Identification of bacteria-derived HLA-bound peptides in melanoma

Article

Identification of bacteria-derived HLA-bound peptides in melanoma

https://doi.org/10.1038/s41586-021-03368-8

Received: 4 November 2019

Accepted: 17 February 2021

Published online: 17 March 2021

Check for updates

Shelly Kalaora[1,27], Adi Nagler[1,27], Deborah Nejman[1], Michal Alon[1], Chaya Barbolin[1], Eilon Barnea[2], Steven L. C. Ketelaars[3], Kuoyuan Cheng[4], Kevin Vervier[5], Noam Shental[6], Yuval Bussi[1,7], Ron Rotkopf[8], Ronen Levy[1], Gil Benedek[9], Sophie Trabish[1], Tali Dadosh[10], Smadar Levin-Zaidman[10], Leore T. Geller[1], Kun Wang[4], Polina Greenberg[1], Gal Yagel[1], Aviyah Peri[1], Garold Fuks[11], Neerupma Bhardwaj[12], Alexandre Reuben[13], Leandro Hermida[4], Sarah B. Johnson[13,14], Jessica R. Galloway-Peña[15], William C. Shropshire[16], Chantale Bernatchez[17], Cara Haymaker[17], Reetakshi Arora[13,14], Lior Roitman[1], Raya Eilam[18], Adina Weinberger[1,7], Maya Lotan-Pompan[1,7], Michal Lotem[19], Arie Admon[2], Yishai Levin[20], Trevor D. Lawley[5], David J. Adams[5], Mitchell P. Levesque[21], Michal J. Besser[22,23], Jacob Schachter[22,24], Ofra Golani[8], Eran Segal[1,7], Naama Geva-Zatorsky[12,25], Eytan Ruppin[4], Pia Kvistborg[3], Scott N. Peterson[26], Jennifer A. Wargo[13,14], Ravid Straussman[1] & Yardena Samuels[1]✉

A variety of species of bacteria are known to colonize human tumours[1–11], proliferate within them and modulate immune function, which ultimately affects the survival of patients with cancer and their responses to treatment[12–14]. However, it is not known whether antigens derived from intracellular bacteria are presented by the human leukocyte antigen class I and II (HLA-I and HLA-II, respectively) molecules of tumour cells, or whether such antigens elicit a tumour-infiltrating T cell immune response. Here we used 16S rRNA gene sequencing and HLA peptidomics to identify a peptide repertoire derived from intracellular bacteria that was presented on HLA-I and HLA-II molecules in melanoma tumours. Our analysis of 17 melanoma metastases (derived from 9 patients) revealed 248 and 35 unique HLA-I and HLA-II peptides, respectively, that were derived from 41 species of bacteria. We identified recurrent bacterial peptides in tumours from different patients, as well as in different tumours from the same patient. Our study reveals that peptides derived from intracellular bacteria can be presented by tumour cells and elicit immune reactivity, and thus provides insight into a mechanism by which bacteria influence activation of the immune system and responses to therapy.

To investigate the influence of intratumoral bacteria on the immune response, we developed an experimental pipeline that comprises a 16S rRNA gene-sequencing platform[5] coupled to HLA peptidomics and applied it to samples from melanoma tumours.

Species of bacteria identified in melanoma

Our sequencing of the 16S rRNA gene of 17 melanoma samples derived from 9 patients with melanoma (Supplementary Table 1) led to the identification of 41 distinct species of bacteria (Supplementary Table 2). Inspection of a microbial phylogenetic tree revealed high similarity in the composition of bacteria found in different metastases from the same patient (Fig. 1, Extended Data Fig. 1), but also among samples from different patients. This finding points to the existence of species of bacteria that are common to melanoma. To validate the presence of bacteria in our tumour cohort, we performed 16S fluorescence in situ hybridization staining (Extended Data Fig. 2).

We complemented this with taxonomic profiling of a whole-genome sequencing dataset of melanoma that comprised 108 paired tumour and blood samples[15], focusing on DNA sequences that do not map to the human genome. Although the proportion of bacterial reads in the tumour and blood samples were the same (paired two-tailed Wilcoxon test, $P = 0.52$), the microbiota richness—measured as the number of species in a sample—was higher in tumour samples (paired two-tailed Wilcoxon test, $P = 0.0045$) (Extended Data Fig. 3a). Moreover, the bacterial composition was more conserved in tumour samples than in blood samples (Bray–Curtis distance, two-tailed Wilcoxon test, $P = 2.8 \times 10^{-8}$) (Extended Data Fig. 3b), and seven genera of bacteria exhibited greater abundance in the tumour samples (Extended Data Fig. 3c, d). Eight species of bacteria from the *Acinetobacter*, *Actinomyces*, *Corynebacterium*, *Enterobacter* and *Streptococcus* genera were also found in our cohort, supporting the microbial composition that we identified in melanoma tumours.

Presentation of bacterial peptides

We performed an HLA peptidome analysis of the HLA-I and HLA-II repertoires of the same tumours that were sequenced by 16S rRNA

A list of affiliations appears at the end of the paper.

