that acquires additional mutation(s) after dissemination that allows the growth of micrometastases into clinically overt metastatic masses. In this case, the metastasis risk cannot be predicted by the analysis of the primary cancer.
different mutations cooperate in forming metastases; (iv) a single cell clone with one or several additional mutations is positively selected and progresses towards metastasis independently of the molecular make-up of the bulk of the tumor; and (v) the metastatic phenotype is acquired only after dissemination through additional mutations not present in the primary tumor (see Fig. 3). These three models are not necessarily mutually exclusive: the driver mutation in the primary tumor might determine the route to metastasis that requires, however, additional mutations not present in the primary tumor or present in only a tiny subpopulation, yet the likelihood of these secondary events might vary with different driver mutations. Cooperative models where cells with different mutations cooperate in at least one step of metastasization have also been reported [102], but despite the high grade of intratumoral heterogeneity typical for CRC, cooperation of subpopulations in CRC has not been addressed.

Deep sequencing of human cancers reveals mutation patterns consistent with a model of branched evolution of the tumors [103, 104]. The difficulty relies in establishing whether a mutation that gives rise to a branch is a mutation that alters the progression potential of the branch (a secondary driver mutation) or not or, in other words, to which extent the fate of the tumor is determined by molecular characteristics present in the bulk of the primary tumor. Tissues that are exposed repeatedly and/or for long time intervals to carcinogenic insults are expected to accumulate mutations in many cells of the tissue (field cancerization) from which eventually single clones arise with the ability to form a tumor [105, 106]. The colonic tissue is exposed to a variety of dietary carcinogens of different chemical nature that induce different molecular lesions. An analysis of the methylation pattern-based lineage tracing of adenomatous crypts as a marker of clonal evolution indicates elevated clonal stability over the long time of adenoma growth with most clones generated at relatively early times of adenoma development [82]. Similar observations have been made for advanced CRC [83] consistent with a model of “flat” clonal expansions. Genomic profiling of many glands form 15 CRCs also indicates that most mutations, even those detected in only a minor fraction of the tumor (private mutations), are acquired early during carcinogenesis and expand slowly during tumor growth to generate a homogenous distribution of clonal population over the whole tumor [107]. This “big bang” model postulates that numerous heterogeneous subclones generated early during tumor development grow as a single expansion. Further, “private” mutations due to replication errors accumulate continuously. Alterations arising at later stages are only present in small regions of the tumor [107] (Fig. 4). Strong evidence in favor of the big bang model has been gathered through massive genomic characterization of serially sampled CRCs. Sottoriva and colleagues analyzed 349 individual glands form 15 CRCs by whole-genome array-based profiling of copy number alterations, whole-exome sequencing, targeted deep sequencing, FISH, and neutral methylation tag sequencing [107]. This study shows that most of the IHT derives from early private alterations and not through clonal selection. The determination of the metastatic potential of the tumor also appears to occur early during cancer development. The central point of these analyses is that the initial genetic instability that generates many clones is overcome in
some clones that regain stability and grow to form the advanced tumor and its metastases. The big bang model thus incorporates the branched evolution model by highlighting the timing of the branching events.

This model is also consistent with other molecular analyses of matched CRC primary lesions and metastases [108–110]. Xie and colleagues performed a whole-genome sequencing analysis of the primary tumors and matched metastases of two CRC patients revealing a complex pattern of molecular alterations that were only partially overlapping between primary lesions and the metastases [109]. The mutational analysis for 189 genes in 13 samples of primary CRC and matched metastases revealed an overall concordance rate of 78%. The exclusion of mutations that were observed only rarely (potential passenger mutations) raised the concordance to 90%. Different primary tumors showed distinct mutation profiles among which distinct KRAS mutations on codons 12 and 13 in one patient and independent KRAS and NRAS mutation in the two primaries of the other patient. Only one of the four primaries of the first patient had an APC mutation indicating that the primary lesions might belong to different molecular classes. The matched metastases showed a mutation profile related to one of the primary lesions, yet the absence of some of the mutations indicated a subpopulation as the likely progenitor of the metastases [108]. Unfortunately, the two latter studies analyzed a very limited set of patient samples using panel re-sequencing rather than whole-exome sequencing which would have allowed for a more global assessment of intratumoral and intertumoral heterogeneity.

Interestingly, in the study of Kogita and coworkers [108], the hepatic metastases analyzed showed no mutation of the primary tumor that was present in all the cells of the metastasis. The highest prevalence was observed for KRAS reaching 56.9% in a metastasis [108]. These observations difficultly fit into a linear evolution model. Metastases composed of cells that resemble various clones of the primary tumor, such as KRAS wild type and KRAS mutated clones, could be derived from subpopulations of the primary cancer that did not carry the specific mutation (KRAS wild type) and have acquired the mutation independently at the metastatic site. Tumor self-seeding [111] or cooperative metastasis would be alternatives. In order to rule out these possibilities, larger studies with whole-exome or whole-genome sequencing are needed.

All genomic alterations, CIN [112], MSI [113, 114], and somatic mutations [115] can be observed in very early stages of CRC carcinogenesis in early adenoma or the putative precursor lesions and non-dysplastic and dysplastic aberrant crypt foci [116] even in the absence of concurrent cancer [113]. However, the ratio of lesions that show a specific mutation increases during CRC carcinogenesis [73]. We have already proposed that a multistep carcinogenesis, dissemination of cancer cells, and colonization of metastasis target tissues may occur from the very beginning of carcinogenesis [117] as it appears from the most thorough analysis so far performed [107]. Under this hypothesis, additional genomic alterations could only marginally influence the risk of metastatic growth, and the primary tumor, long before it becomes clinically overt, could contain all the determinants of metastasis. This would also solve the contradiction between the predictions of the multistep carcinogenesis model and the fact that gene expression profiling of the bulk of the primary tumor can prognosticate the risk to develop metastases [70, 118]. The carcinogenic steps occur early during tumor evolution and are mediated by a pronounced genetic instability of the tumor that is followed by stabilization of several clones that develop into the clinically overt cancer.

5 Conclusions

- CRC shows a clear morphological evolution from aberrant crypt foci over adenoma to invasive carcinoma and metastasis, but the molecular correlate of this evolution is not linear. Somatic mutations, CIN, and MSI are present early during tumor development. The contribution of additional mutations during tumor evolution is not clear but probably limited. IHT is also generated early during carcinogenesis, and clonal populations grow in a relatively stable manner. IHT is at the basis of therapy resistance.
- The location of the primary tumor has prognostic and therapeutic importance but does not correspond to clearly distinct molecularly defined subtypes.
- The big bang model best explains the mutation patterns observed in CRC. The most prevalent mutations are acquired early followed by a flat evolution with a limited expansion of later on acquired private mutations.
- The primary tumor therefore already possesses most of the determinants of tumor evolution consistent with the possibility to predict outcome by analyzing the primary tumor.
- Instability of the genome is most likely limited to early phases of tumor evolution followed by re-stabilization.
- Re-stabilization of initially genetically stable tumors is likely to be an important hallmark of cancer and might constitute the bottleneck in the development of metastatic cancers.
- The most prevalent molecular alterations, CIN, MSI, and somatic mutations converge on the TGFβ and WNT pathways generating epithelial mesenchymal transition with interrelations with cancer stem cells.

Acknowledgments The present work has been supported by a grant from the Italian Ministry of Health “5 x mille” program.
References


