



Published in final edited form as:

Gastroenterology. 2015 January ; 148(1): 10–13. doi:10.1053/j.gastro.2014.11.024.

Molecular Subtyping of Colorectal Cancer - Time to Explore Both Intertumoral and Intratumoral Heterogeneity to Evaluate Patient Outcome

Eric R. Fearon, M.D., Ph.D.^{1,3} and John M. Carethers, M.D.^{2,3}

¹Divisions of Molecular Medicine and Genetics, University of Michigan, Ann Arbor, MI

²Division of Gastroenterology, University of Michigan, Ann Arbor, MI

³Department of Internal Medicine, and Comprehensive Cancer Center, University of Michigan, Ann Arbor, MI

Over the past three decades since molecular analyses of somatic gene alterations in primary human cancer specimens first became tractable, many of the recurrent somatic genetic and epigenetic defects present in colorectal cancers (CRCs) have been identified. The accumulation of multiple loss-of-function defects in selected tumor suppressor genes and gain-of-function defects in selected oncogenes, together with epigenetic alterations, such as DNA methylation changes, is believed to be critical in initiating colorectal tumorigenesis and the progression of dysplastic precursors to invasive and ultimately metastatic lesions (1–4). Among the most common tumor suppressor gene mutations in CRCs are those in the *APC* (adenomatous polyposis coli) and *TP53* genes (2–4). The most common oncogene mutations in CRC include point mutations activating the functions of the *KRAS*, *PI3KCA* (phosphoinositide-3-kinase, catalytic, alpha polypeptide), *BRAF*, and *NRAS* proteins (2–4). The oncogene missense mutations found in the *KRAS* and *NRAS* genes in about 40–45% of CRCs most commonly affect codons 12 and 13, but a subset of CRCs have *KRAS* or *NRAS* codon 61 missense mutations (2, 3). Substitution of glutamic acid for the wild type valine at codon 600 (V600E) accounts for the vast majority of *BRAF* activating mutations in CRC (2, 3). Mutations activating *KRAS* or *NRAS* are mutually exclusive with *BRAF* activating mutations (2–4).

Recent comprehensive sequencing studies suggest that about only 25 different genes are commonly affected by somatic mutations in CRCs, with tumor suppressor genes outnumbering oncogenes on this list by about 4 to 1 (3, 4). About 16% of CRCs manifest a hypermutation phenotype, with a median number of 700 subtle somatic mutations predicted to alter protein products (3). About three-fourths of the hypermutation CRC cases – roughly 12–13% of all CRCs - are constituted by the CRCs that manifest the high frequency of microsatellite instability (MSI-H) phenotype, due to mutation or inactivation of one of several different key proteins functioning in DNA mismatch repair (MMR), most

Correspondence: Eric R. Fearon, M.D., Ph.D., Department of Internal Medicine, University of Michigan, 1504 A. Alfred Taubman Basic Science Research Building, 109 Zina Pitcher Place, Ann Arbor, MI, 48109-2200, Tel: 734-764-1549, fearon@med.umich.edu.

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest are disclosed.

prominently including MLH1, MLH3, and MSH2 (3). The remaining quarter of the hypermutation CRC cases – about 3–4% of all CRCs – do not manifest the MSI-H phenotype and usually harbor somatic mutations in the gene encoding DNA repair polymerase POL ϵ or one or more mismatch repair genes (3). In the remaining 84% of colorectal cancers that do not manifest the hypermutation phenotype [including both the so-called microsatellite stable (MSS) and microsatellite instability low (MSI-L) cases], the median number of subtle somatic mutations in exons predicted to alter protein products is about 60 mutations per tumor (3). Prior work has established that the MSS and MSI-L CRCs often display aneuploidy, albeit with certain recurrent chromosome and sub-chromosome gains and losses seen in considerable fractions of CRCs (2).

In mammalian genomes, DNA methylation covalently modifies the cytosine residue in the majority of 5'-CpG-3' dinucleotide sequences, except for CpG islands, which are localized regions of high CpG content often found in the promoter and upstream regulatory regions of a large fraction of genes (5). Hypermethylation of these CpG islands is associated with gene silencing (5). A subset of colorectal cancers shows extensive DNA hypermethylation at many different CpG islands scattered around the genome and this phenotype has been termed the high frequency CpG island hypermethylation phenotype (CIMP-H) (6, 7). About 20–25% of CRCs manifest the CIMP-H phenotype and a similar fraction of CRCs manifest a lower frequency CIMP (CIMP-L) phenotype, with the remaining 50% of CRCs lacking CIMP (8). Many CIMP-H and a few CIMP-L CRCs have hypermethylation of the promoter region of the *MLH1* MMR gene, and this group of CRCs constitutes the majority of the MSI-H CRCs (7, 8). Of note, the hypermutation CRC subset, including many of the CIMP-H CRCs displaying MSI-H, is almost invariably the subset of CRCs that harbors the *BRAF*^{V600E} oncogenic mutation (7, 8). The overwhelming majority of these CRCs with *BRAF*^{V600E} mutations arise in the right colon (7, 8).