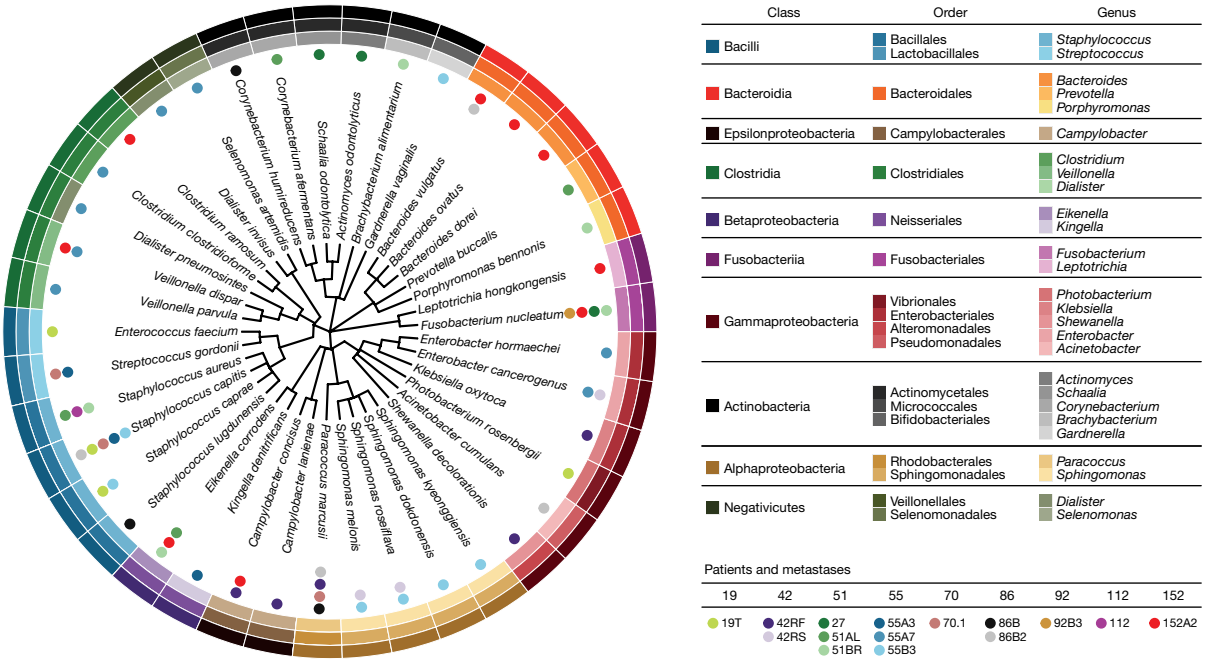


Fig. 1 | Identification of intratumoral bacteria in melanoma. Schematic phylogenetic tree of the bacterial composition of 17 melanoma metastases that originated from 9 patients. The analysis is based on rRNA 16S gene sequencing. The different colours and shades in the circles indicate the different classifications of bacteria at the genus (inner circle), order (middle circle) and class (outer circle) level. Each patient is colour-coded (as in the index), and different metastases from the same patient are depicted in different shades of the same colour.

profiling. The raw data from each HLA peptidome were searched (using MaxQuant software) against the proteomes of the species of bacteria that we identified in the corresponding tumour, together with the human proteome. We filtered the peptides by the quality of their identification and their ability to match the HLA alleles of the patient (Extended Data Fig. 4, Supplementary Information). This analysis revealed 248 unique HLA-I-associated peptides and 35 unique HLA-II-associated peptides (Supplementary Tables 3–5). The clustering of HLA-I peptides of 8–13 amino acids in length that were identified from each patient showed a reduced amino acid complexity of the peptides (as expected for HLA-I peptides), and matched the motifs of the HLA alleles of the patient (Supplementary Fig. 1). The length distribution of the identified peptides was consistent with the expected length of HLA-I and HLA-II peptides (Extended Data Fig. 5, Supplementary Fig. 2). The tandem mass spectra of all of the identified peptides is shown in Supplementary Fig. 3. We validated 48 of the identified peptides by comparing their tandem mass spectra to that of synthetic peptides (Supplementary Fig. 3).

In total, we obtained between 0 and 16 HLA-I and HLA-II peptides from each metastasis, and between 0 and 45 different HLA-I and HLA-II peptides from each species of bacteria (Fig. 2a). We identified 11 recurrent HLA-I-associated peptides that were derived from the bacteria *Fusobacterium nucleatum*, *Staphylococcus aureus* and *Staphylococcus capitis*. Five of the peptides appeared in different metastases from the same patient, and six appeared in different patients (Fig. 2b, Supplementary Table 7). As expected, recurrent peptides shared by patients were predicted to bind to HLA alleles shared by these patients or to contain the same HLA binding motif.

The percentage of bacterial peptides per sample that matched the *HLA-C*03:04*, *HLA-C*03:03* and *HLA-A*02:01* alleles was higher in most samples compared to the percentage of human peptides that matched the same alleles (Fig. 2c, Supplementary Fig. 4). The percentage of bacterial peptides was higher in HLA-C*03:04 and HLA-C*03:03 even when considering the RNA expression of the HLA molecules (Supplementary Fig. 5). When compared to the repertoire of the human-derived HLA peptidome, bacteria-derived peptides were significantly more hydrophobic (unpaired two-sample Wilcoxon test, $P = 3.71 \times 10^{-46}$) (Supplementary Figs. 5, 6, Supplementary Table 8), a property that might make these peptides preferable for antigen presentation[16] and recognition by T cells[17]. Importantly, we found that bacterial proteomes contained a significantly higher fraction of hydrophobic amino acids compared to the human proteome (Extended Data Fig. 6) and that the HLA-C*03:04, HLA-C*03:03 and HLA-A*02:01 alleles bind to more hydrophobic peptides (Extended Data Fig. 7), which provides an explanation for the higher frequency of bacterial peptides that are presented by these alleles.

