Exploring tumor heterogeneity
Fessler, E.

Link to publication

Citation for published version (APA):
Fessler, E. (2016). Exploring tumor heterogeneity

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Part II – Inter-tumor heterogeneity

Chapter 4

Identification of colorectal cancer subtypes.

Evelyn Fessler and Jan Paul Medema.

Submitted.
Abstract
Cancer is a heterogeneous disease and many cancer types do not represent a single entity, but are in fact composed of biologically and clinically diverse subtypes. The subtype affiliation frequently dictates prognosis as well as response to therapy. It is therefore of utmost importance to consider this so-called inter-tumor heterogeneity in clinical management of patients and in preclinical studies. Recently, a consensus molecular classification system has been devised for colorectal cancer, which will be of great benefit for future fundamental and clinical research as it enables uniform categorization of colorectal cancer specimen across different institutions and studies. The biological conformity observed within each consensus molecular subtype holds great promise for the development of subtype-specific treatment regimens. Herein, we describe the classification of colorectal cancer patients, with a focus on the consensus molecular subtypes, and we discuss suitable model systems that will be crucial for in-depth biological characterization of distinct subtypes and design of novel therapeutic strategies.
Introduction

Colorectal cancer (CRC) patients show diversity with respect to prognosis and response to therapy; yet, the highly heterogeneous nature of this malignancy is not reflected in the current CRC staging system (1). Although histopathological parameters like resection margin, tumor differentiation grade, and invasion of lymph and blood vessels play a role (2), the most important factor for prognostication and therapeutic decision making is tumor stage (1). The American Joint Committee on Cancer devised a staging system, which is composed of the T (local tumor invasion), the N (involvement of lymph nodes), and the M (metastases in distant organs) stage (1). This staging system, however, does not allow prediction of prognosis or benefit from chemotherapy on the level of individual patients, which remains a holy grail in clinical practice. Accordingly, a large body of literature is available on how to further classify CRC patients into clinically relevant groups. Nonetheless, to date it remains a challenge to predict treatment response for the early stages of CRC (3). This might in part be due to the fact that predictive signatures were derived from whole patient sets without taking the presence of biologically distinct subtypes into account. Recent advances in derivation and analysis of gene expression profiles make it possible to stratify patients with a given cancer type into biologically homogeneous subgroups. Considering this inter-tumor heterogeneity and classifying CRCs into distinct groups prior to biological and clinical analyses might result in detailed characterization and development of more successful treatment avenues. Herein, we discuss the modalities used for identification of CRC subtypes and compare classical and modern methods of CRC classification, with an emphasis on the potential of recently developed approaches, such as unbiased gene expression-based stratification.

Identification of colorectal cancer subtypes

Molecular markers. Historically, CRC has been classified based on molecular features of the genome into a chromosomally instable and a microsatellite instable group (2). Chromosomal instability (CIN) is present in approximately two thirds of CRC cases and refers to tumors with aneuploid chromosome sets carrying multiple structural and numerical aberrations (2). Although the causes of CIN are not well-defined, it has been described that patients with CIN⁺ tumors suffer a poorer clinical outcome compared to patients with chromosomally stable tumors (4, 5). Yet, this clinical association might be influenced by the microsatellite status of the tumor: CIN and microsatellite instability (MSI) rarely coincide and patients bearing MSI⁺ tumors display better prognosis (6). Therefore, the poor prognosis associated with CIN could be a result of microsatellite stability concurring in these tumors. MSI⁺ tumors are characterized by a near-
diploid genome and instability in form of insertions and deletions in microsatellite regions (2). MSI is caused by a defective mismatch repair system, in the sporadic form most frequently due to inactivation of one of its main components, MLH1, by promoter hypermethylation (7). The limitation associated with MSI as clinical marker results from its low frequency – only ~15% of all CRCs are MSI+. Besides its prognostic value, MSI status also holds information about treatment response, as patients with MSI-high CRC appear to respond poorly to fluorouracil-based therapy (6, 8).