The specific tumor suppressor and oncogenic somatic mutations as well as the chromosome and sub-chromosomal copy number changes and the epigenetic alterations that have critical roles in promoting the outgrowth and/or sustaining the survival of neoplastic cells have been termed “driver” gene lesions. Those gene lesions that do not contribute in a functionally significant fashion to the origin and/or persistence and expansion of the cancer cell population, but may have instead arisen as bystander events during tumorigenesis, are often termed “passenger” gene lesions. Many CRCs have multiple somatic driver gene mutations, often including one or more oncogene mutations together with several different tumor suppressor gene mutations (1–4, 9). Not unexpectedly, in light of the many different possible combinations arising simply from consideration of the commonly mutated oncogenes and tumor suppressor genes and the common chromosome and subchromosomal gains and losses, coupled with the very high number of patient-specific somatic mutations seen in any given CRC, essentially no two CRCs share the same somatic mutation profiles in the bulk of the cancer cells. Moreover, this extensive genetic complexity in a given CRC co-exists with similarly extensive epigenetic complexity in the CRC.

In spite of the tremendous complexity and diversity seen in CRC genomes and epigenomes, prior strategies for attempting to define molecular alterations in CRC that may have utility for likelihood of cancer recurrence and outcome and/or response to conventional or novel

therapeutic agents in patients have not infrequently emphasized efforts to implicate single oncogene or tumor suppressor gene defects or quite limited combinations of gene defects as prognostic or predictive markers. To date, while most single gene mutation markers have modest prognostic or predictive value, *BRAF*^{V600E} mutation has been associated with poorer survival in CRC, and there are convincing data to indicate that the presence of a *KRAS* mutation in a CRC are associated with resistance to therapies targeting the epidermal growth factor receptor (EGFR) (10). Arguably, the most robust prognostic molecular classifier for CRC has been the MSI-H phenotype. The MSI-H phenotype has been linked to improved survival in stage II and stage III CRC patients (11), though the use of 5-fluorouracil (5-FU)-based adjuvant chemotherapy did not appear to show any survival benefit in survival for patients with MSI-H CRC (12–14).

The papers from Sinicrope and colleagues and Phipps and colleagues in this issue of *Gastroenterology* advance some new and important associations between molecular alterations and patient survival (15, 16). The analyses stem in part from prior work from Jeremy Jass that suggested determination of MSI and CIMP status might be very useful for defining distinct histopathological and molecular subsets of CRC, and the distinct subsets of CRC might have significant variations in the prevalence of somatic mutations affecting the *APC* and *TP53* tumor suppressor genes and the *KRAS* and *BRAF* oncogenes (17). In his work on the use of MSI- and CIMP-status to define CRC subsets, Jass also emphasized the view that the key precursor lesions for sporadic CRCs manifesting the CIMP-H phenotype were most likely to manifest a serrated morphology, including sessile serrated adenomas and traditional serrated adenomas (17). In their studies of very large numbers of CRC patients, both the Sinicrope and Phipps groups emphasized the analysis of MSI and CIMP status in the CRC specimens, as well as analysis for mutations of codons 12 and 13 of *KRAS* and codon 600 of *BRAF* (15, 16). Both groups found, with remarkably similar percentages for subgroups, that analysis of MSI-status (MMR-function) and CIMP-status along with *KRAS* and *BRAF* mutations were informative for colorectal cancer-specific mortality (Table 1). Both groups also reported that MMR-proficient CRCs with *KRAS* mutations or especially *BRAF* mutations had poorer outcome than MMR-proficient tumors that were wild type for either gene (Table 1). As the authors appropriately note, the findings offer further evidence that studies of the inter-tumor molecular heterogeneity of CRCs has merit and value.

In spite of the utility of defining molecular phenotypes based on clonal genetic alterations shared by all or nearly all neoplastic cells in a patient's primary CRC, it seems increasingly likely that in-depth and comprehensive analyses of intra-tumoral molecular heterogeneity in each patient's primary CRC may have major ramifications for understanding how molecular defects may contribute individually and collectively to clinical outcomes in CRC patients. Comprehensive sequence-based analyses of cancer cell populations from individual patients where the cancer cell populations were spatially and/or temporally distinct have indicated that significant intra-tumoral genetic heterogeneity in primary cancer lesions and metastases may be the "rule" in cancer, rather than an exception (9, 18). Critical initiating genetic and epigenetic lesions might be shared (i.e., clonally) among all neoplastic clones in a primary cancer, but geographically distinct regions of primary tumors may have distinct mutation and epigenetic profiles from those in other regions of the tumor. Similarly, metastatic cell