As we used bulk melanoma tumours for the HLA peptidomics assay, we were unable to determine whether the bacterial HLA peptides were derived from the melanoma cells or antigen-presenting cells. To address this, we digested two melanoma tumours to recover single cells, which we subsequently separated into two populations on the basis of their CD45 marker: immune cells ($CD45^+$ cells) and nonimmune cells ($CD45^-$ cells, which are primarily melanoma cells) (tumours HG38 and 422 in Supplementary Fig. 7). As previously reported, the melanoma cells in our study not only expressed HLA-I but also HLA-II molecules[18–24] (Supplementary Fig. 8). We subjected the different populations isolated from tumour 422 to HLA peptidomics (Supplementary Table 9), which enabled us to identify both HLA-I and HLA-II peptides in the $CD45^-$ population (with some overlap between $CD45^-$ and $CD45^+$ populations) (Fig. 3a). Although it has previously been shown that bacteria can enter antigen-presenting cells and be presented by HLA molecules[25–28], we wanted to assess whether the bacteria that we identified in our cohort are also presented by antigen-presenting cells. To test this, we cocultured antigen-presenting cells, the B cell line IHW01070 and THP1 cells that were differentiated into macrophages using phorbol-12-myristate-13-acetate[29,30] with *F. nucleatum* and performed HLA peptidomics (Supplementary Fig. 8a–c, Supplementary Table 10). To ensure that the HLA peptides that we identify are indeed HLA ligands, we cocultured the HLA-I-null B cell line 721.221, with and without overexpression of the *HLA-A*01:01* allele, with *F. nucleatum* and performed HLA peptidomics (Supplementary Fig. 8d). As expected, we observed HLA-A*01:01 bacterial peptides only in the cells that overexpressed the HLA-A*01:01 allele; none of these peptides were identified in the 721.221 cells that

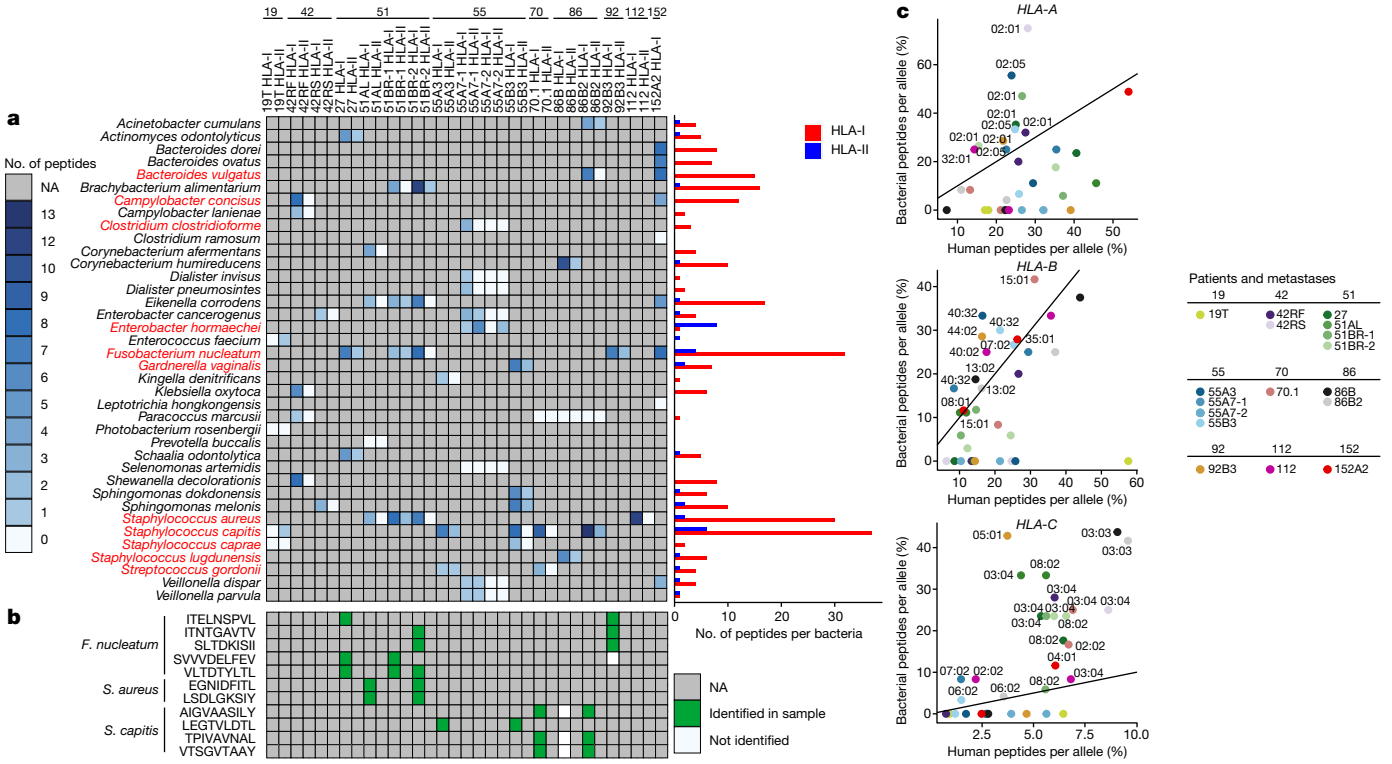


Fig. 2 | Characteristics of bacterial peptides. a, The number of bacterial peptides presented on HLA-I and HLA-II in each patient sample (patient number indicated at top) is indicated in a blue colour scale (left). White indicates that no peptides were identified in the sample, and grey indicates that the bacterium was not identified in this metastasis (NA, not applicable). The total number of bacterial HLA-I and HLA-II peptides from each bacterium is noted in the bar plot on the right. Species names marked in red are known to be intracellular bacteria (Supplementary Table 6). **b**, Bacterial peptides that were identified in a few metastases from the same patient or in different patients are indicated. Peptides identified in the sample are marked green, and white denotes peptides that were not identified in the sample (although the metastasis has the required HLA allele for this peptide presentation and the species of bacteria). Grey indicates samples that lack the HLA allele and bacteria to produce the peptide. **c**, For each metastasis, the percentages of bacterial and human peptides that match each *HLA-A* (left), *HLA-B* (middle) or *HLA-C* (right) allele of the patient is indicated. The allele with the best per cent rank binding prediction (by NetMHCpan) was assigned to each peptide; the full allele list is indicated in Extended Data Fig. 6.

did not overexpress HLA-A*01:01. The HLA-II presentation did not differ between the two lines (Supplementary Table 11).

