A genetic vaccine encoding shared cancer neoantigens to treat tumors with Microsatellite Instability.

Guido Leoni[1], Anna Morena D'Alise[1], Gabriella Cotugno[1], Francesca Langone[1], Irene Garzia[1], Maria De Lucia[1], Imma Fichera[1], Rosa Vitale[1], Veronica Bignone[1], Fabio Giovanni Tucci[1], Federica Mori[2], Adriano Leuzzi[1], Elena Di Matteo[1], Fulvia Troise[1] Adele Abbate[2], Rossella Merone[2], Valentino Ruzza[1], Maria Grazia Diodoro[3], Mahesh Yadav[4], Monica Gordon-Alonso[5], Cristophe Vanhaver[5], Maddalena Panigada[6], Elisa Soprana[6], Antonio Siccardi[6], Antonella Folgori[1], Stefano Colloca[1], Pierre van der Bruggen[5], Alfredo Nicosia[7,8,9], Armin Lahm[1], Maria Teresa Catanese[1] and Elisa Scarselli[1].

[1] *Nouscom S.R.L., Rome, Italy;*
[2] *Reithera S.R.L. Rome, Italy;*
[3] *Department of Pathology, "Regina Elena" National Cancer Institute Rome, Italy;*
[4] *Genentech South San Francisco, California 94080, USA;*
[5] *de Duve Institute and the Université catholique de Louvain, B-1200 Brussels, Belgium;*
[6] *Molecular Immunology Unit, San Raffaele Research Institute, Via Olgettina 58, 20132, Milan, Italy;*
[7] *Nouscom AG, Bäumleingasse, 18 CH-4051, Basel, Switzerland;*
[8] *Department of Molecular Medicine and Medical Biotechnology, University of Naples Federico II, Via Pansini 5, 80131, Naples, Italy;*
[9] *CEINGE, Via Comunale Margherita, 484-538, 80131, Naples, Italy;*

Running title
A vaccine for MSI tumors

Keywords
MSI, Colorectal cancer, gastric cancer, endometrial cancer, genetic vaccine

Corresponding author
g.leoni@nouscom.com
phone: 00390696036210

Conflict of interest
A.N., E.S., A.F., and S.C. are founders of Nouscom. All the other authors are employees of Nouscom, except M.Y., M.G., C.V., M.P., E.So., A.S., P.vdB., A.L., M.T.C. and declare no competing interests.

Abstract

Tumors with microsatellite instability are caused by a defective DNA mismatch repair system that leads to the accumulation of mutations within microsatellite regions. Indels in microsatellites of coding genes can result in the synthesis of frameshift peptides (FSP). FSP are tumor-specific neoantigens shared across MSI patients. In this study, we developed a neoantigen-based vaccine for the treatment of MSI tumors. Genetic sequences from 320 MSI tumor biopsies and matched healthy tissues in TCGA data base were analyzed to select shared FSP. 209 FSP were selected and cloned into non-human Great Ape Adenoviral and Modified Vaccinia Ankara vectors to generate a viral vectored vaccine referred to as Nous-209. Sequencing tumor biopsies of 20 independent MSI colorectal cancer patients revealed that a median number of 31 FSP out of the 209 encoded by the vaccine was detected both in DNA and mRNA extracted from each tumor biopsy. A relevant number of peptides encoded by the vaccine were predicted to bind patient HLA haplotypes. Vaccine immunogenicity was demonstrated in mice with potent and broad induction of FSP-specific CD8 and CD4 T cell responses. Moreover, a vaccine-encoded FSP was processed in vitro by human antigen-presenting cells and was subsequently able to activate human CD8 T cells. Nous-209 is an "off-the-shelf" cancer vaccine encoding many neoantigens shared across sporadic and hereditary MSI tumors. These results indicate that Nous-209 can induce the optimal breadth of immune responses that might achieve clinical benefit to treat and prevent MSI tumors.

Statement of Significance

Findings demonstrate the feasibility of an "off the shelf" vaccine for treatment and prevention of tumors harboring frameshift mutations and neoantigenic peptides as a result of microsatellite instability.

Introduction

Cancer vaccination had limited success in the past likely because of i) the use of tumor associated antigens (TAAs) as targets; ii) the presence of an immunosuppressive tumor microenvironment; and iii) the use of suboptimal vaccination platforms for induction of T cell immunity (1).

Today, vaccination is thought to be a viable therapeutic option in oncology if potent antigens arising from tumor-specific mutations can be selected and combined with drugs reversing tumor-mediated immune suppression (e.g. checkpoint inhibitors) (2).

The vaccine delivery system plays a crucial role in the quality and strength of induced immune responses. A heterologous prime/boost vaccination platform based on Great Ape Adenoviruses (GAd) and Modified Vaccinia Ankara (MVA) vectors has already been tested in over 4,000 volunteers in several clinical trials with different antigen payloads (relevant to Ebola, Malaria, HCV, HIV, and RSV) and shown to generate potent, durable, and high-quality T cell responses (3-8).

MSI tumors, characterized by a defective DNA mismatch repair (dMMR) system, have the highest mutational burden and, accordingly, show a good clinical response to checkpoint inhibitors (9,10). In MSI tumors, the presence of indel mutations in microsatellites regions of coding genes results in the synthesis of shared FSPs that are expected to be very potent and safe neoantigens (11-15) .

Therefore, MSI offers a unique opportunity for the development of an “off-the shelf” neoantigen based cancer vaccine which can be combined with PD-1 checkpoint inhibition in metastatic patients and may lead to improved clinical outcomes. Furthermore, given the extremely high risk of developing tumors for LS mutation carriers, a vaccine inducing a broad immune response could be used to prevent cancer occurrence/recurrence in this population (16,17) .

Materials and Methods

Patients and human samples

To design the Nous-209 vaccine we analyzed a dataset from 320 patients annotated as MSI in TCGA. For the validation, 20 additional samples were examined. NGS data (DNA exome and RNA) from 14 CRC MSI samples from the European Genome-Phenome archive (ID: EGAS00001000288) were downloaded with the permission of Genentech Inc. (1, DNAWay South San Francisco California 94080 USA). 4 tumor Flash Frozen (FF) biopsies of MSI CRC and matched healthy mucosa including 1 diagnosed with Lynch syndrome, were obtained from iSpecimen Inc. (450 Bedford Street, Lexington, Massachusetts 02420). Finally, 2 additional MSI CRC samples with Lynch syndrome, were obtained from “Regina Elena” National Cancer Institute in Rome. DNA and RNA was extracted from each collected sample and sequenced at CeGaT GmbH (Paul-Ehrlich-Straße 23, D-72076 Tübingen, Germany). For all subjects, appropriate Institutional Review Board approval and written informed consent was obtained according to the declaration of Helsinki.

.

Selection of FSMs for Nous-209 vaccine

Chromosomic coordinates of mono nucleotide repeats (MNRs) present in the human genome (Gchr37) with length ≥6 bp were determined by MISA (18). 7,524,435 MNRs were mapped on the human REFSEQ transcriptome (V73) to derive a subset of 30,678 MNRs located in the sequence of protein coding transcripts. Protected Mutation Annotation Format (MAF) files available from TCGA (release date 4.0-October 31 2016) were analyzed for the presence of FSMs overlapping with the 30,678 MNRs. FSMs derived from a 1 nucleotide deletion and with a number of reads harboring the mutation significantly higher in the tumor sample compared to the matched normal sample (FDR-corrected

Fisher-test *p-value* ≤ 0.05) were accepted. The resulting list of FSMs was then further refined by retaining only FSMs that fulfill the following criteria:

- presence in at least 5% of tumor samples in one of the three analyzed tumor types;
- allele frequency of the FSM in the matched normal samples ≤ 25%;
- presence of the FSM in < 2% of alleles of samples from a collection of healthy tissues in the ExAC database (19);
- Expression (RSEM $\log_2$ expression value downloaded from TCGA) of the FSM-carrying gene in all the three tumor types higher than the lower 20th percentile value of all expressed protein-coding genes determined considering all three tumor types.

NGS data analysis of 20 validation samples

Exomeseq and RNAseq reads were aligned on the Gchr37 human genome using HISAT2 2.0.4 (20). Multi-mapping reads were filtered out using Samtools 0.1.19 (21). Optical duplicates were marked using Picard's MarkDuplicates tool v1.14 (https://broadinstitute.github.io/picard/). DNA alignments were further optimized at regions around indels and base scores were recalibrated after the optimization step using GATK software v3.4.46 (22). A "look-up" analysis was performed by identifying MNRs mutated in the exome DNA with a minimum threshold of 3 mutated reads and a minimum mutation allele frequency of 10%. The number of reads at each FSM locus was estimated by using the mpileup utility of samtools 0.1.19 (21). Somatic variant calling of small indels was performed utilizing mutect2 (23), Varscan2 (24) and SCALPEL (25) with default parameters. All FSMs detected by at least one variant caller were considered.

Gene expression was computed by counting the number of reads mapped on each gene annotated in the human REFSEQ transcriptome (V73) with Rsubreads package (26) and

estimating transcript for million values (TPM). Look-up analysis was performed on RNAseq data to estimate the FSM loci covered by mutated reads.

