**SUPPLEMENTARY MATERIALS AND METHODS**

*BRAF V600E mutant patient clustering*

The non-negative matrix factorization was achieved as follows: the few negative values present in the dataset were zeroed to obtain a non-negative matrix, and the NMF was performed for ranks 2 to 6 using the Brunet algorithm. Two hundred runs were performed per rank with random seeding. For each rank, the cophenetic correlation coefficient and silhouette score were computed from the 200 solutions and the factorization achieving the minimum residual error was kept as the solution for further analyses. The cophenetic correlation coefficients and silhouette scores were directly obtained using the *plot* function of the *NMF* package. As a control, the NMF was also performed on a randomized dataset built by taking within each patient a random permutation of the gene expression values, using the *randomize* function of the *NMF* package. After having selected rank 2 as the best one, the NMF was rerun only for rank 2 and for a superior number of runs (1500 runs).

The MSI-adjusted clustering analysis appearing in **supplementary Figure S2B** was performed as described as follows: for each gene, the MSI effect was estimated using a linear model and the effect is subtracted in the MSI group, essentially the average difference between the MSI and the MSS samples is canceled out. The resulting MSI-adjusted data matrix is then used in NMF clustering and the new subgroup calls were compared to the original calls.

*Molecular pathway and biological process analysis*

Pathway analyses were performed by Gene Set Enrichment Analysis (GSEA) and single sample GSEA (ssGSEA) using the GenePattern web interface from the Broad Institute (genepattern.broadinstitute.org/). Individual gene signatures and collections were selected from the MSigDB portal (http://www.broadinstitute.org/gsea/msigdb), we used the Hallmark collection. Enrichment scores from ssGSEA were subjected to differential analysis using the regularized linear model as implemented in the *limma* package (version 3.24.4) (1). Unless otherwise mentioned, the best ranked signatures that were highly significantly different between BM1 and BM2 were kept for plotting (corresponding to a BH-FDR of < 10-11). Enrichment scores were subjected to hierarchical clustering, and signature heatmaps were achieved using the *heatmap.2* function of the *gplots* package (version 2.17.0). For ssGSEA comparisons between WT2, *KRAS* mutant and BM subtypes we randomly selected an equal number of WT2 and *KRAS* mutant patients as in the smaller groups (BM1 and BM2) in order to avoid asymmetric group sizes. Consequently, for comparison of WT2 with BM1 and BM2 we randomly selected 69 and 149 WT2 patients out of 2010 patients (417 from Agendia, 56 from GSE35896, 461 from GSE39582, 227 from Melbourne, 660 from PETACC3 and 189 from TCGA), respectively. Similarly for comparison of *KRAS* mutants with BM1 and BM2 we randomly selected 69 and 149 WT2 patients out of 848 patients (143 from Agendia, 29 from GSE35896, 217 from GSE39582, 94 from Melbourne, 258 from PETACC3 and 107 from TCGA), respectively. The immunological landscape of BM subtypes was assessed using the CIBERSORT tool (2) and the heatmap of macrophage-related genes was done using the macrophage gene signature described by Bindea *et al*. (3).

To remove some ambiguities in the results (e.g. mTOR signatures), we investigated the biological process enrichment of these signatures by GeneTerm Linker analysis which is a new algorithm for functional annotation of gene lists (gtlinker.cnb.csic.es) (4). To do so, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation as reference (minimum support = 4).

*BM classifier and genes specific for BM subtypes*

We constructed a classifier for the two BM groups by training a penalized logistic model with the subtype labels given by the selected two subgroups NMF solution, using the *glmnet* function from the *glmnet* package (version 2.0.2). The model maximized a L1 (lasso) penalized likelihood. One hundred repetitions of a leave-one-out cross-validation were performed, the biggest penalty coefficient corresponding to the minimum of the misclassification error was kept in each run; these values were averaged to compute the penalty coefficient used to train the final model. We obtained a classifier with non-zero coefficients for 44 genes listed in the supplementary Table S2.

