doi:10.1038/nature25492

Tauriello et al.

TGF-beta drives immune evasion in genetically reconstituted colon cancer metastasis

Supplementary Information

Contents

Source data westernblots

Supplementary Discussion

Supplementary Discussion References

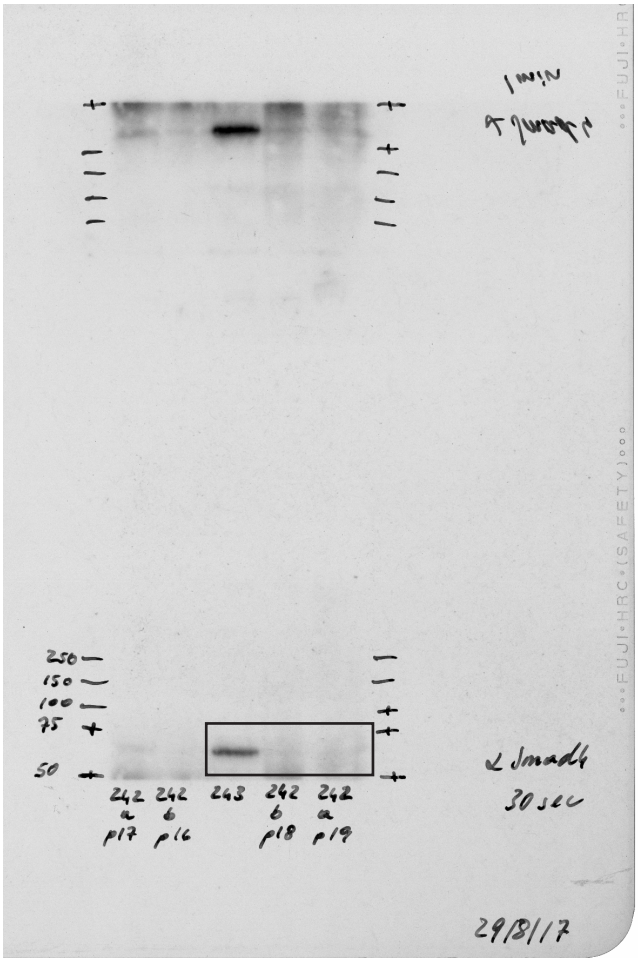
Supplementary Acknowledgements

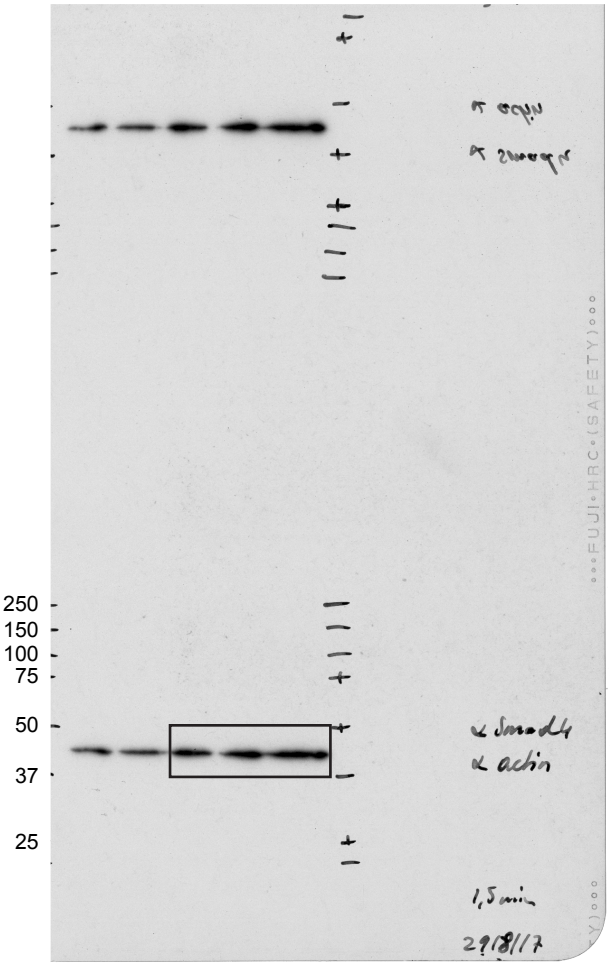
Supplementary Methods

Supplementary Methods References

Supplementary Fig. 1 - Source data westernblots

Fig. 2i raw images





Supplementary Discussion

The mutant mice and MTOs described herein are models to study poor prognosis human MSS CRC. Mutations in the four driver pathways render tumour cells independent of ISC niche factors, a feature that enables autonomous growth in unfavourable environments such those encountered by disseminated tumour cells in foreign organs as previously observed[1–4]. We found that malignancy co-evolves with changes in the TME reminiscent of those associated to poor prognosis CRCs such as T cell exclusion[5,6] and prominent recruitment of TGF-beta activated CAFs[7,89]. The finding that the penetrance of metastasis is ultimately modulated by the TME supports previous studies indicating that features of the immune system and the tumour stroma can help prognosticate the evolution of the disease in patients[5,7,9,10]. Recent consensus molecular classifications of CRC revealed the existence of a poor prognosis patient subset - CMS4 - characterized by prominent desmoplastic reaction, high TGF-beta levels and expression of signatures that indicate immunosuppression[8,11]. We show triple and quadruple mutant MTOs, when grown orthotopically, can be classified as CMS4 CRCs. It is, however, important to consider that frequency of genetic alterations in the main driver genes does not substantially differ between CMS2, CMS3 and CMS4 subtypes (the three subtypes that encompass most of the MSS patients)[8], implying that genotype – or at least mutations in the key oncogenes and tumor suppressor that modulate CRC progression - does not completely define molecular subtypes. It remains vital to characterize which other features and processes drive the acquisition of a TGF-beta activated TME.