MHC binding prediction

For each patient, the HLA class-I and class-II haplotypes were determined from the Exomeseq of the healthy tissue, using the Optitype and PHLAT software's (27) (28). MHC class I binding predictions were performed using the IEDB recommended method included in IEDB 2.17(29). 8-10 mer HLA-binders with a predicted IC50≤ 500nM were considered. MHC class II binding predictions were performed using the NetMHCIIpan method included in IEDB. HLA-II predicted binders with a percentile score ≤10% were considered.

Cell lines and reagents

M9 cells are derivative of Hek293 cells that stably express the Tet repressor gene (TetR). Adherent M9 cells were grown in DMEM supplemented with 10% FBS, 1% Pen/Strep, 2mM L-Glutamine, and 0.4mg/ml Geneticin. Suspension M9 cells were grown in CD293 medium supplemented with 6mM L-Glutamine and 0.4mg/ml Geneticin. For all experiments, M9 cells were used between passage 2 and passage 12 after thawing. Primary Chicken embryo fibroblasts (CEF) cells were prepared from SPF premium plus eggs (Charles River) and grown in VP-SFM medium supplemented with 1% Pen/Strep and 4mM L-Glutamine. CEF cells were frozen after preparation with no additional passages and used for all the experiments between passage 1 and passage 3 after thawing. All cell cultures supernatants were routinely tested with EZ-PCR Mycoplasma Detection Kit (Biological Industries) following manufacturer instruction, to exclude mycoplasma contaminations.

All cell culture reagents were purchased by Gibco, Thermo Fisher Scientific.The EBV-B cell line was established in the laboratory of Pierre van der Bruggen by co-culturing in IMDM (Life Technologies, Gibco) complemented with 10% fetal calf serum (Sigma-Aldrich), the PBMC of the donor with the supernatant of cell line B95-8 that contains infectious particles of EBV.

Media were supplemented with 0.24 mM L-asparagine, 0.55 mM L-arginine, 1.5 mM L-glutamine (AAG), 100 U/ml penicillin and 100 µg/ml streptomycin. Human recombinant IL-2 (Proleukin) was purchased from Novartis, IL-6 and IL-12 from Peprotech (Rocky Hill, NJ, USA), IL-7 from Miltenyi Biotec (Bergish Gladash, Germany) and GM-CSF (Leukine Sargramostim) from Sanofi (Paris, France). Human recombinant IL-4 was produced at de Duve Institute. For immunological analysis, a set of peptides covering the entire neo-antigenic sequence encoded by Nous-209 were obtained from JPT Peptide Technologies and used as antigens for ex-vivo immunological assays. FSPs shorter than 27 amino acids are represented by a single peptide, while for longer FSPs, multiple 15 mers, overlapping by 11 amino acids, were synthesized. 9-mer peptides used for in vitro stimulation of human T cells, were synthesized, purified, and characterized at the de Duve Institute.

GAd-209-FSP vectors production

The amino-acid sequence of the four artificial FSP-A genes (FSPA1-A4) was converted in nucleotide sequence based on human codon usage. At the N-terminus of each artificial gene, a sequence was added, corresponding to the amino acids 1-29 of the human TPA (Tissue Plasminogen Activator) protein (NP_000921.1). At the C-terminus an HA tag was added. The resulting transgenes were synthesized by GeneArt (Thermo Fisher Scientifics) and subcloned via EcoR1-Not1 restriction enzymes (New England Biolabs) into a shuttle plasmid containing the CMV promoter with two Tet Operator repeats and a BGH polyA. The expression cassettes were then transferred into the E1 deletion locus of pGAd20 plasmid by homologous recombination in BJ5183 cells (Agilent). pGAd20 contains the genome of a Gorilla Adenovirus (serotype group C) deleted in E1 and E3 regions. The four GAd vectors were produced by transfection of adherent M9 cells with Lipofectamine 2000 CD (Invitrogen, Thermo Fisher Scientific) and amplification in suspension M9 cells. Vectors were purified from infected cells by Vivapure Adenopack 20 RT (Sartorius). After purification, the four vectors were mixed to generate the GAd-209-FSP vaccine.

MVA-209-FSP vectors production

The four artificial FSP-B genes (FSPB1-FSPB4) were generated as described for the FSP-A genes. TPA and HA sequences (N- and C-term respectively) were also added. The resulting transgenes were synthesized by GeneArt (Thermo Fisher Scientifics) and subcloned via BamH1-Asc1 restriction enzymes (New England Biolabs) into a shuttle plasmid under the control of the P7.5 promoter. In addition, the shuttle plasmid carries an eGFP expression cassette and sequences homologous to the deletion III locus of MVA, to allow insertion of both expression cassettes (eGFP and transgene) in this locus. Recombinant MVA vectors were obtained by homologous recombination in CEF cells as previously described (28). The final recombinant vectors (devoid of any fluorescent gene) were further isolated by a limiting dilution step in CEF cells to obtain a pure viral population (28). After amplification in CEF cells, MVA vectors were purified from infected cells by centrifugation on a Sucrose cushion. After purification, the four vectors were mixed to generate the MVA-209-FSP vaccine.

Animals and vaccinations

Six-week-old female CB6F1 mice were purchased from Envigo. All day-to-day care was performed by trained mouse house staff at Allevamenti Plaisant SRL. All experimental procedures were approved by the Italian Ministry of Health (Authorizations 213/2016 PR) and have been done in accordance with the applicable Italian laws (D.L.vo 26/14 and following amendments) at the Institutional Animal Care and ethic Committee of Allevamenti Plaisant SRL. Viral vectors were administered i.m. in the quadriceps by delivering a total volume of 100 µl (50 µl per side).

Ex vivo immune analysis

IFN-γ ELISpot assays were performed on isolated murine splenocytes as previously described (30). In details we used 16 pools of peptides (P1-P16) encompassing the entire vaccine sequence. P1-P16 are composed of multiple sets of peptides (around 60 peptides per pool) covering the encoded FSPs. FSPs shorter than 27 amino acids are represented by a single peptide, while for longer FSPs, multiple 15 mers, overlapping by 11 amino acids, were synthesized. Peptides were diluted in DMSO and mixed to obtain 4 peptide pools per vector (pool 1 to pool 4 for vector FSP1; pool 5 to 8 for vector FSP2; pool 9 to pool 12 for vector FSP3; pool 13 to pool 16 for vector FSP4). DMSO (Sigma-Aldrich) and concanavalin A (Sigma-Aldrich) were used, respectively, as negative and positive controls. ELISpot data were expressed as IFN-γ Spot Forming Cells (SFCs) per million splenocytes. ELISpot responses were considered positive if all the following conditions occurred: (i) IFN-γ production present in ConA stimulated wells, (ii) the number of spots seen in positive wells was three times the number detected in the mock control wells (DMSO), (iii) at least 30 specific spots/million splenocytes. Intracellular IFN-γ staining was performed by overnight stimulation of splenocytes with peptide pools, as previously described (30).

In vitro stimulation of FSP-specific human CD8 T cells

Peripheral blood was obtained from hemochromatosis patient LB6019 as standard buffy coat preparation, which was laid down on a 15-ml Lymphoprep layer (Axis-Shield PoCAS, Oslo, Norway) in 50-ml Falcon Tubes. Isolation of CD8 T cells was performed as previously described (31). Lymphocyte-depleted PBMCs were left to adhere for 1 hour at 37°C in a culture flask at a density of 2×10^6 cells per cm^2 in IMDM/10% fetal calf serum. Non-adherent cells were discarded, and adherent cells were cultured in the presence of IL-

4 (200 U/ml) and GM-CSF (70 ng/ml) in complete IMDM medium. Cultures were fed on days 2 and 4 by removing half of the medium and adding fresh medium with cytokines. On day 6, monocyte-derived immature dendritic cells were incubated for 6 h with 1 µg/ml of peptide SLMEQIPHL (Ly4_24), IL-4 (200 U/ml), GM-CSF (70 ng/ml), in the presence of 1 µg/ml of Ribomunyl (INAVA, Pierre Fabre Medicament Production, Boulogne, France) and 500 U/ml of IFN-γ to induce their maturation, irradiated (50 Gray) and washed. Recombinant HLA-A*0201 molecules were folded in vitro with β2-microglobulin and peptide Ly4_24 or a control peptide (SLEPWIPYL). Peptides were purified by gel filtration, biotinylated, and mixed with streptravidin-PE (BD Pharmingen). For staining and sorting of CD8 T cells, cells were processed as previously *des*cribed (31). Positive T cells were sorted by MACS technology and cultured in round-bottomed low-adherence microwells (Costar) in presence of 200 µl of IMDM supplemented with AAG, 10% human serum, IL-2 (50 U/ml), IL-6 (1000 U/ml), IL-12 (12 ng/ml) and IL-7 (10 ng/ml). They were stimulated on day 0 and 7 with irradiated peptide-pulsed mature autologous DC at a ratio of 1:1. After 2 rounds of stimulation, approximately 10^5 cells from each microculture were stained with multimer and screened with a FACS Fortessa™ (BD Biosciences) using the DIVA™ software (BD Biosciences) to identify multimer$^+$ cells. The remaining cells in each positive microculture were collected 2 days later, stained with anti-CD8 and fluorescent multimers, and passed through a flow cytometer to isolate positive cells. These cells were restimulated every 7 to 12 days with irradiated (100 Gray) T2 cells loaded with 1 µg/ml of Ly4_24 peptide in presence of IL-2 (50 U/ml) and IL-7 (10 ng/ml).