*Survival analysis*

Univariable and multivariable analyses of relapse-free survival (RFS) and overall survival (OS) were performed using Cox proportional hazard regression models available in the *survival* package (version 2.38.1). Hazard ratios (HR) were estimated with model coefficients and 95% confidence intervals (95% CI) and *p*-values were computed with Wald tests. Survival analyses were performed using the cohort as stratification parameter. The survival data for the different cohorts were provided to this study in the way they were originally defined. Time to event were determined as the time from randomization for PETACC-3, from the time of surgery for Agendia and GSE39582, and from the time of diagnosis for TCGA. Relapse events in RFS analyses were metastasis (local, regional or distant) and death, with the exception of PETACC-3 that also included secondary primary tumors as event. In the Agendia cohort, patients dying during the three months following surgery were excluded from the analyses.

In the gene expression profiling of PETACC-3 (5), the stage II patients were enriched for patient with relapses and consequently for shorter RFS and OS. Thus we have performed the survival analysis once with and once without PETACC-3 stage II patients. The results were almost identical and we report here the results with the PETACC-3 stage II patients. Survival curves were computed with the Kaplan-Meier method. Figure 1 shows which data were used for survival analysis; supplementary Figure S9 summarizes the clinical characteristics of the patients included in the survival analysis, the median censoring times and the number of events in each cohort and for each endpoint.

*Methylation and proteomic analysis*

Level 3 methylation data were obtained from the TCGA portal (https://tcga-data.nci.nih.gov/tcga/). They originate from two different platforms, either JHU-USC HumanMethylation27 (Illumina 27k) or JHU-USC HumanMethylation450 (Illumina 450k). Methylation -values (signal of probes for methylation divided by the sum of signals for methylated and unmethylated cytosines) were used as provided. Differential analysis of -values was achieved using the *limma* package. Positions with an FDR inferior to 0.05 were considered as differentially methylated. Proteomic and phosphoproteomic reverse phase protein array (RPPA) data for TCGA *BRAF* mutant patients were obtained from The Cancer Proteome Atlas (TCPA) platform (http://app1. bioinformatics.mdanderson.org/tcpa/\_design/basic/index.html) (6).

*Drug response prediction on cell lines public data*

Cell lines gene expression profiles and drug screening data were obtained from two public sources: i) the Sanger Institute (http://cancer.sanger.ac.uk/cell\_lines and http://www.cancerrxgene.org/translation/CellLine). The origin and methodology for these data appears here: http://www.cancerrxgene.org/help/; ii) and the Cancer Cell Line Encyclopedia (CCLE) (http://www.broadinstitute.org/ccle) (7). Drug response analyses were done using the data provided by these sources, the values for the area under the curve (AUC) of fitted dose-response curves for the Sanger data and the values of the area above fitted dose-response curves, the inverse measure, for the CCLE. Differential analyses and heatmaps were performed as mentioned in the main text.

**SUPPLEMENTARY FIGURE LEGENDS**

**Supplementary Figure S1**. Clustering of colorectal cancer BRAF V600E mutant patients.

**A**: Principal component analysis of *BRAF* mutant patients. The four first principal components which explain the most of data variation are shown. Patients are labeled according to the cohort to which they belong. **B**: Hierarchical clustering of *BRAF* mutant patients using the Ward method to compute distance between patients.

**Supplementary Figure S2**. Subtyping of CRC *BRAF* mutant patients.

**A**: Heatmap displaying genewise and patientwise double clustering of 218 CRC *BRAF* mutant patients. Clustering was performed using the Ward method to compute distance between genes and between patients. Gene expression is displayed as standardized z-score. **B**: NMF clustering after adjustment for MSI status, the BM subtype calls were then compared to the original calls computed without any adjustment. High agreement is shown between the two methods by the Cohen’s kappa coefficient of 0.82.