It has been proposed that MSS CRCs are immunologically cold due to low mutational burden and therefore that they are unlikely to benefit from immune therapies[12]. In contrast, we found that metastatic cells from this class of CRCs can be effectively killed by the adaptive immune system through a CTL-dependent process, which CRC cells avert by raising TGF-beta levels. This finding provides a rationale for the overarching association between TGF-beta-driven gene programmes in the TME and poor prognosis in CRC[5,7,8,10,13]. We described two TGF-beta-driven mechanisms that cooperatively dampen immune responses. First, TGF-beta impacts directly on the capacity of T cells to mount a robust adaptive anti-tumor response. We observed a tight association between lack of T cell differentiation towards a Th1 effector phenotype and TGF-beta levels in CRCs. In line with this finding, mice deficient for *Tgfbr2* in T cells develop an early onset autoimmune syndrome characterized by premature differentiation of naïve $CD4^+$ T helper cell to a Th1 phenotype[14,15]. Mice that express a dominant negative *Tgfbr2* in T cells showed improved T-cell responses against melanoma and thymoma cell lines[16]. The expression of T-bet, the master transcription factor to drive the Th1 effector phenotype, is directly repressed by TGF-beta signalling[17]. TGF-beta signalling also inhibits $CD8^+$ T responses to T cell receptor stimuli[18–20]. In addition, the

immunosuppressive TME imposed by TGF-beta is likely the result of a coordinated effect on multiple immune and non-immune cell types[10].

Second, our data also reveal that a TGF-beta-activated TME excludes T cells from tumours, a phenomenon that characterizes MSS CRCs as well as many other cancer types[5,21,22]. The mechanisms of immune exclusion remains poorly understood but it may occur through direct action of TGF-beta on T cells or/and indirectly through regulation of certain cytokines and extracellular matrix proteins by TME components such as CAFs and endothelial cells[23,24]. Nevertheless, facilitating T cell infiltration through TGF-beta inhibition appears a pre-requisite for enabling anti-tumour responses by the adaptive immune system. In the particular case of CRC, blockade of immune exclusion by Galunisertib is sufficient to confer susceptibility to PD1/PDL1 checkpoint-based therapies, a strategy that may have broad application for treatment of multiple cancer types that thrive in a TGF-beta rich environment. Our findings can also be interpreted such that TGF-beta-mediated immune evasion operates hierarchically upstream of the PD1/PD-L1 inhibitory checkpoint.

Whereas future studies should be aimed at dissecting the relevance of each of these responses for CRC immune evasion, our results immediately suggest the use of TGF-beta signalling inhibitors as immunotherapy (possibly in combination with other imunnotherapies[25]) for patients with metastastatic CRC and in particular with poor prognosis CMS4 tumours. The tumour suppressor role of TGF-beta in epithelial cancer cells has warned against the use of TGF-beta inhibitors for cancer therapy[10]. Yet, we show that mice bearing metastases with an intact TGF-beta pathway are eradicated by TGF-beta inhibition implying that benefits of unleashing the immune system prevails in this setting and that TGF-beta based-immunotherapies may be safe in a broad range of CRC patients. Finally, although our experiments demonstrate a dependency on a TGF-beta-activated TME during metastasis formation, it remains to be proven that these results truly reflect therapeutic potential to eliminate residual CRC cells present in AJCC Stage II and Stage III CRC patients after surgical removal of the primary tumour. At present there is lack of CRC models that faithfully reproduce disease relapse after therapy.

Supplementary Discussion References

1. Fujii, M. *et al.* A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. *Cell Stem Cell* **18,** 827–838 (2016).

2. Matano, M. *et al.* Modeling colorectal cancer using CRISPR-Cas9–mediated engineering of human intestinal organoids. *Nat. Med.* **21,** 256–62 (2015).

3. Drost, J. *et al.* Sequential cancer mutations in cultured human intestinal stem cells. *Nature* **521,** 43–47 (2015).

4. Fumagalli, A. *et al.* Genetic dissection of colorectal cancer progression by orthotopic transplantation of engineered cancer organoids. *Proc. Natl. Acad. Sci.* **114,** E2357–E2364 (2017).

5. Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313,** 1960–4 (2006).

6. Mlecnik, B. *et al.* The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. *Sci. Transl. Med.* **8,** 327ra26 (2016).

7. Calon, A. *et al.* Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47,** 320–329 (2015).

8. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21,** 1350–1356 (2015).

9. Isella, C. *et al.* Stromal contribution to the colorectal cancer transcriptome. *Nat. Genet.* **47,** 312–9 (2015).

10. Tauriello, D. V. F. & Batlle, E. Targeting the Microenvironment in Advanced Colorectal Cancer. *Trends in Cancer* **2,** 495–504 (2016).

11. Becht, E. *et al.* Immune and Stromal Classification of Colorectal Cancer Is Associated with Molecular Subtypes and Relevant for Precision Immunotherapy. *Clin. Cancer Res.* **22,** 4057–4066 (2016).

12. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372,** 2509–2520 (2015).

13. Calon, A. *et al.* Dependency of Colorectal Cancer on a TGF-β-Driven Program in Stromal Cells for Metastasis Initiation. *Cancer Cell* **22,** 571–584 (2012).

14. Gorelik, L. & Flavell, R. A. Abrogation of TGFbeta signaling in T cells leads to spontaneous T cell differentiation and autoimmune disease. *Immunity* **12,** 171–81 (2000).

15. Marie, J. C., Liggitt, D. & Rudensky, A. Y. Cellular Mechanisms of Fatal Early-Onset Autoimmunity in Mice with the T Cell-Specific Targeting of Transforming Growth Factor-beta Receptor. *Immunity* **25,** 441–454 (2006).

16. Gorelik, L. & Flavell, R. A. Immune-mediated eradication of tumors through the blockade of transforming growth factor-beta signaling in T cells. *Nat. Med.* **7,** 1118–1122 (2001).

17. Gorelik, L., Constant, S. & Flavell, R. A. Mechanism of transforming growth factor beta-induced inhibition of T helper type 1 differentiation. *J. Exp. Med.* **195,** 1499–505 (2002).

18. Thomas, D. A. & Massague, J. TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell* **8,** 369–380 (2005).

19. Mehal, W. Z., Sheikh, S. Z., Gorelik, L. & Flavell, R. A. TGF-beta signaling regulates CD8+ T cell responses to high- and low-affinity TCR interactions. *Int. Immunol.* **17,** 531–538 (2005).

20. Zhang, N. & Bevan, M. J. TGF-beta signaling to T cells inhibits autoimmunity during

lymphopenia-driven proliferation. *Nat Immunol* **13,** 667–673 (2012).

21. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature* **541,** 321–330 (2017).

22. Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science (80-.).* **348,** 74–80 (2015).