MVA infection of EBV-B cells and recognition assay by FSP-specific T cell clones

EBV-B cells (2 x 10^6) from HLA*A2 patient CP50 were resuspended in a 15-ml tube in 250 µl of RPMI 1640 (Gibco) complemented with 2% human serum containing MVA-FSP-B2 at MOI of 5 and unrelated MVA encoding hcRed at MOI of 2 to control for infection efficiency.

Cells were incubated for 1 hours at 37°C, 5% CO_2. Cells were washed, resuspended at 10^6/ml in RPMI 2% HS and plated overnight at 37°C. Infection rate was measured by flow cytometry and cell mortality was analyzed with Trypan Blue. EBV infected cells were then tested for their ability to stimulate IFN-γ secretion by the T cell clones. In total, 10,000 T cells were added in microwells containing 10,000 EBV-B cells either untouched, loaded with 1 µg/ml of Ly4_24 peptide or infected with MVA-FSP-B2 encoding for Ly4_24 in a total volume of 200 µl of complete IMDM supplemented with 25 U/ml of IL-2. After 20 hours, the IFN-γ released in the supernatant was measured by ELISA using anti-IFN-γ capture antibody (clone 350B10G6, Thermofisher) and biotin-conjugated anti-IFN-γ monoclonal antibody (clone 67F12A8, Thermofisher).

Results

Selection of 209 FSPs shared across MSI tumors

Frame Shift Mutations (FSMs) detected in 320 MSI tumors of different histologies, 69 colorectal cancer (CRC), 85 gastric cancer (GC) and 166 endometrial cancer (EC), were subjected to a multistep filtering procedure to identify those shared across the three tumor types (32-34). 1494 FSMs present in the tumor and absent in matched healthy tissue were selected (Fig. 1A and Materials and Methods).

FSMs were mapped on the REFSEQ transcriptome to determine the amino acid sequences of the resulting FSPs, allowing for a single FSM generation of multiple FSPs due to alternatively spliced isoforms. To optimize for the presence of immunogenic epitopes, FSPs shorter than 4*aa* were removed and those shorter than 10*aa* were lengthened by adding up to 4 wild-type (wt) amino acids to the N-terminus to allow for generation of CD8 epitopes involving the junction between wt and FSP sequences. To minimize the risk of inducing auto-immunity, any FSP segment ≥ 8*aa* matching with wild-type human proteome was excluded. Finally, we obtained 1087 FSPs suitable for a vaccine (Fig. 1B). We considered 1087 FSPs too many for vaccine feasibility and developed an algorithm further reducing the number of FSPs to allow for their overall sequence to be encoded by 4 viral vectors. We determined that an antigenic sequence of 400*aa* contains on average 3 immunogenic CD8 epitopes (min 1, max 12) (Fig. S1) by reviewing published data on vaccination in healthy volunteers with antigens from HIV and HCV polyproteins encoded by viral vectors (4,8). Therefore, we developed an algorithm to select FSPs, from the initial collection of 1087, such that the maximum number of MSI tumor biopsies examined had in common with the putative vaccine a protein sequence of at least 400*aa*, intended as the sum of the lengths of FSPs encoded by the FSMs detected in the exome of each individual tumor. The final number of FSPs selected for the vaccine was 209 arising from 204 FSMs and encoding overall sequence of total 6021*aa*

corresponding to the maximum protein size that can be encoded by four viral vectors (Fig. 1A, Table S1). The median length of selected FSPs is 17*aa* (Fig. S2). More than 80% (166 out of 204) of the FSMs selected by the algorithm are shared across the 3 analyzed MSI tumor types (Fig. 1C) and have comparable mutation allele frequency in early and late stage tumors (Fig. S3).

A median number of 50, 46 and 21 FSPs is shared between the 209 FSPs selected for the vaccine and each CRC, GC and EC patient, respectively (Table 1). Moreover, the vaccine has a sequence of at least 400*aa* in common with 98% of CRC, 95% of GC samples and 70% of the EC cancer samples from TCGA database (Table 1).

Validation of 209 selected FSPs

To validate the FSPs' selection, an independent dataset of 20 MSI CRC patient samples, 3 of which were LS, was analyzed.

First, a "look-up" analysis (Materials and Methods) of tumors and matched healthy tissues was performed to confirm that Nous-209 FSMs were present in the tumor but absent as germline mutations. Detected FSMs encode a median number of 51 FSPs for a cumulative length ranging from 65 to 2415 amino acid (Fig. 2A and 2B), confirming data obtained by analyzing TCGA tumors. Across the 20 analyzed MSI patient samples, only 2 Nous-209 FSMs were detected in the DNAs from healthy mucosa of 2 patients (1 FSM/patient) with an allele frequency closed to the threshold of detection (10%; Fig. S4A).

A median of 31 FSPs (62% of FSPs present in tumor DNA; min 1 – max 57 FSPs/sample) was found to be expressed in tumor cells (red bars; Fig. 2C). Importantly, the 2 FSMs present in DNA of healthy tissues were not detected in the RNA (Fig. 2C). Moreover each patient's tumor RNA had in common with the vaccine between 26 and 1475 amino acids (median 718 amino acids), intended as the cumulative length of shared FSPs (Fig. 2D).

Exome sequence data were also analyzed by somatic variant calling confirming almost all FSMs detected in tumor DNAs by the look-up analysis with the only exception of sample 004 for which variant callers identified more FSMs than look-up analysis likely because FSMs detected in this patient have allele frequencies below the threshold used for look-up analysis (Fig. S4B).

Predicted HLA-I and HLA-II binders in FSPs encoded by Nous-209

We estimated the potential of Nous-209 FSPs to encode CD8 binding peptides, by predicting FSP-derived 8-10mer peptides that bind patient's HLA-I haplotypes, with an IC50 ≤ 500 nM (Fig. 3A). The Nous-209 vaccine encodes a median number of 335 HLA-I predicted binder peptides per patient (Fig. 3A; orange bars) of which a median of 27 (minimum 0 and maximum 100) are detected in both DNA and mRNA extracted from each patient's tumor biopsy (Fig. 3A; blue bars). Only in one patient (ID:12919), no CD8 binder peptides were predicted, likely because this patient harbors very rare HLA-I haplotypes for which a limited amount of experimental data are available in IEDB database lowering the performance of the prediction algorithm. By looking at class-II predictions we found a median of 751 predicted binder peptides of which 90 were encoded by FSMs expressed in each patient (Fig. 3B).

Generation of Nous-209 vaccine and immunogenicity in mice

The 209 FSPs were assembled into 4 artificial genes by joining their sequences one after the other in a head-to-tail configuration. Each artificial gene was assembled in two layouts (A and B) containing the same set of FSPs, but in a different order. Genes in the layout A were used for the construction of 4 GAd vectors (GAd-FSP-A1 to A4), while genes in the layout B were encoded in 4 MVA vectors (MVA-FSP-B1 to B4). The scrambling of FSPs

between the 2 layouts was designed to avoid GAd and MVA encoding the same junctional aminoacidic sequences between adjacent FSPs which could boost immune responses against epitopes at the junctions. The 4 GAd and the 4 MVA vectors were then mixed to generate, respectively, the polyvalent GAd-209-FSP and MVA-209-FSP vaccines, which are collectively referred to as Nous-209 (Fig. S5 A-B).

To demonstrate the feasibility of our vaccination approach in generating T cell response against multiple target antigens, the immunogenicity of Nous-209 was evaluated in CB6F1 mice using heterologous prime/boost regimen according to the scheme depicted in Figure 4A. GAd-209-FSP and MVA-209-FSP were administered intramuscularly (IM) at the dosage of $4x10^8$ vp and $4x10^7$ ifu respectively. Mice were primed at week 0 with GAd-209-FSP. Two weeks later, a group of animals was sacrificed to evaluate GAd-primed immune responses, while a second group was boosted with a single dose of MVA-209-FSP. FSP-specific T cell responses post prime (week 2) and post MVA boost (week 3) were measured by IFN-γ ELISpot on splenocytes restimulated with peptides covering the entire neo-antigenic sequence encoded by Nous-209 (Fig. 4B-C). Priming with GAd-209-FSP elicited a strong T cell mediated immunity. Responses were efficiently boosted by MVA-209-FSP, with approximately overall 10.000 spot forming colonies (SFC)/10^6 splenocytes (Figure 4B). Antigen-specific immune responses against each part of the neoantigen sequence covered by 16 peptide pools were detected in vaccinated animals (Fig. 4C and Materials and Methods). These results, although do not directly allow to derive information about immunogenicity in human, demonstrate the feasibility of employing Nous-209 vaccine for effective prime-boost immunization.