Entry of bacteria into melanoma cells

To thoroughly evaluate whether melanoma cells can present bacterial HLA peptides, we first confirmed that the species of bacteria that we identified in our cohort of patients with melanoma are capable of entering melanoma cells. To this end, we cocultured low-passage cell lines derived from the same melanoma tumours with representative intracellular bacteria (bacteria for which we identified HLA peptides), and assessed the entry of bacteria into the cells using several orthogonal methodologies. We performed a gentamicin protection assay[31] on cocultures of 51AL and 55A3 melanoma cells with aerobically grown *Staphylococcus caprae* and anaerobically grown *Actinomyces odontolyticus*, and detected colony-forming units at both 4 and 8 h after infection (Extended Data Fig. 8a, b). As a control, we used *Lactobacillus animalis* grown under aerobic conditions and *Lactobacillus plantarum* grown under anaerobic conditions as representative species of bacteria. Both of these species are expected to possess weaker cell-invasion phenotypes and, using the gentamicin protection assay, we observed less invasion of these bacteria into 51AL and 55A3 cells compared to *S. caprae* and *A. odontolyticus* (Extended Data Fig. 8a, b, Supplementary Table 12).

We further verified the presence of bacteria within melanoma cells by performing immunofluorescence staining of 51AL and 55A3 GFP-expressing cells that were cocultured with *S. caprae*, using anti-lipoteichoic acid (Extended Data Fig. 9). Alternatively, we used copper-free click chemistry to label anaerobically grown *F. nucleatum* and *A. odontolyticus* with DIBO–Alexa Fluor 488[32], and cocultured these bacteria with 51AL and 55A3 cells stained with anti-HLA-I (Fig. 3b, Extended Data Fig. 10). The images presented in Extended Data Figs. 9, 10 are derived from the centre of *z*-stack images of the cells cocultured with the bacteria (Supplementary Video 1–8). We also constructed 3D representations of the *z*-stack images (Supplementary Videos 10–18). We detected all three species of bacteria in both of the melanoma cell lines. We performed correlative light and electron microscopy (CLEM) analysis of cells cocultured with *F. nucleatum*, *A. odontolyticus* and *S. caprae* and stained with anti-lipoteichoic acid, which confirmed that the bacteria indeed entered the melanoma cells (Fig. 3c, Extended Data Fig. 11).

Bacterial presentation is specific

We assessed whether we could use the bacteria that were identified in tumours to identify the bacteria-derived HLA-bound peptides of the corresponding patient in cocultures of 51AL cells with *F. nucleatum*, and 55A3 cells with *S. caprae*. Subjecting these cocultures to HLA peptidomics revealed 105 HLA-I and 130 HLA-II peptides in cells cocultured with bacteria (Supplementary Table 13). Knockout of the β-microglobulin (*B2M*) or class II major histocompatibility complex transactivator (*CIITA*) genes in the cells using genome editing reduced the expression of HLA-I and HLA-II, respectively (Supplementary Fig. 8e), and led to a decrease in the number of HLA-I- and HLA-II-associated peptides after coculture with bacteria compared to a scrambled non-targeting single-guide RNA control (Fig. 3d). Furthermore, coculturing the cells with *L. plantarum* or *L. animalis*—which are less cell-invasive than