23. Kraman, M. *et al.* Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. *Science (80-.).* **330,** 827–830 (2010).

24. Feig, C. *et al.* Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 20212–7 (2013).

25. Tauriello, D. V. F., Calon, A., Lonardo, E. & Batlle, E. Determinants of metastatic competency in colorectal cancer. *Molecular Oncology* **11,** 97–119 (2017).

Supplementary Acknowledgements

We thank Natalia Plana, Carolina Sanchez and Edgar Cristóbal Lecina for technical assistance, Francisco Barriga for materials, Raquel Batlle, Gonzalo Fernández Miranda and Raúl Méndez for mouse strains, Elena Sancho for help with the manuscript and important support, and Cristina Mendez for logistic support. From our outstanding facilities, we wish to thank from histopathology: Neus, Mònica, Begoña, Alicia, Anaïs and Antonia; from functional genomics: Nacho and David; from advanced digital microscopy: Sébastien and Anna; from flow cytometry: Jaume, Sonia, Chary and Ricardo; from the animal facilities: Steve, Barbara, Ferran, Conchi, Jesús, Rosa, Javi, Julia, Jordi, Vanessa and Fernando; and from the CRG genomic unit: Jochen, Anna and Maria.

Supplementary Methods

Cell lines

Murine cell lines MC38, CT26 and L-Wnt3a were obtained from ATCC and cultured in DMEM with 10% FBS. For MC38 and CT26, gDNA and mRNA was harvested from trypsinized, cultured cells and exome/RNA sequenced with the same protocols as for MTOs. Cell lines were regularly tested for mycoplasma contamination.

In vitro crypt niche factor dependency assays

Starting with full MTO-medium (above), we used combinations leaving out components, or adding 5 ng/ml recombinant human TGF-β1 (Peprotech) while removing galunisertib (+TGFB1). Wildtype normal intestinal organoids were obtained from a naïve C57BL/6J mouse, using the crypt shaking method[5], and cultured in MTO-medium supplemented with 1 µg/ml RSPO-1 and 50% v/v Wnt3a-CM. MTOs were never cultured with either supplement and WT organoids did not survive in MTO-medium without them. RSPO-1 was produced in-house[6] and Wnt3a-CM (from L-Wnt3a cells) was produced and tested according to previously described methods[7].

For Fig. 1h: MTOs were seeded in standard 48^{well} plates in triplicate as single cells and growth/death was manually scored at day 2, 4 and 6. Pictures were taken on day 5. Sensitivity was defined as difference in growth/survival compared to the control. This was assessed both manually, where 3 was the maximum level (0 *vs* +++), or using automated imaging (Olympus CellR/ScanR multiwell plate scanner; individual wells were scanned in a 5x4x8 (x-y-z) matrix) followed by ImageJ[8] analysis (Raw images were projected, stitched and analysed for automated organoid counting by ImageJ macros written by Anna LLado and Sébastien Tosi (IRB Barcelona ADM core facility); the result is shown in Extended Data Fig. 3a). Although the analyses were in high concordance, the heatmap is composed of data from the former. For each genotype, between 3 and 7 MTOs were analysed, p-values were calculated using the 0-3 sensitivity scores: for the LA-LAKTP EGF comparison p=0.0001, for the LA-LAKTP TGFB1 comparison: p<0.0001, and for the LAK-LAKTP TGFB1 comparison: p=0.0010 (two-sided MWW tests). Data were analysed and represented using R (v 3.4.2)[9], RStudio (v 1.1.383) and the *ggplot2* package (v 2.2.1)[10]. For diameter calculations, organoid sphericality was assumed.

Whole exome sequencing

Genomic DNA was harvested from cultured MTOs (median passage 5), as well as from mouse tail fragments, and purified using the GenElute kit (Sigma). Library prep was done using NebNext Ultra and the exome selection was done using SeqCapEZ developer (4 reaction, 110624_MM9_exome_ L2R_D02_EZ_Hx1; Cat. No. 06740278001, NimbleGen-Roche). Exome enrichment reaction was done in pools of 5 or 6 libraries. Sequencing was done HiSeq2500, Paired End, 125nts (2x125, v4); one lane per pool.

Reads preprocessing: Paired end reads were aligned to the mm10 version of the mouse genome using the mem algorithm implemented in the bwa software[11] (bwa-0.7.4) with default parameters. SAM files were sorted and indexed using Sambamba (v0.5.9; http://lomereiter.github.io/sambamba/) and duplicated reads removed with the Picard software (v1.128; https://github.com/broadinstitute/picard). Read qualities were recalibrated using the Genome Analysis Tool Kit (GATK, v3.5)[12].

Somatic SNP and INDEL calling: For each control-sample pair processed BAMs were compared using MuTect2 from GATK with default parameters. No extra filtering was applied other than the PASS/REJECT call from the Mutect2 algorithm. We decided to include

mutations marked as "clustered_events" since some positive controls (e.g. Kras (G12V) mutation in CT26) were otherwise excluded. Resulting VCF files were annotated with *snpEff* (v4.1)[13]. Somatic mutation load was computed with respect to the total length of the non-overlaping regions captured in the exome capture kit.

We constructed mutational signatures using the somatic mutations found in each of our samples. We normalized by the overall trinucleotide frequency in the mouse exome and compared against the human mutational signatures[14] using the algorithm implemented in the deconstructSigs software (v1.8.0)[15]. We removed signature "1B" from the initial set since it has been reported to be a linear combination of signatures 1 and 5[16].

RNA sequencing

mRNA was harvested from cultured MTOs and purified using the Ambion PureLink RNA mini kit (Life Technologies). Library prep was done using the Illumina kit Truseq Stranded mRNA Sample Prep kit. Sequencing was done on the HiSeq 2500, Paired End, 50nts (2x50, v4) with pools of 8-10 samples (one pool per lane). Paired end reads were aligned to the mm10 version of the mouse genome using the Star software (v2.3.0e)[17]. Expression was estimated using the R package *casper*[18]. The full expression matrix was normalized by quantiles.

