SUPLEMENTAL METHODS

Establishment of tumor cells cultures

PBMCs were obtained from CRC patients or HLA-matched healthy donors by standard Ficoll separation (Ficoll-Paque PLUS, GE Healthcare Bio-Science). To generate differentiated CRC cell lines, surgical specimens were cut into fragments, filtered through sterile gauze and cultured in RPMI medium supplemented with 5% FCS. CSCs were isolated by mechanical processing of primary tumor tissue and cultured as spheres in the presence of stem cell permissive medium (DMEM/F12) as described [1]. EBV-transformed B-cell lines were established from patients PBMCs by immortalization with EBV containing culture supernatant. HLA-A*0201-positive T2 cells were obtained from the American Type Culture Collection (ATCC, Manassas, USA).

Massive parallel cDNA sequencing

Amplified cDNA pools (3µg) were processed for massive sequencing according to the GS FLX Titanium protocol (454 Life Sciences, Roche, Branfort, CT, USA). Single strand (sstDNA) libraries were quantitated by RiboGreen RNA Quantitation Kit (Invitrogen Inc., Carlsbad, CA, USA), checked for quality by capillary electrophoresis (Agilent Bioanalyzer 2100 with the RNA Pico 6000 LabChip kit; Agilent Technologies, Palo Alto, CA, USA) and amplified in emulsion-based clonal reaction. Each enriched sample was separately loaded onto a PicoTiterPlate (PTP) and pyrosequenced. From each sequencing run, we obtained an average of 126824±42030 high quality sequencing reads (approximately 50Mbases of raw sequence), with an optimal average 1180X depth per sample for all the 20 genes selected in the

panel. Each gene had a depth ≥100X in all the samples. cDNA sequences for the 20 selected genes were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/gene) and used as reference. Raw reads were aligned by using gsMapper software (v. 2.5, Roche) using default mapping parameters. Candidate mutations were first screened keeping in the analysis only those having more than 10 varied reads and an incidence of varied reads on the total of at least 10%. The alignments of the reads supporting all remaining variants were then manually checked in order to filter out false positives in error-prone regions (e.g.: homopolymers). After this step, the analysis was prioritized on the variants having an incidence of ≥40%. Peptide sequences were generated on a sample-by-sample basis introducing in the cDNA sequence the mutations deriving from the previous analysis and, then, traducing the sequence into peptides using EMBOSS [2] function “transeq”; finally, open-reading-frames (ORFs) were predicted by using the function “checktrans” of the EMBOSS suite.

MHC-peptide binding analyses

Quantitative assays to measure the binding of peptides to HLA A*02:01 class I molecules are based on the inhibition of binding of a radiolabeled standard peptide (HBV core 18-27 analog, FLPSDYFPSV). MHC molecules were purified by affinity chromatography from the EBV transformed homozygous cell line JY, and assays performed as described previously [3]. Peptides were tested at six different concentrations covering a 100,000-fold dose range in three or more independent assays, and the concentration of peptide yielding 50% inhibition of the binding of the radiolabeled probe peptide (IC_{50}) was

calculated. Predictions of the HLA A*02:01 binding capacity of peptides 8 to 11 residues in length were performed using the command-line version of the consensus prediction tool available at the Immune Epitope Database (IEDB; www.iedb.org) [4].

T cell cultures

T cell lines and mixed lymphocyte-tumor cell culture (MLTC) from patient 147 were generated from PBMCs as described [5]. To induce MLTCs, 1247 CSC cultures were pre-treated o.n.with hrIFN-γ (1000 IU/ml; PeproTech), before culturing with autologous PBMCs. To induce peptide-specific T cell lines, PBMCs were cultured with either mutated or WT peptides (10 μg/ml) in 96-well flat bottom plates in X-VIVO 15 medium (Lonza) supplemented with 5% normal human serum (EuroClone). After two days, recombinant human IL-2 (150 IU/mL; R&D) and IL-7 (10 ng/mL; R&D) were added to the cultures. At day 7, T cell cultures were re-stimulated with autologous irradiated (50 Gy) PBMCs pulsed with peptides and then re-stimulated every 7 days. At day 14, T cells were assayed for antigen recognition by ELISPOT, which was repeated twice in the next 14 days to confirm the results. Effector $CD8^+$ T cells were enriched from the cultures by negative immunomagnetic sorting using $CD4^+$ T cell Isolation Kit (Miltenyi Biotec). Mixed lymphocyte-tumor cell culture (MLTC) between 1247 PBMCs and autologous CSC were induced as described [1]. PBMCs (10^6 cells/well) of CRC patients were co-cultured with autologous irradiated (150 Gy) 1247 CSC cells in X-VIVO 15 medium (Lonza) supplemented with 10% normal human serum, 120 IU/mL rhIL-2, 5 ng/mL rhIL-7 and 10 μg/ml anti–IL-4 neutralizing mAb 3007 (all from R&D System).

MLTC were stimulated weekly with irradiated autologous CRC/CSC cells. T cell reactivity was tested after the third restimulation.