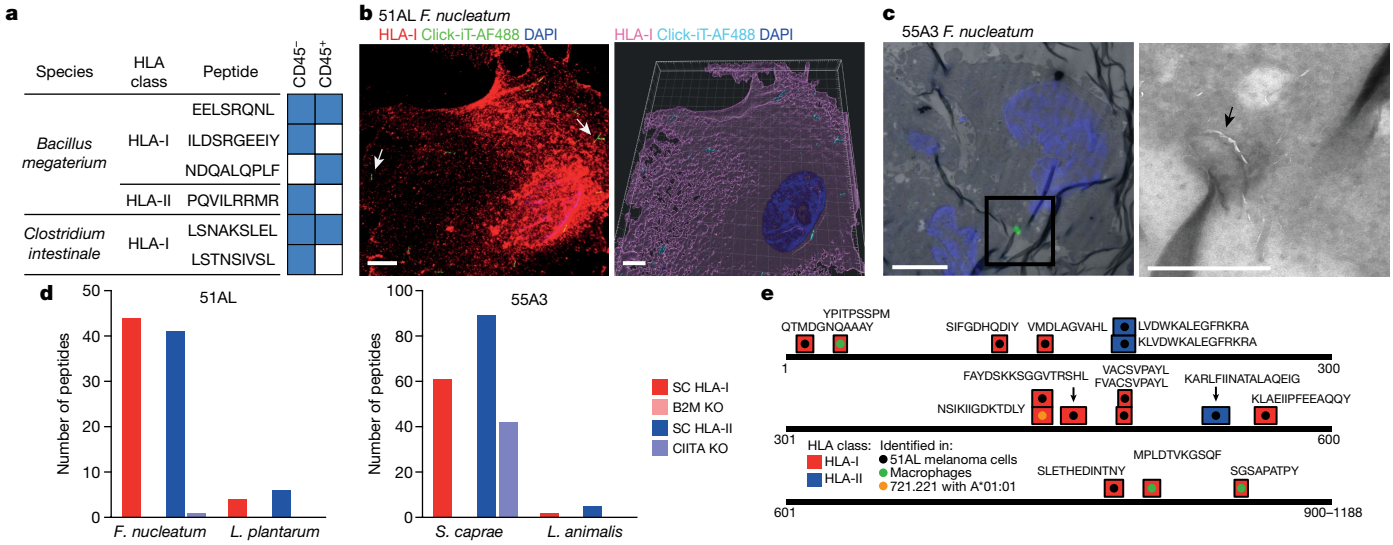


Fig. 3 | Evidence of bacteria entry and presentation by melanoma cells. **a**, Tumour 422 was digested and CD45$^+$ and CD45$^-$ populations were subjected to HLA peptidomics. The table lists the peptides identified in each sample. **b**, Immunofluorescence staining detecting bacteria in 51AL cells stained with an anti-HLA antibody (red) that were cocultured with *F. nucleatum* (green). Cell nuclei were stained with DAPI (blue). Left, representative merged image from the *z*-stack centre, presented at 63× magnification. Right, *z*-stack 3D image (bacteria, turquoise; melanoma cells, magenta; nucleus, blue). Scale bars, 10 µm. **c**, We co-incubated 55A3 cells with *F. nucleatum* (green). Cell nuclei were stained with DAPI (blue). Ultramicrotome sections were analysed by CLEM. Left, the area of intracellular bacteria is marked by a black box. Scale bar, 5 µm. Right, zoomed-in view of the bacteria. Scale bar, 1 µm. **d**, Number of bacterial peptides identified by HLA peptidomics in the 51AL and 55A3 cell lines that were cocultured with *F. nucleatum* and *S. caprae*, respectively. *B2M* and *CIITA* were knocked out (KO) to reduce the levels of HLA-I and HLA-II, respectively. Compared with cells infected with scrambled control single-guide RNA (SC), the cells with *B2M* or *CIITA* knocked out show a lower number of identified bacterial peptides. Less-invasive species of bacteria (*L. animalis* and *L. plantarum*) were used as a control to show that the identified peptides are HLA ligands that resulted from intracellular bacteria. **e**, Different peptides were derived from the pyruvate–ferredoxin (flavodoxin) oxidoreductase protein from *F. nucleatum*. The protein sequence is described until the 900th amino acid rather than the end of the protein (1,188 amino acids in total), as no peptides were identified after this region of the protein.

F. nucleatum and *S. caprae* (Extended Data Fig. 8a, b)—yielded fewer HLA-I- and HLA-II-bound peptides (Fig. 3d). To control for nonspecific contamination in the HLA immune purification process, we performed HLA peptidomics on bacterial pellets using the same amount as was used for the cocultures, and observed none or few peptides—none of which were identified in the coculture samples (Supplementary Table 14). Both the HLA-knockout and less-invasive-bacteria controls suggest that the bacterial peptides that we identified are HLA ligands, as their presence was strongly reduced in both control experiments.

Tumour-isolated bacterial presentation

To further validate the results of our cultured cell studies, we isolated bacteria from melanoma tumours and checked whether the specific isolated strains can invade melanoma cells. Our gentamicin protection assay and immunofluorescence staining of the isolated bacteria from tumours (*S. capitis* from tumour 58A and *Staphylococcus succinus* from tumour 261MS) revealed that both of these bacteria invade melanoma cells (Extended Data Figs. 8c, d, 9, Supplementary Videos 1–18). We cocultured *S. capitis* from tumour 58A with a melanoma-derived cell line produced from this tumour and performed HLA peptidomics on the cells, which led to the identification of 13 HLA-I and 11 HLA-II bacterial peptides (Supplementary Table 15).

For our HLA peptidomics analyses, we used published proteomes (from UniProt[33]) of the species of bacteria that we identified using 16S rRNA sequencing. We performed whole-genome sequencing of the bacteria isolated from the tumours and of the commercial bacteria used for the cocultures. We then constructed a proteome database from the whole-genome sequencing data and analysed the HLA peptidomics data in a similar manner to that from the UniProt database. The bacterial peptides identified in each of the experiments are shown in the peptide tables provided in Supplementary Tables 9–11, 13–15.

The overlap in the identified peptides between the two analyses was high in most experiments, which suggests that the use of published proteomes—even if they are not of the specific strain that was identified in the tumour—can give a sufficiently close peptide identification (Supplementary Table 16). This emphasizes that—although there is a clear advantage in using a database derived from the specific strain of bacteria isolated from the tumour to increase specificity—using a combination of 16S rRNA sequencing and the UniProt database avoids the need to isolate and sequence the bacteria from the tumour (which may be very difficult for many bacteria), thus increasing the number of bacterial HLA-bound peptides that it is possible to identify. A schematic overview of bacteria sequencing stages that we applied before the HLA-I and HLA-II peptidomics workflow is presented in Supplementary Fig. 9.

Recurrently presented bacterial peptides

We were able to identify two HLA-I peptides in the HLA peptidome of the cocultured cell lines that were also identified in the corresponding tumour samples: GLDLGTLTY, which was identified in metastasis 27 of patient 51, and GVDLGTLTY, which was identified in metastasis 51BR of the same patient—these peptides were noted also in 51AL cells that were cocultured with *F. nucleatum*. We identified the peptide ETTLVVTEY both in tumour 27 and in 721.221 cells that were cocultured with *F. nucleatum*. We found two additional peptides, one in the 51AL cell line (peptide IASDVSAIL) and one in macrophages and 721.221 cells (peptide NSIKIIGDKTDLY) cocultured with *F. nucleatum*. This demonstrates that some bacterial peptides can be presented by both melanoma cells and antigen-presenting cells, similar to what we observed in tumour 422 (in which we identified two peptides (EELSRQNL and LSNAKSLEL) that were present in both the CD45$^+$ and CD45$^-$ cell fractions) (Supplementary Table 9).