Neoantigen prediction

Affinity prediction in mouse samples: SNVs annotated as non-synonymous or nonsense and inframe insertions or deletions were selected for each sample. Protein sequences were downloaded from Biomart using the *biomaRt* package[19] and the "mmusculus_gene_ensembl" dataset available on July 2017. A peptide including 10 bases up and downstream from the alteration was extracted from the canonical protein associated with the corresponding gene. Resulting peptides matching any other annotated protein were discarded from the dataset. MHC haplotypes for the C57BL (haplotype b) and BalbC (haplotype d) mouse strains were obtained from the Affymetrix eBioscience mouse haplotype table [http://tools.thermofisher.com/content/sfs/brochures/Mouse_Haplotype_Table.pdf]. All samples in the mouse cohort belong to the C57BL/6J strain except for the CT26 cell line, which has been reported as BalbC[20]. Peptides and the corresponding haplotypes for each sample were submitted to the netMHCpan affinity prediction software (v2.8)[21].

Human colorectal cancer mutations and neoantigens database: Total number of mutations and predicted neoantigens for human colorectal cancer (CRC) samples were downloaded from [https://tcia.at/neoantigens]. Samples were classified as MSS or MSI according to the TCGA consortium[22]. As well as for mouse, neoantigens in the CRC database were predicted using the netMHCpan software.

Affinity thresholds in mouse and human predictions: In order to find comparable thresholds between human and mouse affinity values, we generated 600,000 predictions from 12 human haplotypes and random peptides. We found good equivalence between a cutoff of 500nM and a 2% rank threshold, which has been used to filter the mouse dataset.

Neoantigen filtering: The CRC predicted neoantigens were filtered by a maximum affinity of 500nM and minimum expression of 0 (log2) RPKM[23]. Neoantigens in mouse samples were filtered with the following criteria: 2% Rank affinity as reported by netMHCpan and 0 RPKM of the mean expression of all available replicates for the corresponding gene.

Lentiviral infection of MTOs

For bioluminescent tracking, MTOs were infected with a lentivirus encoding an mCherry-firefly luciferase fusion reporter construct, followed by an IRES-Zeocin or IRES-Puro resistance cassette[3], cloned under control of the Ubiquitin promotor in a FUW vector[24]. Virus was produced using packaging constructs in HEK293T cells in DMEM 10% FBS medium, and filtered. Trypsinized organoids (single cells) were suspended in ultra-low attachment plates (Corning) in MTO-medium and treated with successive rounds of infection (adding virus-containing medium 50% v/v) in the presence of 8 µg/ml polybrene. After a week, MTOs were put back in BME drops and selected with 50 µg/ml zeocin or 0.5-1 µg/ml puromycin (InvivoGen). After 2-3 weeks, selected cells were sorted for mCherry expression.

Tumour dynamics with Bioluminescence imaging

Growth kinetics of luciferase-expressing MTOs were tracked with in vivo bioluminescence, using an IVIS-Spectrum (Perkin Elmer) imager. Animals were anesthetized with 2.5% isofluorane and received a retro orbital injection with 50 ul D-luciferin at 15 mg/ml (Resem BV). Mice were shaved before every measurement using electrical trimmers. For quantification, two or more images per mouse (typically in the 0.5-60s exposure range, bin 4-16) were averaged, quantifying the total flux (photons s^{-1}) of a ROI spanning the abdomen (IC) or lower thorax and upper abdomen (IS). Values were normalized per mouse on the value on day 0, measured 5 min after injection. Data were processed and visualized with R/RStudio[9] and ggplot2[10]. Depicted are longitudinal curves, connecting measurements of individual mice, and the group smooth (*LOESS*, span = 0.5) with 95% confidence interval. LOESS stands for locally weighted regression and is a non-parametric regression method analogous to a moving average in time[25].

Patient data analysis: Transcriptomic datasets

In this study, we used five Affymetrix microarray datasets publicly available in the NCBI GEO repository, which include gene expression and clinical information from a total of 1.194 CRC patients: briefly, GSE14333[26] contains a pool of 290 patients with CRC treated at 2 different hospitals: the Peter MacCallum Cancer Centre (Australia) and the H. Lee Moffitt Cancer Center (United States); the GSE33113[27] data set includes 90 patients with AJCC stage II disease collected at the Academic Medical Center in Amsterdam (the Netherlands); GSE39582[28] includes expression and clinical data for 566 patients with CRC collected for the Cartes d'Identité des Tumeurs (CIT) program, from the French Ligue Nationale Contre le Cancer; GSE3782 (https://www.ncbi.nlm.nih.gov/geo/query/ acc.cgi?acc=GSE37892, accessed 20/10/2016) is a series of stage II and III CRC patients collected at five different hospitals from France (Marseille La Timone, Nice Lacassagne, Marseille Institut Paoli-Calmettes, Paris Lariboisiare, Nancy Brabois and Paris Saint-Antoine); and GSE38832[29] contains transcriptomic and clinical information from 127 patients treated at the H. Lee Moffitt Cancer Center (United States). In order to facilitate the integration of the datasets, samples from centres contributing with less than 10 samples (GSE37892: Marseille La Timone and Nancy Brabois; GSE38832: Nashville Veterans Affaires Medical Center) were excluded, leaving a total of 1.179 samples for downstream analyses.

Patient data analysis: Microarray processing

Processing of microarray samples was carried out separately for tumour samples of each dataset using packages *affy*[30] and *affyPLM*[31] from Bioconductor[32]. Raw CEL files were normalized using RMA background correction and summarization[33]. Standard quality controls were performed in order to identify abnormal samples[34] regarding: a) spatial artefacts in the hybridization process (scan images and pseudo-images from probe level

models); b) intensity dependences of differences between chips (MvA plots); c) RNA quality (RNA digest plot); d) global intensity levels (boxplot of perfect match log-intensity distributions before and after normalization and RLE plots); e) anomalous intensity profile compared to the rest of samples (NUSE plots, Principal Component Analyses). Technical information concerning samples processing and hybridization was retrieved from the original CEL files: date of scanning were collected in order to define scan batches in each dataset separately; technical metrics described by Eklund AC and Szallasi Z[35] were computed and recorded as additional features for each sample. Probeset annotation was performed using the information available in Affymetrix web page (Affymetrix Analysis Center. Netaffx. https://www.affymetrix.com/analysis/ index.affx, accessed 07/27/2016). Each dataset was a-priori corrected by potential sources of bias due to technical variability. For doing so, a linear model was fitted to each gene and dataset separately including centre, three Eklunds' metrics (PM IQR, RMA IQR and RNA DEG), scanning day and the interaction between centre and Eklunds' metrics. This correction was carried out using a mixed-effect model in which gender, age at diagnosis, tumour location and microsatellite instability (MSI) status were also included as covariates, when available. Scanning day was modelled as a random effect in these models. The five transcriptomic datasets were then merged in a unique expression matrix after applying quantile normalization[36]. Finally, expression data from each dataset was standardized at probeset level using GSE39582 as a reference; for each dataset and probeset, we selected randomly a subset of samples from GSE39582 matching the same clinical characteristics regarding gender, age, tumour location and MSI, and then centred and scaled the dataset according to these parameters.