To exclude immunological interference due to vectors co-administration, we selected FSP1 vector for proof of concept. Therefore, we measured the immune responses induced after vaccination with GAd-FSP-A1 as a single vector ($1x10^7$vp) or co-administered with

the 3 other GAd vectors (each of them at the dose of $1x10^7$vp). Animals receiving GAd-FSP-A1 were boosted with the corresponding MVA vector (MVA-FSP-B1) at $1x10^7$ pfu, while the second group of mice, receiving the mixture of 4 GAds, was boosted with the 4 MVAs. T cell responses against the pools P1-P4 covering the FSPs in GAd-FSP-A1 and MVA-FSP-B1 were measured in both groups and found comparable, thus excluding immunological interference in case of concomitant, multivalent vaccination with the 4 viral vectors (Figure 4D). Moreover, the quality of induced T cell responses was investigated by flow cytometric intracellular staining (ICS) on splenocytes harvested post MVA GAd-FSP-A1/MVA-FSP-A1 vaccine. Cells were stimulated with pools 1 and 3, which yielded the strongest IFN-γ ELISpot responses against the FSPs encoded by the first vector. Both CD8 and $CD4^+$ IFN-γ^+ FSP-specific T cells were induced in vaccinated mice (Fig. 4E).

Isolation of FSP-specific CD8 T cell clones recognizing a naturally processed epitope encoded by Nous-209

To demonstrate that Nous-209-encoded epitopes can be processed, loaded in the MHC pocket of human cells and recognized by specific human T cells, we performed an in vitro experiment. 95 FSPs present in Nous-209 and observed in the majority of CRC patients (≥ 25% of patients) were analyzed for epitopes predicted to bind the HLA-A*02:01 molecule. Predicted epitopes were ranked by predicted IC50 (Table S2). Among the top scoring peptides, we selected the Ly4_Ly24 peptide (derived from FSP CKAP2) for proof of concept *in vitro* stimulation study, after confirming its binding to HLA-A*02:01 by T2 binding assay. We synthesized an HLA multimer folded with the Ly4_Ly24 peptide to enrich for FSP-specific CD8 T cells from an HLA-A*02:01 subject (35). The sorted T cells were expanded by two rounds of in vitro antigenic stimulation with autologous dendritic cells (DCs) pulsed with the Ly4_Ly24 peptide. The microcultures were screened with a fluorescent HLA-A2/Ly4_Ly24-multimer and 3 out of 301 were found positive. Clonal

412 populations were isolated from each of the 3 microcultures by flow cytometry, and
413 amplified in vitro by several rounds of antigenic stimulation. Using flow cytometry, we
414 isolated 3 T cells clones able to bind specifically the fluorescent HLA-A2/ Ly4_Ly24
415 multimer, but not a multimer with an unrelated peptide (Fig. 5A, 5C and Fig. S6).
416 Moreover, we showed that those clones selectively secrete IFN-γ when incubated with
417 HLA-matched EBV-B cells, either loaded with the neoantigen peptide or infected with the
418 MVA-FSP-B2 encoding the Ly4_Ly24 peptide (Fig. 5B).

Discussion

The knowledge that passive transfer of neoantigen-specific T cells resulted in complete remissions of patients with advanced malignancies makes neoantigens a highly desirable targets for cancer vaccination (36-39).

MSI solid tumors are characterized by a high number of indel mutations in coding genes that can give rise to FSPs that are entirely novel, non-self peptides and, as such, potentially the most immunogenic and safe type of neoantigens (14,15) .

Interestingly, a positive correlation between the total number of FSPs detected in MSI CRCs and the CD8 TIL density was observed (13), highlighting an important role for FSP-specific CD8 T cell response in controlling MSI tumors and providing a rationale for developing cancer vaccines targeting FSPs to strengthen and broaden the spontaneous immune response. Moreover, this type of mutations were shown to be shared to a variable extent across MSI patients (40,41), offering a unique opportunity for an “off-the-shelf” vaccination approach. Indeed, some shared FSPs derived from frequently mutated genes have been tested as peptide vaccine in MSI CRC patients (NCT01885702 and NCT01461148) and vaccination was shown to be safe and immunogenic (42,43).

Our work describes the widespread occurrence of shared FSPs in MSI tumors and translates those findings into the generation of Nous-209, the first genetic vaccine for MSI tumors. We selected 209 tumor-specific FSPs using TCGA dataset by taking into account several parameters to maximize the probability of inducing a tumor-specific immune response intra-patient and across different patients’ HLA haplotypes. Nous-209 was conceived to achieve a very broad coverage by targeting many different FSPs to face highly heterogeneous and rapidly evolving tumors.

Off-the-shelf cancer vaccines represent a very attractive approach as they overcome the complexity of a personalized cancer vaccine, including potential delays between biopsy, vaccine manufacturing, and release tests, as well the high costs associated with vaccine

production. To date, the current “off-the-shelf” neoantigen-based cancer vaccines target single or few frequently shared driver mutations (44). This implies: i) the need for patient pre-selection based on the presence of the targeted mutations and ii) the high likelihood of tumor escape if the mutation(s) in common with the vaccine is lost under the immune pressure. Our vaccine does not require MSI patient screening for specific mutations by NGS, and has a high probability of inducing effective antitumor activity by targeting a wide number of tumor neoantigens.

To validate the selection of FSPs included in vaccine, we used an independent cohort of MSI CRC patients samples. The presence of vaccine-encoded mutations was confirmed in this cohort with the same frequencies observed in the TCGA dataset. Importantly, none of the FSMs was expressed in the normal adjacent mucosa of MSI patients, highlighting the anticipated safe profile of the vaccine.

The value of Nous-209 is further strengthened by the use of a potent genetic vaccine platform based on GAd and MVA viral vectors. This platform has been validated in the field of infectious diseases for its potent induction of T cells in humans (4,8). Here we demonstrated a powerfull immunogenicity of the Nous-209 vaccine in mice in which we could demonstate the prime/boost effect with a balanced induction of both CD4 and CD8 T cells and the lack of interference among the four co-administred vectors. Although we could not evaluate the efficacy of the Nous-209 vaccination in an MSI model we recently provided a proof of concept of our vaccination platform encoding neoantigens in murine cancer models. Vaccination as standalone treatment was demonstrated to be 100% effective in a prophylactic setting, and to synergize with anti-PD-1 in advanced therapeutic settings of tumor-bearing mice, with a 3-fold increase of the cure rate compared to anti-PD1 monotherapy (30).

MSI tumors showed indeed a good response to anti-PD-1 treatment and the expectation is to increase the response rate by expanding T cells against relevant tumor neoantigens (9,10).

The presence of a human T cell repertoire against FSPs has been demonstrated both in healthy volunteers and cancer patients(13). Here, we provide further evidence by reporting an additional reactivity. Althought this finding is limited to only one FSP in Nous-209, we provided a demonstration that human APCs can process and present an epitope derived from a Nous-209 vaccine encoded FSP, resulting in activation of human T cell clones specific for that reactivity.

The breadth of the immune response inducible by Nous-209 vaccine was investigated by predicting the number of potential MHC binders present in the vaccine according to the patient's HLA class I and II haplotypes. Nous-209 targets a median of 27 class-I and 90 class-II predicted MHC binders in FSPs expressed by each patient's tumor and median number of 335 class-I and 678 class-II additional epitopes in FSPs potentially arising over the course of disease progression/relapse.

Mutations in the Beta-2-Microglobulin (B2M) gene, frequently observed in these patients, could represent an obstacle to vaccine effectiveness (45). Nevertheless, recently published evidence indicates that 85% of B2M mutated MSI CRC tumors do benefit from checkpoint inhibitor therapy (46). Based on these results, we expect that only a limited number of patients won't be able to respond to the vaccine because of impaired HLA-I presentation in the tumor.

We previously showed that benign polyps acquire a hypermutated phenotype and start to accumulate FSPs derived from indel mutations as they become more advanced (47). Nous-209 could therefore represent a prophylactic approach to attack these types of lesions and prevent their transformation into malignant tumors in LS carriers for which preventive options are today limited to aspirin that was shown to slightly reduce cancer

incidence in the CAPP2 randomized trial (48). This study opens new avenues for treatment and prevention of MSI tumors. Nous-209 vaccine is currently under clinical evaluation in metastatic CRC, gastric and gastro-esophageal cancer patients in combination with Pembrolizumab (NCT04041310).

Acknowledgements

We are grateful to Professor Riccardo Cortese, Nouscom founder and former CEO. Most of this work would have not been possible without his ideas and guidance. We thank Marina Udier, CEO of Nouscom for critical reading of the manuscript and helpful discussion.

We acknowledge the animal facility of Plaisant in Castel Romano (Rome) for the maintenance and care of the mice used in this study.