In addition, we identified 39 bacterial protein groups from which we received multiple bacterial peptides (2–15 different peptides per gene)

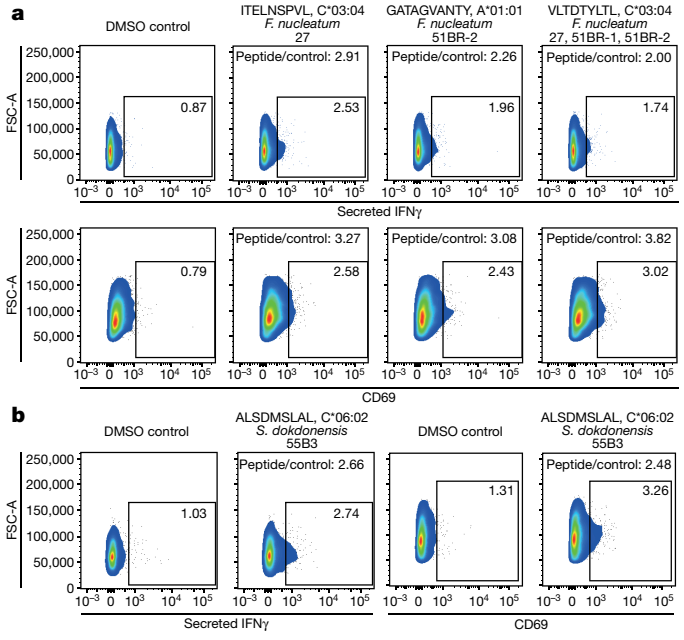


Fig. 4 | TIL reactivity towards bacteria-derived antigens. IFNγ-secreting 51AL and 55A3 TILs were detected after 6 h of coculture with B cells loaded with a bacterial peptide or dimethyl sulfoxide (DMSO) control, using flow cytometry. TILs were also tested for the presence of the CD69 reactivity marker. The image presents 4 representative immunogenic peptides (3 out of 7 immunogenic peptides of patient 51 and 1 out of 1 peptides of patient 55) that showed at least a 2-fold change between the peptide and DMSO control. The percentage of positive IFNγ-secreting or CD69-expressing TILs is an average of three independent experiments (Extended Data Fig. 12, Supplementary Figs. 10, 11). **a**, Patient 51. **b**, Patient 55.

(Supplementary Table 17). The bacterial protein with the corresponding highest number of different peptides was pyruvate–ferredoxin (flavodoxin) oxidoreductase from *F. nucleatum* (Fig. 3e). Twelve of these peptides were HLA-I peptides, and three were HLA-II peptides. Most of the peptides that we identified were from the 51AL cell line cocultured with *F. nucleatum*, although we also identified some in macrophages and the B cell line 721.221 overexpressing HLA-A*01:01. One peptide was found in both the 51AL melanoma cells and the 721.221 B cells, and two different pairs of peptides had an overlapping sequence. Other genes showed a higher number of overlapping peptides, which created presentation 'hot spots'. Among these genes was ATP synthase subunit-δ from *S. caprae* (from which 10 different HLA-II peptides were derived, protein group 9) (Supplementary Table 17).

Bacterial peptides are immunogenic

As recent studies have reported that tumour microbiota can affect responses to therapy by modulating the immune system[12–14] and as we identified antigens derived from bacteria on tumour HLA-I and HLA-II molecules, we hypothesized that intratumoural bacteria may not only shape the immune tumour microenvironment but also affect T cell immune reactivity.To evaluate this hypothesis, we analysed the reactivity of tumour-infiltrating lymphocytes (TILs) isolated from the analysed tumours towards the HLA-I-bound bacterial antigens that we identified. Specifically, we assessed the reactivity of bacterial peptides identified in the tumour that were derived from bacteria (such as *F. nucleatum* and *S. aureus*) that are known to negatively affect the host and immune system response by mechanisms other than peptide presentation[34–39], as well as the recurrently presented bacterial peptides (Supplementary Table 7), owing to their biological relevance and potential prevalence in the population. We pulsed synthetic peptides onto Epstein–Barr-virus-transformed B cells that expressed the tumour-matched HLA alleles. We then cocultured the loaded B cells with the autologous TILs, and analysed TIL reactivity by using flow cytometry to detect IFNγ-secreting TILs (Extended Data Fig. 12). We detected an increase in IFNγ-secreting TILs of twofold or more for eight different bacterial peptides, compared to control B cells that were not loaded with these peptides (Fig. 4, Extended Data Fig. 12, Supplementary Fig. 10). These included seven immunogenic peptides that were derived from patient 51 and one peptide that was derived from from patient 55. Of the reactive antigens, one peptide (VLTDTYLTL) was identified in two different metastases of patient 51, two peptides (ITELNSPVL and SLTDKISII in patients 51 and 92) were identified in two different patients, one peptide was also identified in the matching in vitro coculture of cells from the same patient (GVDLGTLTY in patient 51), and one peptide was identified from two potential species of bacteria that were identified in the same metastasis (ALSDMSLAL was derived from *Sphingomonas dokdonensis* or *Sphingomonas melonis*; both of these species were observed in metastasis 55B3). We confirmed these results by flow cytometry analysis of the CD69 T cell reactivity marker (Fig. 4, Supplementary Fig. 11).

Discussion

In summary, here we demonstrate the presence of HLA peptidomic signatures derived from bacteria in tumours from patients with melanoma, and characterize the species of bacteria involved. We demonstrate that bacterial HLA-I and HLA-II peptides can be presented by both antigen-presenting cells (as has previously been shown[25–28]) and by melanoma cells, by applying HLA peptidomics to a tumour sample that was separated into $CD45^-$ and $CD45^+$ populations as well as to antigen-presenting cells and melanoma cell lines that were cocultured with bacteria. Among the peptides that we identified were ones shared by different metastases from the same patient or by metastases from different patients. Some of these recurrent antigens bind to HLA alleles that are highly prevalent in the melanoma cohort of The Cancer Genome Atlas (Supplementary Fig. 12). Moreover, a few of the identified peptides that elicited an immune response from the autologous TILs of the patient are recurrent peptides. As the bacterial antigens are non-self, they could serve as targets for immunotherapy. The selection of the species of bacteria for immunotherapy should be carefully considered, and should favour species of bacteria that are known to negatively affect the response of the host and immune systems (rather than more 'protective' bacteria[40–42]). When focusing on the presented bacterial proteins, we observed genes that have overlapping peptides, which create presentation hots pots that possibly indicate the existence of protein domains and may further assist investigations into selecting the targets for immunotherapy.