Retroviral transduction of mutated and WT SMAD4 minigenes

Two minigenes enconding 27 aa-long polipeptides encompassing either the SMAD4^{V370A} mutation expressed by the 1247 CRC, or the corresponding SMAD4$^{V370-WT}$ residue, flanked on each side by 14 and 12 aminoacids of the SMAD4 WT sequence, were cloned using BglII and XhoI restriction sites in the retroviral vector MSCV-IRES-GFP upstream of an IRES sequence preceding a GFP tag (kindly provided by Dr. K. Murphy, Washington University). Each minigene sequences contain a 5' Kozak consensus box and a 3' STOP codon:

```
Minigene SMAD4 Mut 1247
ATGtctggaggagatcgcttttgtttgggtcaactctccaatgCccacaggacagaagccattgagagagcaaggttgcacTGA
M  S  G  G  D  R  F  C  L  G  Q  L  S  N  A  H  R  T  E  A  I  E  R  A  R  L  H  STOP

Minigene SMAD4 WT
ATGtctggaggagatcgcttttgtttgggtcaactctccaatgTccacaggacagaagccattgagagagcaaggttgcacTGA
M  S  G  G  D  R  F  C  L  G  Q  L  S  N  V  H  R  T  E  A  I  E  R  A  R  L  H  STOP
```

Each retrovirus was produced upon transfection into human embryo kidney HEK293t packaging cells and transduced into different HEK293t cells, which were used as recipient cells because they are HLA-A*02:01^{+} and negative for the SMAD4^{V370A} mutation. Transduced HEK293t cells were sorted for the expression of EGFP (90%) and tested for the capacity to stimulate T cells from the 1247 patient that were specific for the SMAD4^{V370A} mutation.

PCR typing of mutated and WT SMAD4

The indicated tumor cell lines were screened for the expression of mRNA encoding either the SMAD4^{V370A} or SMAD4^{R361C} mutation, expressed by the

CRC cell lines 1247 and 1869, respectively. Total RNA was extracted from each tumor cell line, reverse transcribed into cDNA typed by PCR with upstream oligonucleotide primers ending exactly on the mutated nucleotide present in each sequence, coupled with downstream primers complementary to the WT sequence. The expression of the WT gene segment corresponding to either mutation was detected by similar PCR typing, utilizing upstream oligonucleotide primers ending exactly on the WT nucleotide present in each sequence, coupled with downstream primers complementary to the WT sequence. The upstream primer sequences are listed in Supplemental Table 4. PCR conditions were:

- SMAD4 mut and WT 1247: 35 cycles, 30" 94°C, 20" 64°C, 30" 72°C;
- SMAD4 mut and WT 1869: 30 cycles, 30" 94°C, 30" 65°C, 30" 72°C

PCR cycles were preceded by 10' denaturation at 94°C and terminated by 10' at 72°C.

ELISPOT assays

ELISPOT assay for IFNγ production by unique neo-antigen specific T cells were performed as described [1]. T cells were plated in duplicates at 5×10^3 cells/well in 96-well plates (MAIPS4510; Millipore, Bedford, Mass., USA) precoated with anti-IFNγ capture mAb (Mabtech) and stimulated for 4h with: CRC or CSC cell, LCL or T2 cell pre-loaded with each mutated or wild-type peptides (10 μg/ml), ±10 μg/ml of anti-HLA class I and/or anti-HLA-DR mAb. Biotinylated anti-IFNγ mAb (Thermo Scientific Pierce, Rockford, IL, USA), streptavidin–alkaline phosphatase conjugate (Amersham Pharmacia Biotech Europe GmbH, Freiburg, Germany,) and chromogenic substrate (nitroblue

tetrazolium-BCIP [5-bromo-4-chloro-3-indolylphosphate] (Sigma, St. Louis, Mo., USA) were subsequently added. The individual spots were counted using an automated ELISA-Spot Assay Video Analysis System (Eli-Scan) with the software Eli.Analyse V4.2 (A.EL.VIS, Hannover, Germany). As positive control, effector T cells were stimulated with PHA (1 µg/ml) (Sigma, St. Louis, Mo., USA). Unstimulated T lymphocytes represented the negative control.

REFERENCES

1 Volonte A, Di Tomaso T, Spinelli M, Todaro M, Sanvito F, Albarello L, *et al.* Cancer-initiating cells from colorectal cancer patients escape from T cell-mediated immunosurveillance in vitro through membrane-bound IL-4. Journal of immunology 2014;192:523-32.

2 Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000;16:276-7.

3 Sidney J, Southwood S, Moore C, Oseroff C, Pinilla C, Grey HM*, et al.* Measurement of MHC/peptide interactions by gel filtration or monoclonal antibody capture. Curr Protoc Immunol 2013;Chapter 18:Unit 18 3.

4 Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR*, et al.* The immune epitope database (IEDB) 3.0. Nucleic acids research 2015;43:D405-12.

5 Campi G, Crosti M, Consogno G, Facchinetti V, Conti-Fine BM, Longhi R*, et al.* CD4(+) T cells from healthy subjects and colon cancer patients recognize a carcinoembryonic antigen-specific immunodominant epitope. Cancer Res 2003;63:8481-6.