References

1. Hollingsworth RE, Jansen K. Turning the corner on therapeutic cancer vaccines. NPJ Vaccines **2019**;4:7 doi 10.1038/s41541-019-0103-y.
2. Hu Z, Ott PA, Wu CJ. Towards personalized, tumour-specific, therapeutic vaccines for cancer. Nat Rev Immunol **2018**;18(3):168-82 doi 10.1038/nri.2017.131.
3. Barnes E, Folgori A, Capone S, Swadling L, Aston S, Kurioka A, *et al.* Novel adenovirus-based vaccines induce broad and sustained T cell responses to HCV in man. Sci Transl Med **2012**;4(115):115ra1 doi 10.1126/scitranslmed.3003155.
4. Borthwick N, Ahmed T, Ondondo B, Hayes P, Rose A, Ebrahimsa U, *et al.* Vaccine-elicited human T cells recognizing conserved protein regions inhibit HIV-1. Mol Ther **2014**;22(2):464-75 doi 10.1038/mt.2013.248.
5. Ewer K, Rampling T, Venkatraman N, Bowyer G, Wright D, Lambe T, *et al.* A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA. N Engl J Med **2016**;374(17):1635-46 doi 10.1056/NEJMoa1411627.
6. Green CA, Scarselli E, Sande CJ, Thompson AJ, de Lara CM, Taylor KS, *et al.* Chimpanzee adenovirus- and MVA-vectored respiratory syncytial virus vaccine is safe and immunogenic in adults. Sci Transl Med **2015**;7(300):300ra126 doi 10.1126/scitranslmed.aac5745.
7. Sheehy SH, Duncan CJ, Elias SC, Choudhary P, Biswas S, Halstead FD, *et al.* ChAd63-MVA-vectored blood-stage malaria vaccines targeting MSP1 and AMA1: assessment of efficacy against mosquito bite challenge in humans. Mol Ther **2012**;20(12):2355-68 doi 10.1038/mt.2012.223.
8. Swadling L, Capone S, Antrobus RD, Brown A, Richardson R, Newell EW, *et al.* A human vaccine strategy based on chimpanzee adenoviral and MVA vectors that primes, boosts, and sustains functional HCV-specific T cell memory. Sci Transl Med **2014**;6(261):261ra153 doi 10.1126/scitranslmed.3009185.
9. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science **2017**;357(6349):409-13 doi 10.1126/science.aan6733.
10. Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA Approval Summary: Pembrolizumab for the Treatment of Microsatellite Instability-High Solid Tumors. Clin Cancer Res **2019**;25(13):3753-8 doi 10.1158/1078-0432.CCR-18-4070.
11. Cortes-Ciriano I, Lee S, Park WY, Kim TM, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. Nat Commun **2017**;8:15180 doi 10.1038/ncomms15180.
12. Kim TM, Laird PW, Park PJ. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. Cell **2013**;155(4):858-68 doi 10.1016/j.cell.2013.10.015.
13. Maby P, Tougeron D, Hamieh M, Mlecnik B, Kora H, Bindea G, *et al.* Correlation between Density of CD8+ T-cell Infiltrate in Microsatellite Unstable Colorectal Cancers and Frameshift Mutations: A Rationale for Personalized Immunotherapy. Cancer Res **2015**;75(17):3446-55 doi 10.1158/0008-5472.CAN-14-3051.
14. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, *et al.* An immunogenic personal neoantigen vaccine for patients with melanoma. Nature **2017**;547(7662):217-21 doi 10.1038/nature22991.

15. Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, *et al.* Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. Lancet Oncol **2017**;18(8):1009-21 doi 10.1016/S1470-2045(17)30516-8.
16. Chen W, Swanson BJ, Frankel WL. Molecular genetics of microsatellite-unstable colorectal cancer for pathologists. Diagn Pathol **2017**;12(1):24 doi 10.1186/s13000-017-0613-8.
17. Hampel H, de la Chapelle A. How do we approach the goal of identifying everybody with Lynch syndrome? Fam Cancer **2013**;12(2):313-7 doi 10.1007/s10689-013-9611-5.
18. Thiel T, Michalek W, Varshney RK, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). Theor Appl Genet **2003**;106(3):411-22 doi 10.1007/s00122-002-1031-0.
19. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. Nature **2016**;536(7616):285-91 doi 10.1038/nature19057.
20. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods **2015**;12(4):357-60 doi 10.1038/nmeth.3317.
21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics **2009**;25(16):2078-9 doi 10.1093/bioinformatics/btp352.
22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res **2010**;20(9):1297-303 doi 10.1101/gr.107524.110.
23. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol **2013**;31(3):213-9 doi 10.1038/nbt.2514.
24. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res **2012**;22(3):568-76 doi 10.1101/gr.129684.111.
25. Fang H, Bergmann EA, Arora K, Vacic V, Zody MC, Iossifov I, *et al.* Indel variant analysis of short-read sequencing data with Scalpel. Nat Protoc **2016**;11(12):2529-48 doi 10.1038/nprot.2016.150.
26. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res **2019**;47(8):e47 doi 10.1093/nar/gkz114.
27. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics **2014**;30(23):3310-6 doi 10.1093/bioinformatics/btu548.
28. Bai Y, Wang D, Fury W. PHLAT: Inference of High-Resolution HLA Types from RNA and Whole Exome Sequencing. Methods Mol Biol **2018**;1802:193-201 doi 10.1007/978-1-4939-8546-3_13.
29. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics **2016**;32(4):511-7 doi 10.1093/bioinformatics/btv639.

30. D'Alise AM, Leoni G, Cotugno G, Troise F, Langone F, Fichera I, *et al.* Adenoviral vaccine targeting multiple neoantigens as strategy to eradicate large tumors combined with checkpoint blockade. Nat Commun **2019**;10(1):2688 doi 10.1038/s41467-019-10594-2.
31. Lonchay C, van der Bruggen P, Connerotte T, Hanagiri T, Coulie P, Colau D, *et al.* Correlation between tumor regression and T cell responses in melanoma patients vaccinated with a MAGE antigen. Proc Natl Acad Sci U S A **2004**;101 Suppl 2:14631-8 doi 10.1073/pnas.0405743101.
32. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature **2012**;487(7407):330-7 doi 10.1038/nature11252.
33. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. Nature **2014**;513(7517):202-9 doi 10.1038/nature13480.
34. Cancer Genome Atlas Research N, Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, *et al.* Integrated genomic characterization of endometrial carcinoma. Nature **2013**;497(7447):67-73 doi 10.1038/nature12113.
35. Ottaviani S, Zhang Y, Boon T, van der Bruggen P. A MAGE-1 antigenic peptide recognized by human cytolytic T lymphocytes on HLA-A2 tumor cells. Cancer Immunol Immunother **2005**;54(12):1214-20 doi 10.1007/s00262-005-0705-2.
36. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat Rev Cancer **2019**;19(3):133-50 doi 10.1038/s41568-019-0116-x.
37. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, *et al.* Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. Mol Cancer Ther **2017**;16(11):2598-608 doi 10.1158/1535-7163.MCT-17-0386.
38. Ott PA, Dotti G, Yee C, Goff SL. An Update on Adoptive T-Cell Therapy and Neoantigen Vaccines. Am Soc Clin Oncol Educ Book **2019**;39:e70-e8 doi 10.1200/EDBK_238001.
39. Schrock AB, Ouyang C, Sandhu J, Sokol E, Jin D, Ross JS, *et al.* Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. Ann Oncol **2019**;30(7):1096-103 doi 10.1093/annonc/mdz134.
40. Woerner SM, Benner A, Sutter C, Schiller M, Yuan YP, Keller G, *et al.* Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. Oncogene **2003**;22(15):2226-35 doi 10.1038/sj.onc.1206421.
41. Woerner SM, Yuan YP, Benner A, Korff S, von Knebel Doeberitz M, Bork P. SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. Nucleic Acids Res **2010**;38(Database issue):D682-9 doi 10.1093/nar/gkp839.
42. Kloor M, von Knebel Doeberitz M. The Immune Biology of Microsatellite-Unstable Cancer. Trends Cancer **2016**;2(3):121-33 doi 10.1016/j.trecan.2016.02.004.
43. Westdorp H, Fennemann FL, Weren RD, Bisseling TM, Ligtenberg MJ, Figdor CG, *et al.* Opportunities for immunotherapy in microsatellite instable colorectal cancer. Cancer Immunol Immunother **2016**;65(10):1249-59 doi 10.1007/s00262-016-1832-7.
44. Pan J, Zhang Q, Sei S, Shoemaker RH, Lubet RA, Wang Y, *et al.* Immunoprevention of KRAS-driven lung adenocarcinoma by a multipeptide vaccine. Oncotarget **2017**;8(47):82689-99 doi 10.18632/oncotarget.19831.

45. Ozcan M, Janikovits J, von Knebel Doeberitz M, Kloor M. Complex pattern of immune evasion in MSI colorectal cancer. Oncoimmunology **2018**;7(7):e1445453 doi 10.1080/2162402X.2018.1445453.
46. Middha S, Yaeger R, Shia J, Stadler ZK, King S, Guercio S, *et al.* Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High. JCO Precis Oncol **2019**;3 doi 10.1200/PO.18.00321.
47. Chang K, Taggart MW, Reyes-Uribe L, Borras E, Riquelme E, Barnett RM, *et al.* Immune Profiling of Premalignant Lesions in Patients With Lynch Syndrome. JAMA Oncol **2018**;4(8):1085-92 doi 10.1001/jamaoncol.2018.1482.
48. Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, *et al.* Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet **2011**;378(9809):2081-7 doi 10.1016/S0140-6736(11)61049-0.

663 Tables

664

Tumor	Patients	% tumors sharing ≥ 400aa with Nous-209	Cumulative length of Nous-209 FSPs/patient (aa)	Nous-209 FSPs/patient
CRC	69	98	1322	50
GC	85	95	1178	46
EC	166	70	512	21

665

666 **Table 1: Coverage of MSI tumors in TCGA database with Nous-209.** In table is
667 indicated, for each MSI tumor type, the percentage of tumor samples having at least 400
668 amino acids (aa) in common with Nous-209, the cumulative length (median) of all FSPs
669 encoded by FSMs detected in tumor DNA, the number of FSPs shared with Nous-209.