**Supplementary Figure S3**. Gene Set Enrichment Analysis of *BRAF* mutant patients.

The 476 *BRAF* mutant specific genes were used to perform a standard GSEA analysis between BM1 and BM2. EMT (MSigDB ref: M5930) and KRAS (MSigDB ref: M5953) signatures were found enriched in BM1 while G2M (MSigDB ref: M5901) and E2F (MSigDB ref: M5925) signatures were enriched in BM2. Venn diagrams on the right side highlight the number of genes that are shared or specific between the respective signatures and the indicated clinical factors.

**Supplementary Figure S4**. Pathway analysis comparison between BM subtypes, *KRAS*/*BRAF* wild-type and *KRAS* mutant CRCs.

Single sample gene set enrichment analysis (ssGSEA) was performed for BM1 versus *KRAS*/*BRAF* wild-type (WT2) patients (top panel, top heatmap, 69 patients per group) or versus *KRAS* mutant patients (bottom panel, top heatmap, 69 patients per group) and similarly for BM2 versus WT2 patients (top panel, bottom heatmap, 149 patients per group) or versus *KRAS* mutant patients (bottom panel, bottom heatmap, 149 patients per group). These analyses were performed using the hallmark gene signature collection from MSigDB and using the only the 476 BM-specific genes.

**Supplementary Figure S5**. Molecular characterization of BM subtypes.

Heatmap displaying the average ssGSEA hallmark signature scores for BM1, BM2, *KRAS*/*BRAF* wild-type (WT2) and *KRAS* mutant using all gene profiles (compared to Figure 3 where only the 476 BM-specific genes were used). False discovery rates adjusted p-values (FDR) for BM1 versus BM2 statistical analyses appear on the right under the form of a heatmap for both all genes and for the 476 BM-specific genes.

**Supplementary Figure S6**. MSigDB gene set overlap.

Venn diagram showing the number of genes that are overlap between collections of signatures tested in Figure 3.

**Supplementary Figure S7**. Immune profiles of BM subtypes.

A CIBERSORT analysis that reveals the immune landscape of BM subtypes in terms of relative percentage for each immune subset (top panel). The bottom panel displays the relative percentage of immune subsets that were found significantly enriched (FDR < 0.01) for each BM subtypes. Bottom heatmap displays the expression of macrophage-associated genes according to BM subtypes. MSI status is displayed as well.

**Supplementary Figure S8**. Methylation profiles of *BRAF* mutant subtypes.

**A**: In TCGA patients, general methylation profiles were assessed between *BRAF* wild-type and *BRAF* mutant patients. More than 95% of the probes had similar similar -values between these two groups. **B**: The analysis of two different methylation platforms used by TCGA shows that the methylation profiles between BM1 and BM2 patients is overall highly similar with few differences that are obviously driven by the MSI status. **C**: As the MSI status was enriched according to BM subtypes, we assessed the methylation status of several MLH1 probes and did not find any difference between BM subtypes. Rather, the differences were driven by the MSI status. **D**: List of gene probes that are differentially methylated between BM1 and BM2 and according to analyses made in different methylation platforms.

**Supplementary Figure S9**. Characteristics of patients enrolled in the survival analysis.

In the four cohort used in the survival analysis, the tables at the top of the figures display how the relapses and time zeros are defined, the number of survival events for overall survival (OS) and relapse-free survival (RFS) and the median of the censoring time (Med. Cens.) in days. For each clinical variables used in the survival analysis, the percentage of features is shown for each cohort. Heatmaps also show what the clinical features of each patient are.

**Supplementary Figure S10**. Performance of the BM classifier.

We built a generalized logistic model classifier for BM1/BM2. We performed one hundred leave-one-out cross-validation steps and selected a penalty coefficient lambda that minimizes the misclassification error. This plot shows the misclassification error rate of the hundred cross-validations according to the lambda (in the log scale). Error bars represents standard deviations. The selected lambda results in the selection of 44 non-zero coefficient genes (see list in supplementary Table S2) with a misclassification error of 11%.