populations in a patient may have genetic and epigenetic divergence from the non-metastatic cells, and there may be heterogeneity among different metastatic lesions in a patient and within individual metastases (9, 18). Branched evolutionary growth may be an important feature in both primary tumors and metastatic lesions, with multiple competing clonal populations evolving over time and in space. This more recent view of the potentially quite extensive intra-tumoral genetic heterogeneity in any given cancer and the contributions of intra-tumoral heterogeneity to tumor progression contrasts with some earlier views. Prior to the recent studies (9, 18), it was suspected that the cell populations in many primary cancers might be more homogeneous, where somatic mutations were accumulated in a more stepwise fashion as a result of multiple sequential clonal sweeps of each variant cell populations in the primary cancer, with metastases perhaps often arising from the clonally dominant cell population in the primary tumor. Consistent with the notion that rare variant cell populations may have important roles in clinical outcome, a recent study indicated that in some CRC patients where the primary cancer cell population was wild type for *KRAS*, there is strong biological selection for outgrowth of CRCs with mutations in *KRAS* or other mitogen-activated protein kinase pathway proteins when EGFR blockade is used therapeutically (19, 20). New strategies, such as future deep-sequencing of primary CRC cell populations, comprehensive single cell analyses, and/or analyses of circulating tumor-derived DNA will likely be needed for future molecular approaches to better define prognosis and predict likely responses to existing and new targeted therapies in patients with CRC.

Acknowledgements

This work was supported by the United States Public Health Service (DK067287, CA162147 and CA082223) and the A. Alfred Taubman Medical Research Institute of the University of Michigan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990; 61:759–767. [PubMed: 2188735]
2. Fearon ER. Molecular genetics of colorectal cancer. *Ann Rev Pathol*. 2011; 6:479–507. [PubMed: 21090969]
3. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012; 487:330–337. [PubMed: 22810696]
4. Sehagiri S, Stawiski EW, Durinck S, et al. Recurrent R-spondin fusions in colon cancer. *Nature*. 2012; 488:660–664. [PubMed: 22895193]
5. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev*. 2011; 25:1010–1022. [PubMed: 21576262]
6. Toyota M, Ahuja N, Ohe-Toyota M, et al. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci USA*. 1999; 96:8681–8686. [PubMed: 10411935]
7. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006; 38:787–793. [PubMed: 16804544]
8. Hinoue T, Weisenberger DJ, Lange CPE, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res*. 2012; 22:271–282. [PubMed: 21659424]
9. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science*. 2013; 339:1546–1558. [PubMed: 23539594]

10. Herzig DO, Tsikitis VL. Molecular markers for colon diagnosis, prognosis and targeted therapy. *J Surg Oncol*. 2014 [Epub ahead of print].
11. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol*. 2005; 23:609–618. [PubMed: 15659508]
12. Carethers JM, Smith EJ, Behling CA, et al. Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. *Gastroenterology*. 2004; 126:394–401. [PubMed: 14762775]
13. Kim GP, Colangelo LH, Wieand HS, et al. Prognostic and predictive roles of high-degree microsatellite instability in colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project collaborative study. *J Clin Oncol*. 2007; 25:767–772. [PubMed: 17228023]
14. 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. [PubMed: 20498393]
15. Phipps AI, Limburg PJ, Baron JA, et al. Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology*. 2015 (**this issue**).
16. Sinicrope FA, Shi Q, Smyrk TC, et al. Molecular markers identify prognostically distinct subtypes of stage III colon cancer patients. *Gastroenterology*. 2015 (**this issue**).
17. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology*. 2007; 50:113–130. [PubMed: 17204026]
18. Hiley C, de Bruin EC, McGranahan N, et al. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. *Genome Biol*. 2014; 15:453–462. [PubMed: 25222836]
19. Diaz LA Jr, Williams RT, Wu J, et al. The molecular evolution of acquired resistance to targeted EGFR blockage in colorectal cancers. *Nature*. 2012; 486:537–540. [PubMed: 22722843]
20. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early and late-stage human malignancies. *Science Transl Med*. 2014; 6:224ra24.

Summary of genotypes, their percentages, and relative patient outcome from Phipps *et al.* and Sinicrope *et al.* among primary colorectal cancers.

Table 1

Cancer Genotype	MSS	MSS/ <i>BRAF</i> ^{V600E} mutant	MSS/ <i>KRAS</i> mutant	MSI-H/ <i>BRAF</i> ^{V600E} mutant or <i>hMLHI</i> hypermethylation	MSI-H without <i>BRAF</i> ^{V600E} mutant or <i>hMLHI</i> hypermethylation
Approximate Phenotype	Traditional/CIN	Serrated/CIMP	Alternate	Serrated/Sporadic CIMP/MSI	Familial MSI
Phipps, <i>et al.</i> N=2080 (% of cancers)	(Type 4) 47%	(Type 2) 4%	(Type 3) 26%	(Type 1) 7%	(Type 5) 4%
Sinicrope, <i>et al.</i> N=2720 (% of cancers)	49%	6.9%	35%	6.8%	2.6%
Outcome	Referent	Poor 1.5-2X worse	Poor 1.5X worse	Favorable	Favorable 0.3X
Comments	More likely distal	More likely proximal, female, older age	Slightly more proximal; high representation among African Americans	More likely proximal, female, older age	More likely proximal, young age

MSS=microsatellite stable, MSI-H=microsatellite instability-high, CIN=chromosomal instability, CIMP=CpG island methylator phenotype