670

Figures

Figure 1: Selection and prioritization of FSPs for Nous-209 vaccine. A, Filtering procedure to select FSPs included in Nous-209 vaccine. **B,** Venn diagrams showing the overlap of the initial collection of 1059 FSMs corresponding to 1087 FSPs, among the three MSI tumor types. **C,** Venn diagram showing the overlap among the three tumor types of the 204 FSMs coding for the 209 FSPs in Nous-209 vaccine.

Figure 2: Validation in MSI patients. A, Number of FSMs in Nous-209 that were detected by the look-up analysis performed on Exomeseq of tumor and healthy tissue from 20 MSI CRC patients. **B,** Cumulative length in amino acids (aa) of the FSPs present in DNA of tumor tissues that are in common with Nous-209. **C,** Number of FSMs for which evidence of expression was found in RNAseq of tumor tissue. In dark red is reported the number of FSMs expressed (TPM ≥ 2, mutated reads in RNAseq ≥ 1). **D,** Cumulative length in amino acids (aa) of the FSPs encoded by the FSMs expressed in tumor tissue.

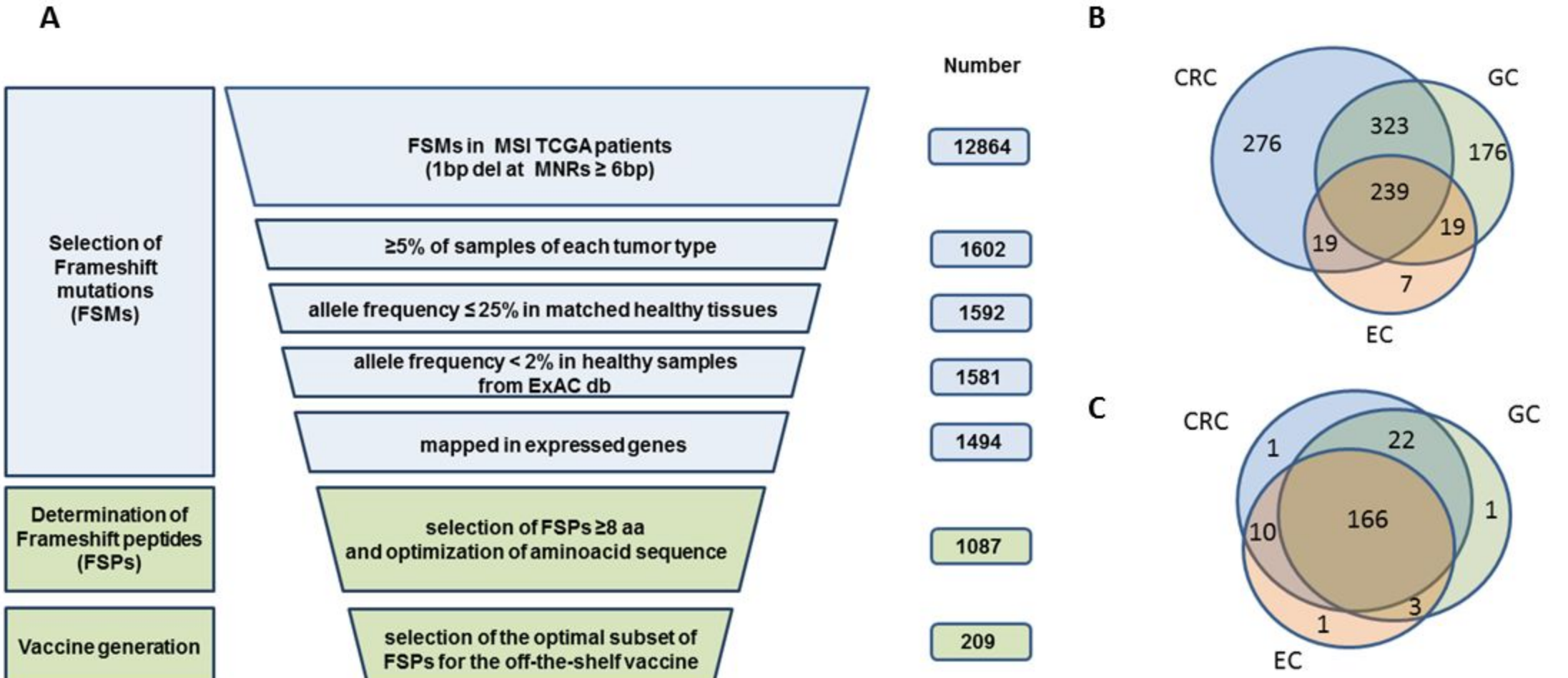
Figure 3: Nous-209 predicted MHC binders in 20 MSI CRC patients. A, Number of predicted CD8 binder peptides in FSMs detected in tumor DNA (orange bars) or expressed in tumor RNA (blue bars; TPM ≥ 2, mutated reads in RNAseq ≥ 1) of 20 MSI CRC patients**. B,** Number of predicted CD4 binder peptides in FSMs detected in tumor DNA (orange bars) or expressed in tumor RNA (blue bars; TPM ≥ 2, mutated reads in RNAseq ≥ 1).

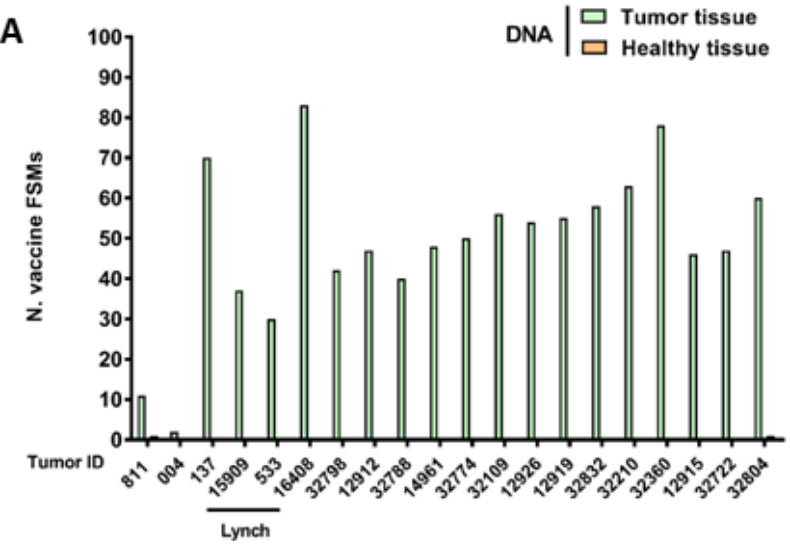
Figure 4: Immunogenicity of Nous-209 vaccine in mice. A, Schematic of immunization schedules and animal groups used for experiments in panels B and C. Nous-209

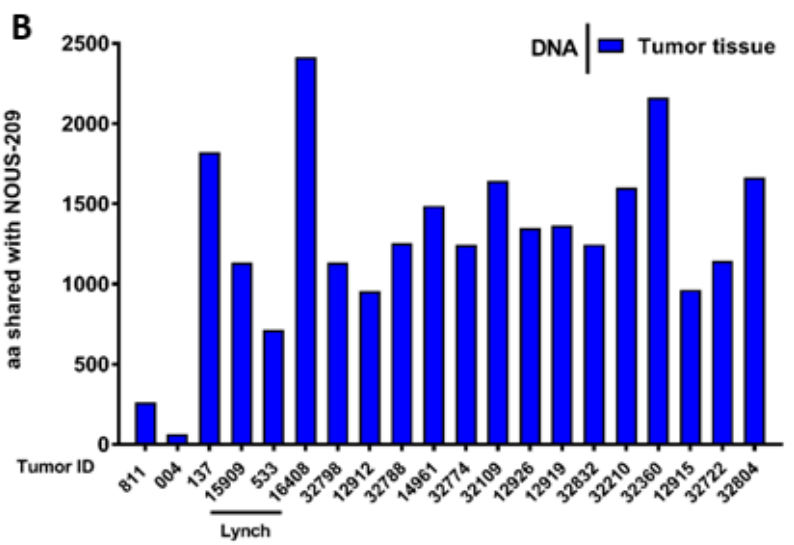
immunogenicity was measured on splenocytes of vaccinated mice by IFN-γ ELISpot at 2 weeks post prime with the 4 GAd vectors (GAd-209-FSP) and 1 week post boost with the 4 MVA vectors (MVA-209-FSP). **B,** T cell responses after prime and prime/boost are expressed as the sum of IFN-γ Spot Forming Cells (SFC)/10^6 splenocytes to the 16 pools of synthetic peptides covering the entire sequence of 209 FSPs. Each bar represents the average + standard error mean (SEM) responses in 6 animals. **C,** responses to each of the 16 peptide pools after prime/boost are shown. **D,** IFN-γ ELISpot responses to peptide pools 1 to 4 (covering FSP-1 vector) were measured post prime (week2) and post boost (week3) in mice immunized with a single vector (FSP-1) or the mix of 4 vectors (Mix). Responses are expressed as the number of IFN-γ-producing T cells per millions of splenocytes. Statistics are calculated by non-parametric Mann-Whitney U test (ns: $p > 0.05$; **$p < 0.01$). **E,** The percentages of $CD4^+$/IFN-γ^+ (gray) and CD8/IFN-γ^+ (black) FSP-specific T cells responses to peptide pools representative of FSP1 vector were measured on splenocytes of mice vaccinated with GAd-FSP-A1/MVA-FSP-B1.

