Supplementary Methods:

Pathologist Scoring

Scoring was performed by three trained observers (KT, MJ, GO), blinded from automated classification results. Cores were eligible for scoring when at least 50 percent of the core was intact. Individual cores were scored for both intensity and percentage of KER. Intensity was scored semi-quantitatively as 0 = absent, 1 = weak, 2 = moderate and 3 = strong. The percentage of KER was scored as the percentage of positive cells compared to negative cells in the whole core, to identify the epithelial load of the cores. In this epithelial load CDX2, FRMD6, HTR2B and ZEB1 were scored. For CDX2 and FRMD6 the percentage of intra-tumoral epithelial positive cells were evaluated. For CDX2 and HTR2B the intensity of intra-tumoral epithelial positive cells was scored semi-quantitative. For ZEB1 intra-tumor epithelial positivity was evaluated and scored as either present or absent. In the AMC-AJCCII-90 set, triplicates of each tumor specimens were available. Cores were scored on an individual basis and afterwards a majority consensus was applied to define scoring at a patient level.

Supplementary Tables

Table S1. Description of features used in the quantitative classifier

|  |  |
| --- | --- |
| Feature | Description |
| StainArea | Number of pixels in a core classified as ‘positive’ |
| StainInt | Average intensity of pixels in a core classified as ‘positive’ |
| StAreaFrac | Number of pixels in a core classified as ‘positive’/Number of pixels in a core |
| Brown.total | StainInt x StAreaFrac |
| StAreaFrac.norm | StAreaFrac/ KeratinStAreaFrac |
| Brown.total.norm | Brown.total / Keratin.Brown.total |

Table S2. Multivariate survival analysis of the CAIRO and CAIRO2 cohorts with respect to treatment arm

|  |
| --- |
| **CAIRO2: KRAS/BRAF wt tumors** |
|  |  | **CMS2/3 (epithelial)** | **CMS4 (mesenchymal)** |
| **Clinical Factor** |   | N (n event) | HR | CI | P value | N (n event) | HR | CI | P value |
| Age |  | 91 (62) | 1.01 | 0.98-1.03 | 0.48 | 68 (56) | 1.02 | 0.98-1.06 | 0.274 |
| Sex | F | 35 (18) | 1 |  |  | 27 (14) | 1 |  |  |
|   | M | 56 (41) | 1.19 | 0.70-2.03 | 0.51 | 41 (34) | 0.79 | 0.45-1.40 | 0.42 |
| Treatment | Control | 39 (30) | 1 |  |  | 34 (24) | 1 |  |  |
|   | Cetuximab | 52 (32) | 0.52 | 0.31-0.86 | 0.0108 | 34 (32) | 1.56 | 0.91-2.65 | 0.10 |
| Log rank Test |  |  |  |  | 0.049 |  |  |  | 0.11 |
| **CAIRO2: KRAS/BRAF mutant tumors** |
|  |  | **CMS2/3 (epithelial)** | **CMS4 (mesenchymal)** |
| **Clinical Factor** |   | N (n event) | HR | CI | P value | N (n event) | HR | CI | P value |
| Age |  | 57 (52) | 1.02 |  | 0.21 | 88 (75) | 1.01 | 0.99-1.04 | 0.32 |
| Sex | F | 24 (17) | 1 |  |  | 41 (27) | 1 |  |  |
|   | M | 33 (31) | 0.93 | 0.51-1.69 | 0.81 | 47 (40) | 1.13 | 0.69-1.84 | 0.62 |
| Treatment | Control | 27 (24) | 1 |  |  | 44 (35) | 1 |  |  |
|   | Cetuximab | 30 (28) | 1.57 | 0.86-2.85 | 0.142 | 44 (40) | 1.75 | 1.08-2.84 | 0.022 |
| Log rank Test |  |  |  |  | 0.34 |  |  |  | 0.064 |
| **CAIRO** |
|  |  | **CMS2/3 (epithelial)** | **CMS4 (mesenchymal)** |
| **Clinical Factor** |   | N (n event) | HR | CI | P value | N (n event) | HR | CI | P value |
| Age |  | 253 (120) | 0.99 | 0.98-1.01 | 0.65 | 154 (141) | 0.99 | 0.98-1.02 | 0.74 |
| Sex | F | 91 (79) | 1 |  |  | 52 (45) | 1 |  |  |
|   | M | 161 (141) | 1.12 | 0.84-1.48 | 0.43 | 102 (96) | 1.33 | 0.93-1.91 | 0.13 |
| Treatment | Sequential | 124 (111) | 1 |  |  | 77 (75) | 1 |  |  |
|   | Combination | 129 (109) | 0.91 | 0.69-1.19 | 0.48 | 77 (66) | 0.71 | 0.51-0.99 | 0.042 |
| Log rank test |  |  |  |  | 0.78 |  |  |  | 0.09 |