Next to instability afflicting the genome, global changes can also occur on the epigenetic level. This manifests either as an overall DNA hypomethylation or as the so-called CpG island methylator phenotype (CIMP). CIMP is defined as hypermethylation of CpG islands, which are frequently located in promoter regions of genes. Hence, CIMP can lead to epigenetic silencing of tumor suppressor genes such as CDKN2A (9). CIMP has also been shown to affect the promoter region of the MLH1 gene, leading to inactivation of the mismatch repair system and thus to MSI, which rationalizes the overlap of MSI and CIMP found in CRC (10).

In addition to genome-wide aberrations, single genetic alterations – such as mutations in the KRAS and BRAF oncogenes – present important clinical parameters. The KRAS mutation status, for instance, holds predictive value for patients with metastatic disease that would be eligible for anti-epidermal growth factor receptor (EGFR) therapy (11, 12). KRAS-mutant tumors do not respond to this targeted approach due to activation of the mitogen-activated protein kinase (MAPK) pathway downstream of the EGFR. Of note, the type of mutation afflicting the KRAS gene seems to influence treatment response: retrospective analyses revealed that patients with KRASG13D-mutant CRCs benefit from treatment with the anti-EGFR monoclonal antibody Cetuximab in contrast to tumors displaying other KRAS mutations (13, 14). In vitro studies underlined this observation in showing that CRC cell lines carrying the KRASG13D mutation are sensitive to EGFR inhibition (13, 15). However, the conclusion that anti-EGFR therapy provides benefit for this particular subset of CRC patients cannot yet be drawn and warrants results from prospective trials (16). The use of the KRAS mutation status as predictive marker is further complicated by the fact that only ~35% of non-KRAS-mutant tumors respond to this treatment option, which could be explained by the presence of mutations in other downstream signaling components of the MAPK pathway – such as BRAF or NRAS – and by mutations in PIK3CA activating a parallel pathway (17). This suggests a benefit from combining information about multiple molecular markers, which has been shown to allow more refined classification of patients and more accurate prognostication (18-20). This approach also revealed that the prognostic value of individual mutations relies on the presence or absence of other molecular
markers – the $BRAF^{V600E}$ mutation, for example, is only associated with poor prognosis in tumors that are not microsatellite unstable (19).

Even though progression of CRC has been well-defined on the genetic level (21) and whole mutational landscapes have been drafted for this malignancy (22, 23), defining the mutational spectrum of a cancer does not always reveal all its characteristics. Recently, it was, for instance, shown that a striking overlap in gene expression exists between tumors carrying a $BRAF^{V600E}$ mutation and a subset of tumors – dubbed $BRAF$mutant-like – which lack this genetic event, but display the same gene expression patterns and overlapping clinical features (24). This analogy could result from mutational activation of the same signaling pathway and could predict similar response to targeted therapy (25), propelling whole pathways rather than single mutations into the spotlight (26). In line with this is the observation that poor response to anti-EGFR therapy can be observed in tumors characterized by a gene expression signature indicative of active MAPK or PI3K signaling pathways. Surprisingly, also samples lacking mutations in above-mentioned oncogenes can fall into this category (27), highlighting the need for assessing pathway activity next to individual mutations.

Importantly, also on protein level single markers can carry prognostic information, as was shown for the expression of caudal-type homeobox transcription factor 2 (CDX2) (28-31). CRCs expressing no CDX2 protein display significantly worse clinical outcome and seem to benefit from adjuvant therapy (31). The predictive validity of this approach, however, awaits further testing in appropriately-designed clinical trials. Furthermore, in accordance with results on the genetic level, the information provided by individual markers does not permit to draw conclusions about underlying biology and thus hampers the identification of novel targets and treatment avenues. In this respect, gene expression-based studies might aid to obtain a more comprehensive picture and to group tumors based on the activity and complexity of biological programs.