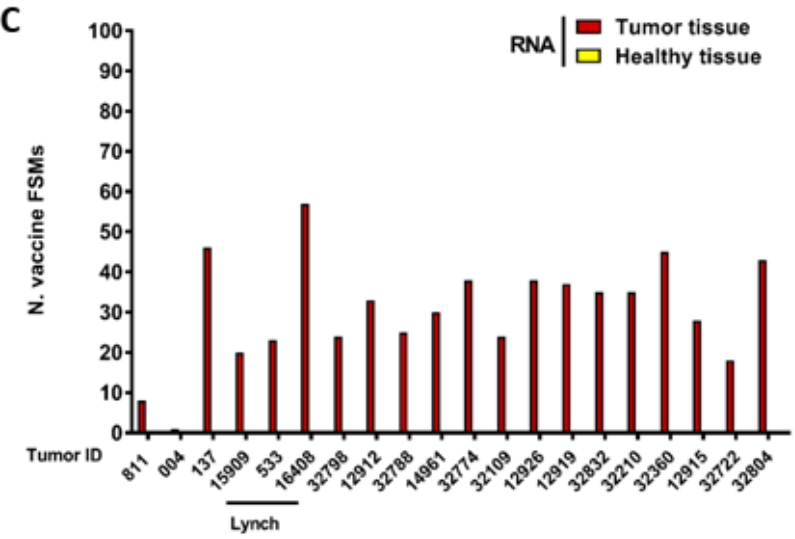
Figure 5: Isolation and functional analysis of FSP-specific human T cell clones. A, T cells amplified by weekly rounds of stimulation with peptide-pulsed T2 cells and EBV-B feeders cells were labelled with the fluorescent HLA-A2 multimers and CD8ß. Binding to PE-labelled HLA-A2 multimers folded to the Ly4_24 peptide or to a control peptide is shown for the three CD8 T cell clones (D5, H3 and F1) recognizing the FSP-derived epitope Ly4_24. **B,** Functional analysis of T cell clones. T cells were stimulated with HLA-A2 EBV-B cells without any peptide, pulsed with the Ly4_24 peptide, or infected with the MVA-FSP-B2 carrying the sequence coding for Ly4_24 peptide. The IFN-g secreted in the supernatants after 20h of co-culturing was measured by ELISA. Results are representative of 3 independent experiments and the mean ± s.d is shown. **C,** Gating

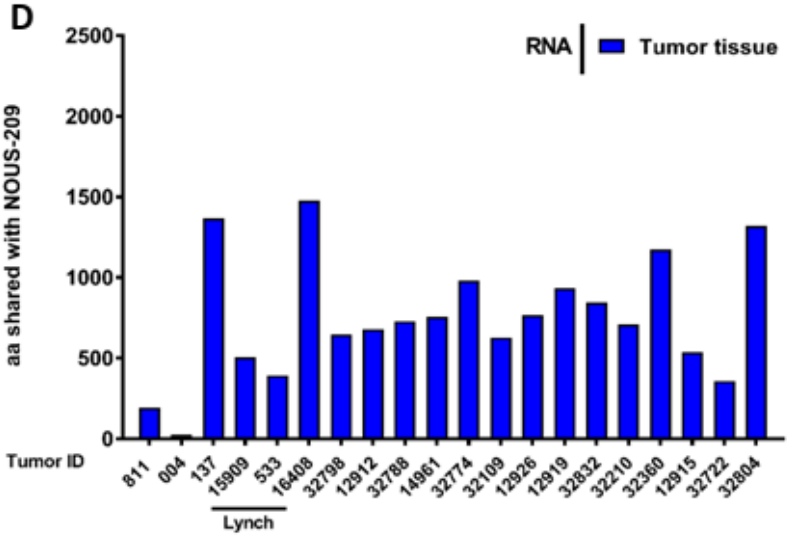
strategy used to obtain the data shown in A. Live T cells were selected on the FSC-SSC dotplot, the singlets were subsequently selected to finally select only the CD8ß-positive cells.

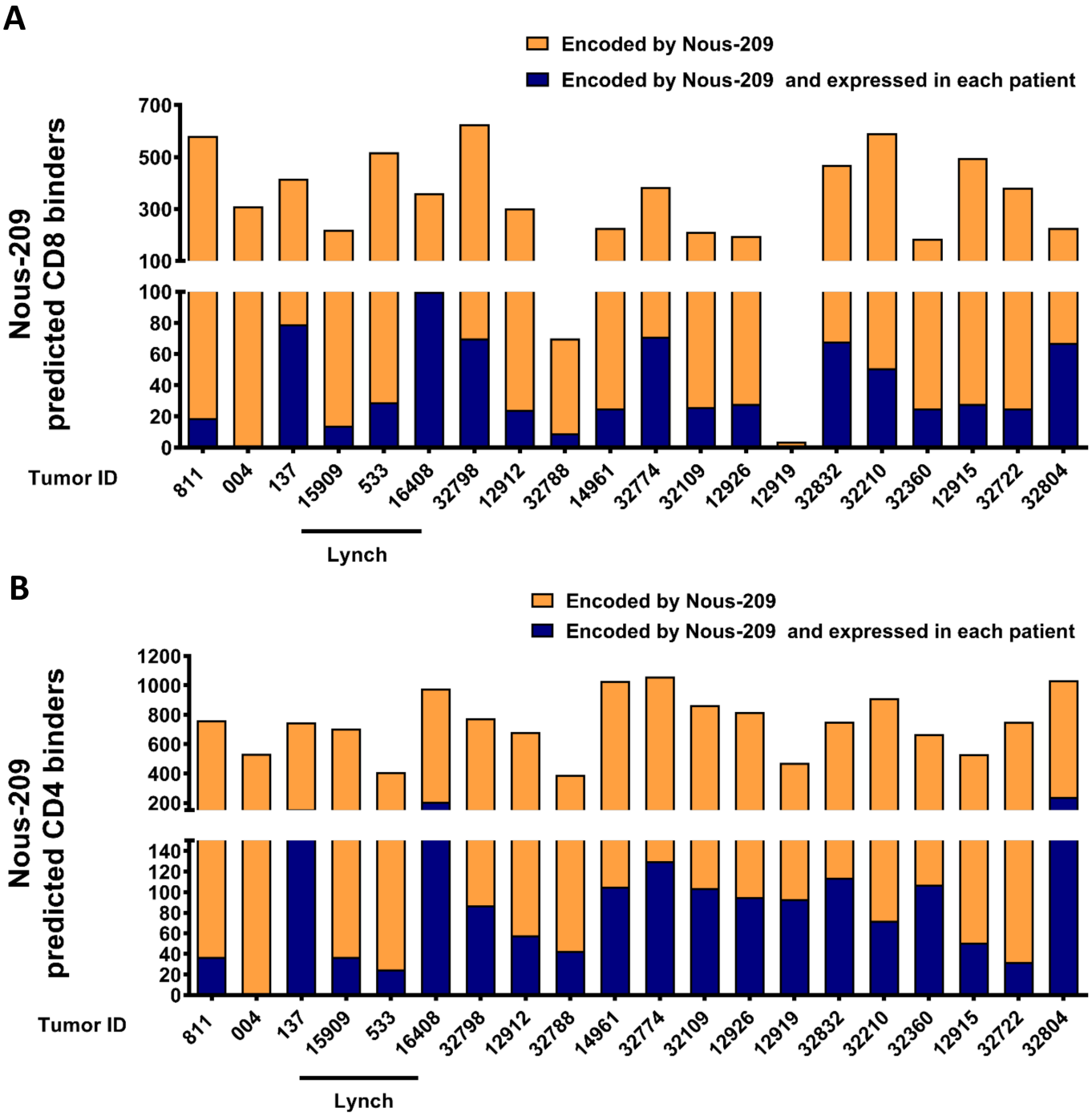

A
Number
Selection of Frameshift mutations (FSMs)
FSMs in MSI TCGA patients (1bp del at MNRs ≥ 6bp)
12864
≥5% of samples of each tumor type
1602
allele frequency ≤ 25% in matched healthy tissues
1592
allele frequency < 2% in healthy samples from ExAC db
1581
mapped in expressed genes
1494
Determination of Frameshift peptides (FSPs)
selection of FSPs ≥8 aa and optimization of aminoacid sequence
1087
Vaccine generation
selection of the optimal subset of FSPs for the off-the-shelf vaccine
209
B
CRC
GC
EC
276
323
176
239
19
19
7
C
CRC
GC
EC
1
22
1
166
10
3
1