**Supplementary Figures**

**Fig. S1. Patient selection and omission**

**(A)** Flowchart of patient selection criteria.Omitted patients are highlighted in yellow, patients which were included in the final cohort are indicated in blue. ArmA in the CAIRO cohort refers to sequential chemotherapy with capecitabine, oxaliplatin and irinotecan. ArmB in the CAIRO series refers to combination therapy. CAIRO2 ArmA refers to treatment with capecitabine, oxaliplatin and bevacizumab. Arm B refers to capecitabine, oxaliplatin, bevacizumab with adjuvant cetuximab. \*MSI+ patients in the AMC-AJCCII-90 series were removed from this step. MSI+ mesenchymal-like patients were used to train the IHC classifier. (**B**) Distribution of clinical variables in each cohort prior to and after patient selection for the study

**Fig. S2. Overview of CMS classification scheme**

(**A**) AMC-AJCCII-90 dataset with Colon Cancer Subtypes previously determined by gene expression is used to train a TMA IHC-based classifier. This is then used to classify cores from the LUMC, CAIRO and CAIRO2 patient cohorts as epithelial-like or mesenchymal-like. MSI patients are identified by microsatellite instability (**B**) Image Analysis Pipeline: The main TMA area is detected in each image. Color deconvolution is then applied to separate the haematoxylin and DAB channels, of which the latter is quantified. Results are normalized with epithelial keratin staining (**C**) IHC data from different cohorts is then normalized by rescaling features to a [0, 1] range. A PCA plot shows the overall distribution of features before and after rescaling.

**(D)** Classification using a random forest classification algorithm that employs a collection of decision trees using the staining derived variables (**E**) Cores are mapped to patients. In the case of conflicting core classification, the majority class is used.

**Fig. S3. Overview of web-based classifier**

Pathologist-based classifier is available online at crcclassifier.shinyapps.io/appTesting/. Semi-quantitative scoring of intensity and content are required as inputs. Examples of scoring for each stain are available in separate tabs.

**Fig. S4. Pathologist scoring of TMAs**

**(A)** Association between tumour differentiation reported by pathologist scoring and staining of CDX2 and ZEB1.

**(B)** Associations between pathologist scoring of stains and CMS subtypes.

**(C)** Intra-class correlation as a measurement of consistency of inter-pathologist scoring for each stain

**(D)** Prediction errors of individual features using 100 iterations of 3 fold-cross validation. The error distribution of the combined classifier is highlighted in red.

**(E)** Keratin staining of regions containing tumor budding and corresponding CDX2 staining.

**Fig. S5. Comparison of CMS-classification to the serrated adenoma pathway**

**(A)** Comparison of the Leggett and Whitehall (LW) classification scheme described by CIMP, MSI, KRAS and BRAF status to the CMS classification.

**(B)** Contingency table of CMS with LW and the corresponding survival analysis using the LW classification system in the AMC-AJCII-90 cohort

**Fig. S6. Treatment arms stratified by colon cancer subtype in the CAIRO cohort**