**Supplementary Figure S11**. Cancer cell line drug sensitivity analyses.

Cell lines from the Sanger (**A**) and CCLE (**B**) databases were classified as BM1 or BM2 cell lines and their sensitivity to a panel of different drugs was assessed by differential analyses. Drugs from the top display the most different sensitivity between BM1 and BM2 cell lines.

**SUPPLEMENTARY TABLE LEGENDS**

**Supplementary Table S1**. Summary of mutations used as clinical factors in this study.

This table shows which amino acid change were considered as mutations for the *KRAS*, *BRAF* and *PIK3CA* genes according to the cohort.

**Supplementary Table S2**. Non-zero coefficients genes obtained by training a penalized generalized linear model to classify BM subtypes.

Table displaying the genes with non-zero coefficients after having trained a generalized logistic (binomial method) model via lasso penalty on the BRAF mutant dataset and using the BM1 and BM2 calls as response variables. These genes can be used to classify any sample with a gene expression profile into BM1 or BM2.

**Supplementary Table S3**. Association between *BRAF* mutant subtypes and other clinicopathological factors.

This table shows the association between *BRAF* mutant subtypes and clinical factors assessed by Fisher’s exact tests. P-value were adjusted for multiple comparisons using the Bonferroni method.

**Supplementary Table S4**. List of genes that are specific for *BRAF* mutant subtypes, MSI status and gender.

Tables displaying the list of genes that are specific for *BRAF* mutant subtypes, MSI status and gender after a multivariable linear regression. Each clinical factor is displayed in a separate spreadsheet. Each spreadsheet displays, for each gene, the coefficient, t-statistics and *p*-values in the multivariable regression. Genes were ordered with the ones at the top being the most associated with the clinical variables.

**Supplementary Table S5**. List of KEGG biological processes represented in mTOR M5924 and M7561 signatures.

Tables displaying the lists of KEGG biological processes found significantly represented in the MTORC1 hallmark signature (MSigDB ref: M5924) and KEGG MTOR signaling pathway (MSigDB: M7561). The silhouette scores and the adjusted *p*-values are displayed.

**Supplementary Table S6**. Classification of CRC cell lines into BM subtypes.

Table showing the *BRAF* mutant classification of CRC cell lines from Sanger, CCLE and GSK databases. BM1 cell lines are highlighted in red while BM2 in blue. If a cell lines was found called from the same subtype across several database, it was defined as consistently classified and thus kept for drug response analyses.

**Supplementary Table S7**. Drug responses of BM1 and BM2 CRC cell lines.

Tables displaying the drug responses of CRC cell lines classified as BM1 and BM2. The BM1\_Sanger and BM1\_CCLE spreadsheets show the drugs (ordered by efficiency) that display activity towards BM1 cell lines in Sanger and CCLE, respectively. Similarly, the BM2\_Sanger and BM2\_CCLE spreadsheets display the same information but towards BM2 cell lines. Differential analyses were performed to find whether BM subtypes display sensitivity to different drugs and the results appear in the DE\_Sanger and DE\_CCLE spreadsheets.

REFERENCES

1. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:1-25.

2. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453-7.

3. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013;39:782-95.

4. Fontanillo C, Nogales-Cadenas R, Pascual-Montano A, De las RJ. Functional analysis beyond enrichment: non-redundant reciprocal linkage of genes and biological terms. PLoS One 2011;6:e24289.

5. Popovici V, Budinska E, Tejpar S, Weinrich S, Estrella H, Hodgson G, et al. Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. J Clin Oncol 2012;30:1288-95.

6. Li J, Lu Y, Akbani R, Ju Z, Roebuck PL, Liu W, et al. TCPA: a resource for cancer functional proteomics data. Nat Methods 2013;10:1046-7.

7. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603-7.