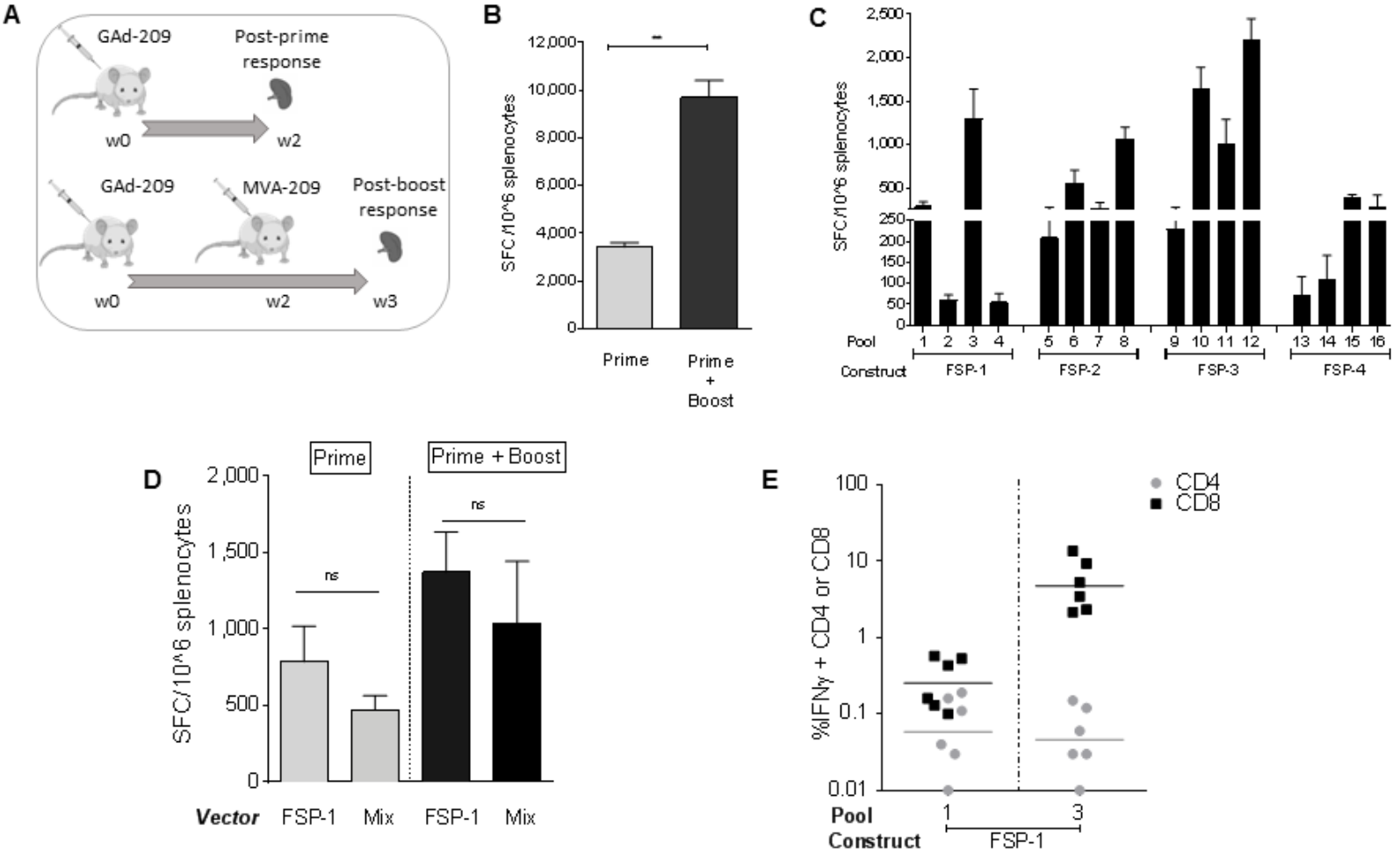

A
DNA
Tumor tissue
Healthy tissue
N. vaccine FSMs
100
90
80
70
60
50
40
30
20
10
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch

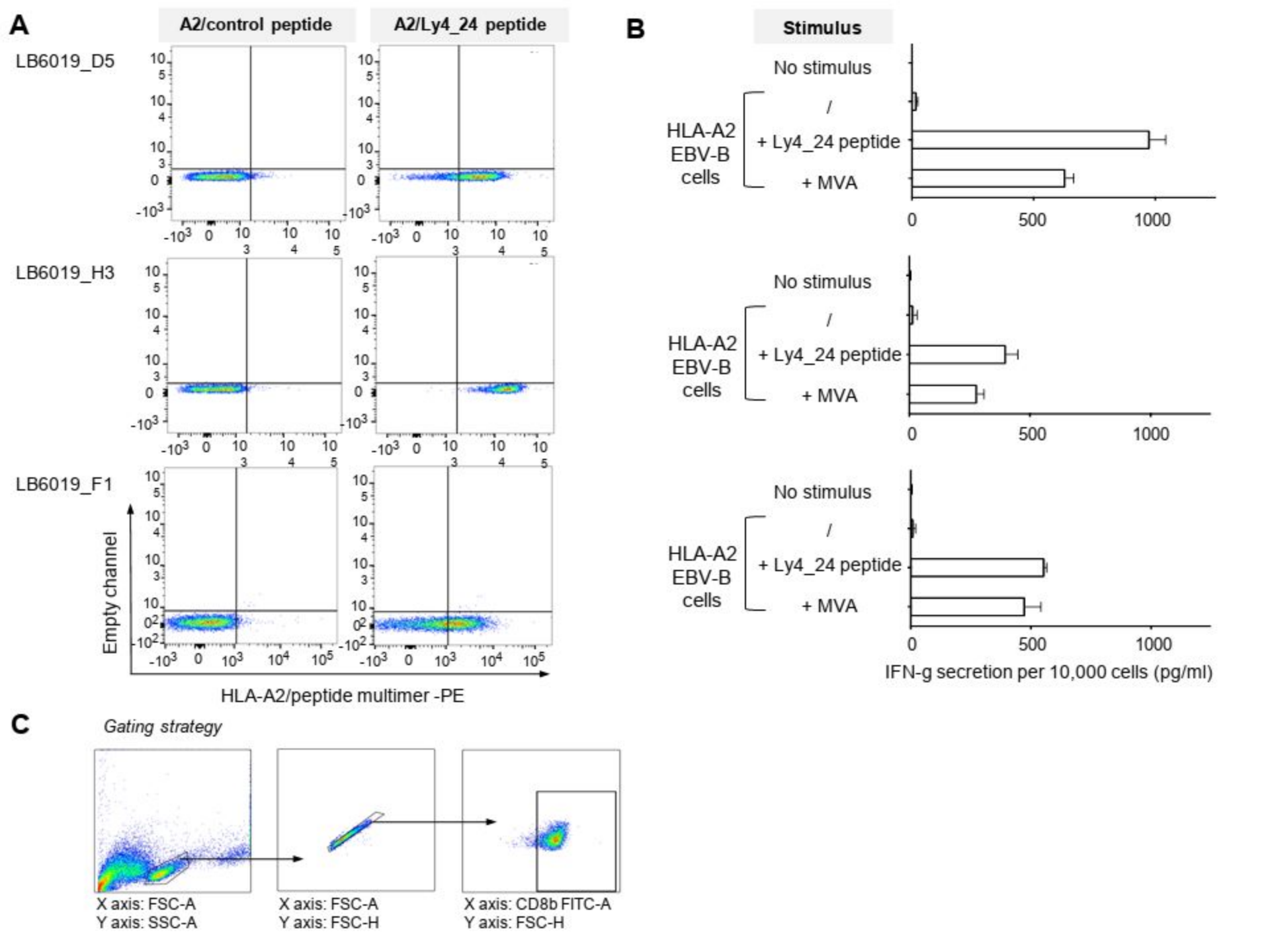

B
DNA
Tumor tissue
aa shared with NOUS-209
2500
2000
1500
1000
500
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch


C
RNA
Tumor tissue
Healthy tissue
N. vaccine FSMs
100
90
80
70
60
50
40
30
20
10
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch


D
RNA
Tumor tissue
aa shared with NOUS-209
2500
2000
1500
1000
500
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch


A
Encoded by Nous-209
Encoded by Nous-209 and expressed in each patient
Nous-209 predicted CD8 binders
700
500
300
100
100
80
60
40
20
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch
B
Encoded by Nous-209
Encoded by Nous-209 and expressed in each patient
Nous-209 predicted CD4 binders
1200
1000
800
600
400
200
140
120
100
80
60
40
20
0
Tumor ID
811
004
137
15909
533
16408
32798
12912
32788
14961
32774
32109
12926
12919
32832
32210
32360
12915
32722
32804
Lynch


A
GAd-209
Post-prime response
w0
w2
GAd-209
MVA-209
Post-boost response
w0
w2
w3
B
SFC/10^6 splenocytes
12,000
10,000
8,000
6,000
4,000
2,000
0
Prime
Prime + Boost
C
SFC/10^6 splenocytes
2,500
2,000
1,500
1,000
500
250
200
150
100
50
0
Pool
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Construct
FSP-1
FSP-2
FSP-3
FSP-4
D
Prime
Prime + Boost
SFC/10^6 splenocytes
2,000
1,500
1,000
500
0
ns
ns
Vector
FSP-1
Mix
FSP-1
Mix
E
%IFNγ + CD4 or CD8
100
10
1
0.1
0.01
CD4
CD8
Pool
1
3
Construct
FSP-1


A
A2/control peptide
A2/Ly4_24 peptide
LB6019_D5
LB6019_H3
LB6019_F1
Empty channel
HLA-A2/peptide multimer -PE
B
Stimulus
No stimulus
/
+ Ly4_24 peptide
+ MVA
HLA-A2 EBV-B cells
0
500
1000
IFN-g secretion per 10,000 cells (pg/ml)
C
Gating strategy
X axis: FSC-A
Y axis: SSC-A
X axis: FSC-A
Y axis: FSC-H
X axis: CD8b FITC-A
Y axis: FSC-H