Despite its strengths, our study has several shortcomings. Specifically, the resolution needed for the identification of specific species of bacteria is limited, owing to the high similarity of the 16S sequences between some species of bacteria. We also acknowledge the limitations of identifying bacterial peptides using HLA peptidomics, as it relies on peptide intensity: this may sometimes be below the detection level, especially when the amount of sample material is limited. This also explains the higher number of identified bacterial peptides in cell lines, as the amount of cells used was higher in these experiments. Additional studies that assess intratumoral bacteria-derived peptides—in combination with clinical and therapeutic information on large-scale cohorts—are likely to shed further light on the clinical role of bacterial peptides and to provide a higher-resolution view of the fundamental trends outlined in this Article.

Our comprehensive analysis of the intratumour microbiota in melanoma and the HLA peptides derived from these bacteria demonstrates that the bacteria that colonize melanoma tumours can enter melanoma cells, and that their peptides can be presented by the HLA-I and HLA-II

molecules of melanoma tumours. As both classes of HLA molecule have previously been shown to have a central role in $CD8^+$ and $CD4^+$ T cell immunity[43–49], they are expected ultimately to modulate immune function. Finally, as the gut and tumour microbiota can affect the survival of patients with cancer and their responses to therapy[8,14,40–42,50], our findings are of particular relevance, as they suggest that a mechanism that involves antigen presentation may underlie these effects.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03368-8.

1. Zheng, J. H. et al. Two-step enhanced cancer immunotherapy with engineered *Salmonella typhimurium* secreting heterologous flagellin. *Sci. Transl. Med.* **9**, eaak9537 (2017).
2. Silva-Valenzuela, C. A. et al. Solid tumors provide niche-specific conditions that lead to preferential growth of *Salmonella*. *Oncotarget* **7**, 35169–35180 (2016).
3. Geller, L. T. et al. Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science* **357**, 1156–1160 (2017).
4. Pushalkar, S. et al. The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. *Cancer Discov.* **8**, 403–416 (2018).
5. Nejman, D. et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* **368**, 973–980 (2020).
6. Castellarin, M. et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299–306 (2012).
7. Drewes, J. L. et al. High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes* **3**, 34 (2017).
8. Bullman, S. et al. Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. *Science* **358**, 1443–1448 (2017).
9. Hieken, T. J. et al. The microbiome of aseptically collected human breast tissue in benign and malignant disease. *Sci. Rep.* **6**, 30751 (2016).
10. Mrázek, J. et al. Melanoma-related changes in skin microbiome. *Folia Microbiol. (Praha)* **64**, 435–442 (2019).
11. Jin, C. C. et al. Commensal microbiota promote lung cancer development via γδ T cells. *Cell* **176**, 998–1013 (2019).
12. Kostic, A. D. et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**, 292–298 (2012).
13. Kostic, A. D. et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **14**, 207–215 (2013).
14. Riquelme, E. et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell* **178**, 795–806 (2019).
15. Hayward, N. K. et al. Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017).
16. Vigneron, N., Abi Habib, J. & Van den Eynde, B. J. Learning from the proteasome how to fine-tune cancer immunotherapy. *Trends Cancer* **3**, 726–741 (2017).
17. Chowell, D. et al. TCR contact residue hydrophobicity is a hallmark of immunogenic $CD8^+$ T cell epitopes. *Proc. Natl Acad. Sci. USA* **112**, E1754–E1762 (2015).
18. Deffrennes, V. et al. Constitutive expression of MHC class II genes in melanoma cell lines results from the transcription of class II transactivator abnormally initiated from its B cell-specific promoter. *J. Immunol.* **167**, 98–106 (2001).
19. Ruiter, D. J. et al. Immunohistochemical analysis of malignant melanomas and nevocellular nevi with monoclonal antibodies to distinct monomorphic determinants of HLA antigens. *Cancer Res.* **44**, 3930–3935 (1984).
20. Anichini, A. et al. Association of antigen-processing machinery and HLA antigen phenotype of melanoma cells with survival in American Joint Committee on Cancer stage III and IV melanoma patients. *Cancer Res.* **66**, 6405–6411 (2006).
21. Johnson, D. B. et al. Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. *Nat. Commun.* **7**, 10582 (2016).
22. Axelrod, M. L., Cook, R. S., Johnson, D. B. & Balko, J. M. Biological consequences of MHC-II expression by tumor cells in cancer. *Clin. Cancer Res.* **25**, 2392–2402 (2019).
23. Yue, F. Y. et al. Interleukin-10 is a growth factor for human melanoma cells and down-regulates HLA class-I, HLA class-II and ICAM-1 molecules. *Int. J. Cancer* **71**, 630–637 (1997).
24. Bernsen, M. R. et al. On the biological relevance of MHC class II and B7 expression by tumour cells in melanoma metastases. *Br. J. Cancer* **88**, 424–431 (2003).
25. Bettencourt, P. et al. Identification of antigens presented by MHC for vaccines against tuberculosis. *NPJ Vaccines* **5**, 2 (2020).
26. Pfeifer, J. D. et al. Phagocytic processing of bacterial antigens for class I MHC presentation to T cells. *Nature* **361**, 359–362 (1993).
27. Kovacsovics-Bankowski, M. & Rock, K. L. A phagosome-to-cytosol pathway for exogenous antigens presented on MHC class I molecules. *Science* **267**, 243–246 (1995).
28. Lewinsohn, D. M. et al. Characterization of human $CD8^+$ T cells reactive with *Mycobacterium tuberculosis*-infected antigen-presenting cells. *J. Exp. Med.* **187**, 1633–1640 (1998).
29. Genin, M., Clement, F., Fattaccioli, A., Raes, M. & Michiels, C. M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide. *BMC Cancer* **15**, 577 (2015).
30. Tsuchiya, S. et al. Induction of maturation in cultured human monocytic leukemia cells by a phorbol diester. *Cancer Res.* **42**, 1530–1536 (1982).
31. Elsinghorst, E. A. Measurement of invasion by gentamicin resistance. *Methods Enzymol.* **236**, 405–420 (1994).
32. Geva-Zatorsky, N. et al. In vivo imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria. *Nat. Med.* **21**, 1091–1100 (2015).
33. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
34. Gur, C. et al. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* **42**, 344–355 (2015).
35. Yang, Y. et al. *Fusobacterium nucleatum* increases proliferation of colorectal cancer cells and tumor development in mice by activating Toll-like receptor 4 signaling to nuclear factor-κB, and up-regulating expression of microRNA-21. *Gastroenterology* **152**, 851–866 (2017).
36. Parhi, L. et al. Breast cancer colonization by *Fusobacterium nucleatum* accelerates tumor growth and metastatic progression. *Nat. Commun.* **11**, 3259 (2020).
37. Menzies, B. E. & Kourteva, I. *Staphylococcus aureus* α-toxin induces apoptosis in endothelial cells. *FEMS Immunol. Med. Microbiol.* **29**, 39–45 (2000).
38. Esen, M. et al. Mechanisms of *Staphylococcus aureus* induced apoptosis of human endothelial cells. *Apoptosis* **6**, 431–439 (2001).
39. Lowy, F. D. *Staphylococcus aureus* infections. *N. Engl. J. Med.* **339**, 520–532 (1998).
40. Gopalakrishnan, V. et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).
41. Matson, V. et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104–108 (2018).
42. Routy, B. et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).
43. Spitzer, M. H. et al. Systemic immunity is required for effective cancer immunotherapy. *Cell* **168**, 487–502 (2017).
44. Kreiter, S. et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* **520**, 692–696 (2015).
45. Veatch, J. R. et al. Tumor-infiltrating $BRAF^{V600E}$-specific $CD4^+$ T cells correlated with complete clinical response in melanoma. *J. Clin. Invest.* **128**, 1563–1568 (2018).
46. Alspach, E. et al. MHC-II neoantigens shape tumour immunity and response to immunotherapy. *Nature* **574**, 696–701 (2019).
47. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
48. Ott, P. A. et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* **547**, 217–221 (2017).
49. Sahin, U. et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **547**, 222–226 (2017).
50. Vetizou, M. et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015).