**Gene expression-based classification.** The list of publications on gene expression-based categorization of CRC patients is too extensive to be discussed herein. Of note are two gene expression-based signatures that were successfully translated into the clinic to guide prognostication for stage II patients: the ColoPrint and Oncotype DX signatures (32, 33). One major drawback of these two signatures is the supervised approach: in case of the ColoPrint signature, survival was considered as an a priori variable and the genes assessed in the Oncotype DX assay have previously been associated with disease recurrence. Many more signatures have been proposed for the identification of high risk CRC patients; however, small sample sizes
frequently result in lack of statistical significance. Moreover, absence of pre-stratification – e.g. patients with MSI+ tumors might have to be excluded due to better prognosis – makes many studies difficult to interpret.

Unsupervised approaches allow unbiased classification of cancer patients into distinct groups yielding valuable biological insight (34). Each cancer subgroup identified this way can subsequently be analyzed individually, allowing the detection of underlying driver mutations or pathways – information that can be used to design subtype-specific prediction signatures and therapies. This strategy has led to promising treatment options for breast cancer (35).

Unbiased classification approaches have been applied to distinct CRC data sets and several studies report the existence of 3-6 CRC subtypes (29, 36-42). In an effort to generate a unified classification system for CRC, the international CRC subtyping consortium was formed (43).

Indeed, of a total of 3,962 CRC samples 3,443 are classifiable into one of four main consensus molecular subtypes (CMSs) and each of the subtypes associates with distinct biological programs (Figure 1) (43). CMS1 tumors display a diffuse immune infiltrate and are frequently associated with MSI and CIMP, and the \( \text{BRAF}^{V600E} \) mutation, therefore this group was dubbed the MSI, immune subtype of CRC. The canonical CMS2 is characterized by high expression of targets of the WNT pathway and the transcription factor MYC, as well as an epithelial signature. This group was recognized as the classical type of CRC due to high levels of CIN and activated WNT signaling, suggesting it to follow the canonical path of CRC development (21). Based on gene expression, CMS3 tumors also possess a strong epithelial component and additionally show deregulation of metabolic processes, thus representing the metabolic subtype. On the molecular level these tumors frequently display mutations in the \( \text{KRAS} \) oncogene and are generally CIMP low. The fourth group is characterized by a large fraction of stromal cells and the activation of genes associated with epithelial-mesenchymal transition (EMT), angiogenesis, transforming growth factor-\( \beta \) (TGF\( \beta \)) signaling, and matrix remodeling. These processes have been associated with poor disease outcome before (44, 45), and indeed, patients with CMS4 tumors show worse relapse-free and overall survival (43). Importantly, the consensus classification reveals signaling pathways that are selectively active in CMS1 (JAK-STAT), CMS2 (WNT and SRC), and CMS4 (integrin-\( \beta3 \), TGF\( \beta \), and VEGF/VEGFR), further highlighting the biological foundation of the subtyping approach and pointing to the potential of developing subtype-specific targeting regimens (43).

Intriguingly, an overlap with subtypes in other solid cancers can be observed, suggesting that deregulation of similar pathways is a common theme with a comparable phenotypic outcome in malignancies of distinct organs. Lessons might thus be learned from detecting analogies between
Identification of colorectal cancer subtypes.

tumors of the same subtype arising in different organs and from extrapolating biological characteristics and treatment response. Parallels can for instance be drawn between CMS3 of CRC and a metabolic, genomically stable group of gastric cancers (46, 47). Since no specific pathways were identified for CMS3 of CRC, findings in gastric cancer might help to further define this particular CRC subtype. A universal trait of cancer seems to be the development of a mesenchymal subtype that is frequently associated with poor prognosis, as such a subgroup was identified in a large panel of solid cancers, such as CRC (CMS4) (43), glioblastoma (48), ovarian cancer (49), and gastric cancer (46). Interestingly, Hoadley and colleagues recently described the

---

Figure 1. The consensus molecular subtypes of CRC as described by Guinney et al. 2015 (43). Each subtype presents a biologically and clinically distinct subgroup. Precursor lesions are assigned to the CMSs based on gene expression (Fessler et al. 2016 (59)). Active pathways and programs are linked with arrows based on their potential connection. CAFs – cancer-associated fibroblasts; CIMP – CpG island methylator phenotype; CIN – chromosomal instability; DFS – disease-free survival; EMT – epithelial-mesenchymal transition; MSI – microsatellite instability.
identification of 11 ‘integrated subtypes’ across 12 cancer types revealing a striking conformity between gene expression in cancers derived from different organs. Incorporating this subtype information in the current prognostication scheme improves survival prediction, highlighting the clinical significance of this approach (50). This observation indeed indicates that biological traits are shared between distinct tumor types and suggests the existence of oncogenic driver pathways whose aberrant activation installs similar traits in tumors originating in different tissues.