Patient data analysis: MSI imputation

MSI status was imputed in each dataset separately, based on the expression of genes included in a published transcriptomic signature[37]. For doing so, clustering analysis was performed on this signature via non-parametric density estimation[38,39]. Accuracy of this imputation was evaluated in dataset GSE39582, which included annotation of microsatellite-stable (MSS) and -instable (MSI) samples in their clinical information (97% and 80% accuracy for MSS and MSI samples, respectively).

Patient data analysis: Signatures summarization

To summarize the signatures analysed in this work, Z-scores were computed for each gene and each sample in the unique corrected expression matrix, which were then averaged across all genes included in the profile. The resulting score was centred and scaled across samples.

Patient data analysis: CD4 Th1 and Naive signatures

In order to obtain CD4 naive and CD4 th1 profiles, we carried out a differential expression analysis on gene expression dataset GSE22886[40]. Analyses were performed separately for HG-133A and HG-133B platforms using a linear model with empirical shrinkage[41] as implemented in *limma* R package[31]. We defined the CD4 naive signature to include genes that are at least 3 fold upregulated both in CD4 naive vs. Th1 and in CD4 naive vs. Th2, and Benejamini and Hochbergs' False Discovery Rate (FDR)[42] < 5% in both comparisons. Similarly, the CD4 Th1 signature included genes that were at least 3 fold upregulated compared to CD4 naive, and limma FDR < 5%. These signatures were summarized as scores as described in the previous sections. Relative expression of Th1 and Naive cells were then measured by means of the difference of these scores, which corresponds to a log-ratio in the original scale of the microarray intensities.

Patient data analysis: Association analyses

To test the association between expression profiles, a mixed-effect model was fitted to each gene/signature independently in which scan day was included as a random effect, and centre, Eklunds' metrics and the interaction between both were introduced as covariates. Association was assessed using the corresponding Wald and Likelihood Ratio Test (LRT) provided by the linear model. As measure of association and for continuous variables, we computed the Pearson's Correlation Coefficient after correction by technical effects using the mixed-effect model described in the previous section; for categorical variables, adjusted group means and corresponding intervals at 95% confidence were calculated. Association with Cancer Associated Fibroblasts (CAFs) was assessed using a FAP signature derived from human CRC sorted cells (see section *Gene expression in tumour cell populations*). To assess association with overall TGFB levels, a signature was created using microarray probesets *203085_s_at*, *220406_at* and *209747_at*, respectively, as these probesets correlated better with measurements of *TGB1*, *TGFB2* and *TGFB3* obtained by qRT-PCR data[3]. When assessing correlation between TGF-beta and any signature, probesets mapping to *TGFB1*, *TGFB2* or *TGFB3* were previously excluded from the later. For visualization purposes, scatter plots and boxplots were yielded using the technical adjusted values for expression intensities.

The estimates of the association between Th1 and Naive cells ratios and the *TGF-beta* and the FAP signatures were compared with the overall trend observed in the whole genome. For doing so, we generated a null distribution for the correlation coefficients between *TGF-beta* and the scores derived for random signatures. We randomly selected the same number of genes as those corresponding to the original naïve and Th1 signatures and computed the score's ratio. *TGF-beta* probes were excluded for the *TGF-beta* correlation analysis. Bilateral p-values were computed from 10,000 repetitions. The threshold for statistical significance was set at 5%. All these analyses were carried out using the R[9,31] packages *lme4*[43] and *lmerTest*[44].

Patient data analysis: Survival analyses

Association with metastasis was evaluated using a frailty Cox proportional hazards model as previously described[45] and implemented[46]. Statistical significance was assessed by means of a Log-likelihood Ratio Test (LRT), while Wald tests were used for pairwise comparisons when necessary. Association of expression intensities was evaluated as continuous variables assuming a linear relationship with the logarithm of the relative risk. Sample groups of low, medium and high expression levels were defined using the tertiles of the intensity distribution after correction by technical effects, as described in the previous section. When evaluating expression intensities, technical effects were included in the models as covariates. In an analogous way to the correlation analyses, scan batch was included as a random effect in the frailty Cox models. Hazard Rations (HR) and their corresponding 95% confidence intervals were computed as a measure of association. For visualization purposes, Kaplan-Meier curves were estimated for groups of tumours showing low, medium and high gene or signature expression. Only samples from patients diagnosed in stages I, II and III were taken into consideration for these analyses. The threshold for statistical significance was set at 5%. All analyses were carried out using R[9,31].

To compare the risk estimations with the general trend observed in the dataset, we generated a null distribution in the same way described in the previous section. As with the original signature, a Cox model was used to compute the statistic of the three pairwise comparisons for random scores and a bilateral p-value was computed from 10,000 repetitions.

MTO-derived whole tumour expression data

cDNA Library preparation and amplification were performed with the WTA2 kit (Sigma-Aldrich) from 25 ng of starting material. The cDNA was amplified for 17 cycles and purified using PureLink Quick PCR Purification Kit (Invitrogen). Quantification of amplified cDNA was done on a Nanodrop ND-1000 spectrophotometer (Thermo-Fisher Scientific, Waltham, MA, USA). 8.5 ug of the cDNA from each sample was fragmented and labelled with the GeneChip Mapping 250 K Nsp assay kit (Affymetrix). Hybridization was performed using the GeneAtlas Hyb, Wash and Stain Kit for 3' IVT arrays. Samples were denatured at 96°C for 10 min prior to incubation in the Affymetrix GeneAtlas Mouse MG-430 PM Array Strip. Hybridization was performed for 16 h at 45 °C in the GeneAtlas Hybridization Oven. Washing and Stain steps after hybridization were performed in the GeneAtlas Fluidics Station, following the specific script for Mouse MG-430 PM Arrays. Arrays were scanned with GeneAtlas Scanner using default parameters, and the generation of Cel files for bioinformatics analysis was done with GeneAtlas software (all Affymetrix).