**Model systems.** The biological conformity observed within cancer subtypes holds promise for the design of subtype-specific treatment strategies. This may be facilitated by the fact that cell line panels can be classified into the same subtypes as primary CRCs and thus permit subtype-specific screening modalities (36). Making use of this approach, it has, for instance, been revealed that cell lines belonging to the mesenchymal colon cancer subtype are more resistant to the anti-EGFR antibody Cetuximab (29). Importantly, this finding is mirrored in the clinic: patients with mesenchymal colon cancers do not seem to benefit from Cetuximab treatment irrespective of their KRAS mutation status (29). Cell line panels are an invaluable tool for high throughput screening approaches (51-53); yet, their advantages – such as easy propagation and extensive annotation – are counterbalanced by numerous drawbacks – for instance adaptation to culture conditions during continuous propagation. Recent advance in ex vivo culture systems nowadays allows the generation of so-called organoid cultures from patient-derived specimen. Organoids are defined as self-organizing systems that contain both stem cells and their progeny and thus constitute an in vitro model of either the healthy organ (in case of normal tissue used for isolation) or the patient tumor (54). CRC organoids display the same mutations and copy number variations in similar frequencies as CRC specimen, implying that collections of organoid cultures will reflect the same heterogeneity observed on the population level (55). The generation of organoid biobanks might soon permit therapeutic subtype screening in these ex vivo culture systems with the goal of designing personalized treatment strategies (55). Moreover, tissue specimen obtained from patient tumors can also be transplanted into immunodeficient mice to form so-called patient-derived xenografts. This strategy has successfully been used to determine the drug sensitivity of patient-derived tumor tissue (56-58).

**Outlook**
It is becoming increasingly clear that subtypes of a given type of cancer are as heterogeneous as unrelated malignancies arising in different organs (3). Each subtype should therefore be considered as distinct entity and characterized as such. Combining information from patient material with data derived from model systems will greatly facilitate this endeavor by allowing
the identification of subtype-specific drivers and underlying regulatory mechanisms. Subtype-specific vulnerabilities identified this way can be exploited for the design of tailored treatment regimens. While this approach has been successfully employed in breast cancer (35), it remains to be investigated whether devising treatments specific for CRC subtypes will be similarly beneficial. To understand effectiveness or failure of specific agents, we believe that it is of utmost importance to not only consider the current state of the tumor by analyzing the end stage carcinomatous tissue derived from patients, but to complement this knowledge with information from early stage disease. We think that studying the developmental route of distinct subtypes will shed light on the causes of subtype-specific peculiarities. Furthermore, we advocate the use of integrative approaches which acknowledge that each cancer subtype is in fact a complex mélange of non-transformed and malignant cells. At first glance, this detailed biological characterization of each subtype seems like an unsurmountable challenge, yet, sophisticated model systems will keep on being developed that will eventually succeed in pinpointing subtype-specific weaknesses.

Acknowledgements
This work was supported by an AMC Graduate School PhD Scholarship (E.F.). The authors would like to thank Louis Vermeulen for useful comments.

Conflict of interest
The authors declare no conflict of interest.

References
5. T. Watanabe et al., Chromosomal instability (CIN) phenotype, CIN high or CIN low, predicts survival for colorectal cancer. J Clin Oncol 30, 2256-2264 (2012).


37. B. Perez-Villamil et al., Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior. *BMC Cancer* 12, (2012).


Identification of colorectal cancer subtypes.