As for human transcriptomic datasets, CEL files from organoids-derived mice whole tumours were normalized with RMA[33] using the Bioconductor R packages *affy*[30] and *affyPLM*[31]. Standard quality controls (see section Patient data analysis: Microarray processing) discarded the presence of abnormal samples. For probeset annotation we used the informationa available in the Affymetrix web page (Affymetrix Analysis Center. Netaffx. https://www.affymetrix.com/analysis/index.affx, accessed 08/01/2017).

TBRS comparisons:

Previous to the summarization of TBRS signatures, the expression matrix was corrected a-priori by Eklund's metrics[35] RMA.IQR and RNA.DEG, as they were observed to be an important source of technical variability. For such correction, a linear model was used with no other variables included as covariates. Signatures scores were computed independently for samples with three or four mutations. Scores between treatment conditions were compared using a two-sided Mann-Whitney-Wilcoxon test.

Classification of mice tumours according to CMS (extended)

To assess the similarity to the poor-prognosis CMS4 phenotype[47], we first attempted to use the Single-Sample-Predictor (SSP) provided in their original work. Nevertheless, this analysis did not succeed in classifying any of the samples being tested. For this reason, we hypothesized that the differences of platform and/or species between the training (human) and test data (mice) could be harming the performance of this classifier.

As an alternative method, we used *k*-Top Scoring Pairs (*k*TSP)[48] as implemented in the switchBox R package[49]. *k*TSP is a ranked-based method that relies in the ordering in small number of features within each sample and, thus, is very suitable in multi-platforms settings in comparison with other top-performer methods[50]. This *k*TSP classifier was trained for a binary outcome (CMS4 vs not-CMS4) in 960 samples with CMS annotation provided by the original work (168 CMS1, 421 CMS2, 127 CMS3 and 244 CMS4)[47]. The concordance between the SSP and *k*TSP methods was assessed in two human datasets that did not participate in the definition of the CMS classification: GSE38832 and GSE44076[51]; phi coefficients were 0.83 and 0.80 for these series, respectively:

GSE38832		*k*TSP	
		CMS1/2/3	CMS4
SSP	CMS1/2/3	67	7
	CMS4	0	23
	Not classified	16	9

GSE44076		*k*TSP	
		CMS1/2/3	CMS4
SSP	CMS1/2/3	61	4
	CMS4	0	8
	Not classified	13	12

Overall, these results suggested that, as *k*TSP could be more suitable for handling platforms and species differences between training and test datasets, it still retained the good classification performance attributable to SSP.

Previous to any attempt of classification, both mice and human CRC data were summarized from probeset to gene level (Entrez). For doing so, we computed the first principal component from all probesets mapping to the each gene. This component was then centred and scaled to the weighted mean of the means and standard deviations of the probesets using the corresponding contribution to the component as weight. The sign of the component was eventually changed to be congruent to the sign of the probeset contributing the most to the component. Next, mice genes were translated to their corresponding human homologous using the Mouse Genome Informatics Database[52]. Only high confidence homologous genes according to Biomart[53,54] were kept for downstream analyses (Mouse Orthology Confidence=1 and Mouse Gene-order conservation score=100; 10.015 genes). Genes to be included in the model were pre-selected from those over-expressed (1.5 fold-change minimum difference; 293 genes) or under-expressed (1.25 fold-change minimum difference; 215 genes) in CMS4 samples compared with any other subtype in the patients transcriptomic dataset. The classifier was internally tested in the training set of human samples by cross-validation leaving one dataset out for test at each iteration. A classifier trained with all the human samples available was then applied to the mice tumours using a simple rule of majority of votes. The decision cut-off was set at the midpoint of the maximum number of votes, which corresponds to the number of gene pairs used for classification. For visualization purposes, the number of votes for each classifier (cross-validation and training) were re-scaled to a score ranging between zero and one in order to represent all the classification results at the same scale.

Gene pairs:

UP	DOWN
MSRB3	*EPT1*
FERMT2	*RMI1*
EFEMP2	*ASF1B*
SPOCK1	*STIL*
DDR2	*CCNA2*
TAGLN	*HMGB2*
CCDC80	*CDC45*
GLI3	*KIF18A*
TNS1	*UNG*
STON1	*WHSC1*
PTRF	*KIF18B*
SFRP2	*HK2*
MGP	*RBM47*
GAS1	*DONSON*
BNC2	*PLK4*

SLIT2	*CCDC134*
DPYSL3	*EIF4E*
AEBP1	*RHPN2*
PCDH7	*FAM83F*
MAP1B	*FANCD2*
CRYAB	*HOOK1*
PRRX1	*CDCA2*
FBN1	*EZH2*
MXRA8	*GMCL1*
PTGIS	*CENPA*
ZFPM2	*ORC1*
MLLT11	*SPAG5*
MYL9	*TMEM54*

In agreement with the reasoning of a single-sample classifier, SSP and *k*TSP were both applied to a version of the mice data previous to any correction by technical effects. In addition, *k*TSP was also trained before technical corrections were applied to the human transcriptomic dataset. For the concordance analyses between SSP and *k*TSP, GSE44076 cel files were downloaded from GEO and processed with RMA using packages *affy*[30] and *affyPLM*[31] from Bioconductor[32].

Flow cytometry: single cell suspension preparation

Livers with tumours were removed, lobules (and in case of micro-dissection: individual liver metastases) were carefully dissected and finely minced with scalpels. The tissue was enzymatically digested in 10 ml of DMEM supplemented with 10% FBS, 1% HEPES, sodium pyruvate, glutamine, streptomycin and penicillin and 0.1% β -mercaptoethanol (Gibco) and containing 1 mg/ml Collagenase A (Roche), 0.2 mg/ml Dispase II (Sigma) and 0.2 mg/ml DNAse I (Roche), during 25 min at 37ºC with rotation. The enzymatic reaction was quenched by the addition of 30 ml of ice-cold DMEM (10% FBS, supplement). Cell suspension was filtered through a 70 µm cell strainer (BD). The filter was washed with 10 ml of ice-cold 10% FBS DMEM and the cells were pelleted at 280 g for 5 min at 4ºC. Lysis of erythrocytes was performed in Red Cell Lysis Buffer (RCLB, 155 mM NH_4Cl, 12 mM $NaHCO_3$, 0.1 mM EDTA) during 4 min at room temperature and immediately washed with ice-cold 10% FBS DMEM. After filtration through a 70 µm cell strainer and centrifugation. Sequentially, cells were purified by centrifugation 30 min at 2,400 rpm in 40/80 Percoll (Sigma) gradient. Cells were resuspended in-DMEM (10% FBS, supplement; FACS buffer).

Flow cytometry: immunophenotyping analysis and sorting

Cells were incubated 10 min on ice in FACS buffer in the presence of anti-CD16/CD32 (clone 93, eBioscience) to block Fc receptor. Mix of conjugated antibodies was added in the presence of anti-CD16/CD32 and the cells were stained during 20 min on ice. After staining, cells were washed and labelled for cell viability with the LIVE/DEAD Fixable Cell Dead stain kit (Life Technologies) during 7 min. For cytokine analysis, cell suspensions were incubated 4 h in PMA/ionomycin (Sigma-Aldrich) and brefeldin A (eBioscience) at 37 °C. Intracellular staining was performed using IC fixation/permeabilization kit (eBioscience). Cells were stained using PBS, 1% FBS, 1% HEPES and 0.6% EDTA (Gibco).

Gating strategies: Figure 1g (MTO biobank generation): To sort Lgr5-GFP+ tumour cells, we gated: Cells/Single cells/Living cells/GFP+. Figures 2d, S9b andS11a: To assess genetic recombination in the *UbC-CreER^{T2}; Tgfbr2$^{fl/fl}$; R26^{mTmG}* model, or to sort tumour cell populations, we gated: Living cells/Cells/Single cells/PDGFRb+ (CAFs). From the negative population, we followed: Endothelial cells CD31+/CD45- and Leukocytes CD31-/CD45+, further refined to small size and good viability. Figure 3h: T cell activation markers were assessed inside: Living cells/Cells/Single cells, than CD45+/CD3+, then CD4+/CD8- or CD4-/CD8+. Figure 3i and 4j: T cell cytokine levels were assessed inside: Living cells/Cells/Single cells/CD45+/CD3+/NK1.1-, then CD4+/CD8- or CD8+/CD4-. Figure 3j: CD8+ T cells were sorted using gates: Living cells/Cells/Single cells/CD45+/CD19-/MHCII-/CD3+/CD4-/CD8+. Figure 4d: PD-L1 expression was assessed inside: Living cells/Cells/Single cells/CD45+/CD3+, then CD4+/CD8- or CD8+/CD4-. Figure 4e: PD-L1 populations were determined using: Living cells/Cells/Single cells/PD-L1+. Within these we measured for CAFs (PDGFRb+/Epcam-) or Epithelial cells (PDGFRb-/Epcam+). From the double-negative population, we followed: Endothelial cells (CD31+/CD45-) and Leukocytes (CD45+/CD31-). Figure 4g: PD-L1 was measured inside Living cells/Cells/Single cells/CD45+. Figure 4i: For T cell activation markers we gated: Living cells/Cells/Single cells/CD45+/CD3+/NK1.1-, then CD4+/CD8- or CD8+/CD4-. Figure S13d: Myeloid populations were assessed inside: Living cells/Cells/Single cells/CD45+/PD-L1+/Ly6G-/Ly6C+/CD11b+/CD11c+/-.

Flow cytometry analysis and cell separation were performed in a FACSAriaFusion flow cytometer (Beckton Dickinson). Data were analyzed using FlowJo software (v 10.4).

The following antibodies were used for the staining: anti-CD45 (clone 30-F11), anti-CD4 (clone GK1.5), anti-aCD8 (clone 53-6.7), anti-CD69 (clone H1.2F3), anti-CD104b (PDGFRb, clone APB5), and anti-MHCII (clone M5/114.15.2) were obtained from eBioscience; Epcam (clone EBA-1), anti-CD3e (clone 145-2C11), anti-CD44 (clone G44-26), CD62L (clone MEL-14), anti-IFNγ (clone XMG1.2), anti-CD274 (PD-L1, clone MIH-5), and anti-Ly6C (clone AL-32) were obtained from BD Pharmingen; anti-CD31 (clone 390) was obtained from Abcam; and anti-CD11b (clone M1/70), anti-CD11c (clone H418), anti-GZMB (GB11), anti-CD279 (PD-1, clone 29F.1A12), anti-CD8a (clone 53-6.7), anti-T-bet (clone 4B10), anti-Ly6G (clone 1A8) were obtained from BioLegend.

Gene expression in tumour cell populations

RNA from sorted cells was processed and amplified as previously described[55]. For mouse liver metastases: To assess *Tgfb1*, *-2* and *-3* mRNA levels, we performed RT-qPCR with triplicate reactions (each 5 ng of cDNA) in a StepOne instrument (Thermo Fisher) with Taqman probes (*Tgfb1* Mm01178819_m1; *Tgfb2* Mm00436955_m1; *Tgfb3* Mm00436960_m1; endogenous control *Actb* Mm00607939_s1). Other probes used are: *Aurkb* Mm_01718140_m1, *mKi67* Mm01278617_m1, *Gzma* Mm01304452_m1, *Gzmb* Mm00442837_m1, *Pdcd1* Mm01285676_m1, and *Gapdh* Mm99999915_g1.

Sorted CRC populations from human samples were described previously[1,3]. Briefly, two GEO data sets were used to characterize gene profiles according to specific gene expression in tumour cell subpopulations: GSE39395 and GSE39396[3]. In these datasets and as described previously[3], FACS was used to separate the following populations from 14 fresh CRC samples: CD45+/EpCAM−/CD31−/FAP− ,CD45−/EpCAM+/CD31−/FAP−, CD45−/EpCAM−/CD31+/FAP− and CD45−/EpCAM−/CD31−/FAP+. To homogenize these two expression matrices, dataset GSE39396 was centred and scaled pairwise to the mean and standard deviation of GSE39395. A signature of Cancer Associated Fibroblasts (CAFs) was derived from the resulting expression matrix. For doing so, we selected probesets

overexpressed in the CD45−/EpCAM−/CD31−/FAP+ samples with a three minimum fold and raw p-value < 0.05 compared to any other cell population. These analyses were performed using a linear model with empirical shrinkage[41] as implemented in *limma* R package[31]. Also TGF-b levels by cell populations were assessed in this dataset using Kruskal Wallis (KW) and two-sided Mann-Whitney tests.

Supplementary Methods References

1. Calon, A. *et al.* Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47,** 320–329 (2015).

2. Warren, R. S., Yuan, H., Matli, M. R., Gillett, N. A. & Ferrara, N. Regulation by vascular endothelial growth factor of human colon cancer tumorigenesis in a mouse model of experimental liver metastasis. *J. Clin. Invest.* **95,** 1789–1797 (1995).

3. Calon, A. *et al.* Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. *Cancer Cell* **22,** 571–584 (2012).

4. Céspedes, M. V. *et al.* Orthotopic microinjection of human colon cancer cells in nude mice induces tumor foci in all clinically relevant metastatic sites. *Am. J. Pathol.* **170,** 1077–85 (2007).

5. Merlos-Suárez, A. *et al.* The Intestinal Stem Cell SignatureIdentifies Colorectal Cancer Stem Cells and Predicts Disease Relapse. *Cell Stem Cell* 511–524 (2011). doi:10.1016/j.stem.2011.02.020

6. Jung, P. *et al.* Isolation and in vitro expansion of human colonic stem cells. *Nat Med* **17,** 1225–1227 (2011).

7. Tauriello, D. V. F. *et al.* Loss of the tumor suppressor CYLD enhances Wnt/beta-catenin signaling through K63-linked ubiquitination of Dvl. *Mol. Cell* **37,** 607–19 (2010).

8. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **82,** 518–529 (2015).

9. Team, R. C. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* (2016). at <http://www.r-project.org>

10. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009). at <http://ggplot2.org>

11. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).

12. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43,** 491–8 (2011).

13. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w 1118; iso-2; iso-3. *Fly (Austin).* **6,** 80–92 (2012).

14. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500,** 415–421 (2013).

15. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17,** 31 (2016).

16. Alexandrov, L. B. *et al.* Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47,** 1402–1407 (2015).

17. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

18. Rossell, D., Stephan-Otto Attolini, C., Kroiss, M. & Stöcker, A. Quantifying alternative splicing from paired-end RNA-seq data. *Ann. Appl. Stat.* **8,** 309 (2014).

19. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4,** 1184–91 (2009).

20. Castle, J. C. *et al.* Immunomic, genomic and transcriptomic characterization of CT26 colorectal carcinoma. *BMC Genomics* **15,** 190 (2014).

21. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32,** 511–517 (2016).

22. Cancer Genome Atlas Network, T. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487,** 330–7 (2012).

23. Charoentong, P. *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* **18,** 248–262 (2017).

24. Lois, C., Hong, E. J., Pease, S., Brown, E. J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295,** 868–72 (2002).

25. Cleveland, W. S. & Devlin, S. J. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *J. Am. Stat. Assoc.* (2012).

26. Jorissen, R. N. *et al.* Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. *Clin. Cancer Res.* **15,** 7642–7651 (2009).

27. De Sousa E Melo, F. *et al.* Methylation of cancer-stem-cell-associated wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell* **9,** 476–485 (2011).

28. Marisa, L. *et al.* Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Med.* **10,** e1001453 (2013).

29. Tripathi, M. K. *et al.* Nuclear factor of activated T-cell activity is associated with metastatic capacity in colon cancer. *Cancer Res.* **74,** 6947–6957 (2014).

30. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. Affy - Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20,** 307–315 (2004).

31. Gentleman, R. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. *Journal of the American Statistical Association* **102,** (Springer New York, 2005).

32. Gentleman, R. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80 (2004).

33. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4,** 249–264 (2003).

34. Gentleman, R., Huber, W., Carey, V., Irizarry, R. & Dudoit, S. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. *Book* (Springer Science+Business Media, 2005).

35. Eklund, A. C. & Szallasi, Z. Correction of technical bias in clinical microarray data improves concordance with known biological information. *Genome Biol.* **9,** R26 (2008).

36. Bolstad, B. M., Irizarry, R. A., Åstrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193 (2003).

37. Jorissen, R. N. *et al.* DNA copy-number alterations underlie gene expression differences between microsatellite stable and unstable colorectal cancers. *Clin. Cancer Res.* **14,** 8061–8069 (2008).

38. Azzalini, A. & Torelli, N. Clustering via nonparametric density estimation. *Stat. Comput.* **17,** 71–80 (2007).

39. Azzalini, A. & Menardi, G. Clustering via Nonparametric Density Estimation: The *R* Package pdfCluster. *J. Stat. Softw.* **57,** 1–26 (2014).

40. Abbas, A. R. *et al.* Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun.* **6,** 319–331 (2005).

41. Smyth, G. K. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Stat. Appl. Genet. Mol. Biol.* **3,** 1–26 (2004).

42. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57,** (1995).

43. Bates, D., Maechler, M., Bolker, B. M. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. (2015). at <http://arxiv.org/abs/1406.5823>

44. Kuznetsova, A., Bruun Brockhoff, P. & Christensen, Haubo Bojesen, R. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. (2016). at <https://cran.r-project.org/package=lmerTest>

45. Therneau, T. M., Grambsch, P. M. & Pankratz, V. S. Penalized Survival Models and Frailty. *J. Comput. Graph. Stat.* **12,** 156–175 (2003).

46. Therneau, T. M. coxme: Mixed Effects Cox Models. R package version 2.2-5. (2015). at <https://cran.r-project.org/package=coxme>

47. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21,** 1350–1356 (2015).

48. Tan, A. C., Naiman, D. Q., Xu, L., Winslow, R. L. & Geman, D. Simple decision rules for classifying human cancers from gene expression profiles. *Bioinformatics* **21,** 3896–3904 (2005).

49. Afsari, B., Fertig, E. J., Geman, D. & Marchionni, L. switchBox: an R package for k–Top Scoring Pairs classifier development. *Bioinformatics* **31,** 273–274 (2015).

50. Marchionni, L., Afsari, B., Geman, D. & Leek, J. T. A simple and reproducible breast cancer prognostic test. *BMC Genomics* **14,** 336 (2013).

51. Sanz-Pamplona, R. *et al.* Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer. *Mol. Cancer* **13,** 46 (2014).

52. Blake, J. A. *et al.* Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* **45,** D723–D729 (2017).

53. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21,** 3439–3440 (2005).

54. Smedley, D. *et al.* The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res.* **43,** W589-98 (2015).

55. Gonzalez-Roca, E. *et al.* Accurate Expression Profiling of Very Small Cell Populations. *PLoS One* **5,** e14418 (2010).