[1]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. [2]Department of Biology, Technion – Israel Institute of Technology, Haifa, Israel. [3]Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. [4]Cancer Data Science Laboratory (CDSL), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA. [5]Wellcome Sanger Institute, Cambridge, UK. [6]Department of Mathematics and Computer Science, Open University of Israel, Raanana, Israel. [7]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. [8]Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel. [9]Tissue Typing and Immunogenetics Unit, Hadassah Medical Center, Jerusalem, Israel. [10]Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel. [11]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel. [12]The Ruth and Bruce Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel. [13]Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [14]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [15]Department of Veterinary Pathobiology, Texas A&M University, Houston, TX, USA. [16]Department of Infectious Diseases, MD Anderson Cancer Center, Houston, TX, USA. [17]Department of Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [18]Department of Veterinary Resources, Weizmann Institute of Science, Rehovot, Israel. [19]Sharett Institute of Oncology, Hadassah Hebrew University Medical Center, Jerusalem, Israel. [20]The de Botton Institute for Protein Profiling, The Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot, Israel. [21]Faculty of Medicine, University of Zurich Hospital, University of Zurich, Zurich, Switzerland. [22]The Ella Lemelbaum Institute for Immuno Oncology and Melanoma, Chaim Sheba Medical Center, Tel Hashomer, Israel. [23]Department of Clinical Microbiology and Immunology, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. [24]Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. [25]MaRS Centre, Canadian Institute for Advanced Research (CIFAR) Azrieli Global Scholar, Toronto, Ontario, Canada. [26]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA. [27]These authors contributed equally: Shelly Kalaora, Adi Nagler. ✉e-mail: yardena.samuels@weizmann.ac.il

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All raw mass spectrometry files, as well as human and bacteria proteomes and the MaxQuant version used for analysing the data, have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD022150. Raw sequence data of the whole-genome sequencing of isolated bacteria have been deposited in the NCBI Sequence Read Archive (SRA), under BioProject accession number PRJNA669827. Source data are provided with this paper.

Acknowledgements We thank the UT MDACC clinical TIL laboratory for processing the tumour specimens; E. Elinav and. J. Suez for providing bacteria cultures; S. Jung and S. Trzebanski for access to their flow cyometer; and S. Cheriyamundath for his help with confocal imaging. This work was supported by the Intramural Research Programs of the National Cancer Institute. Y.S. is supported by the Israel Science Foundation grant no. 696/17, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 770854), MRA (no. 622106), Rising Tide Foundation, Henry Chanoch Krenter Institute for Biomedical Imaging and Genomics, Estate of Alice Schwarz-Gardos, Estate of John Hunter, Knell Family, Peter and Patricia Gruber Award and the Hamburger Family. J.A.W is supported by generous philanthropic contributions to the University of Texas MD Anderson Moon Shots Program for support of tumor-line generation, from the Estate of Mady Dukler, the Joel and Mady Dukler Fund for Cancer Research, the Estate of Judith Safirstein, the Estate of Elaine S. Scheye, the Estate of David Levinson, the Hadar Impact Fund, the Fannie Sherr Fund, the Erica Drake Fund, the Estate of Bernard Berkowitz, the Bernard and Audrey Jaffe Foundation and the Jacques Asseoff Trust. K.V. and T.D.L. are supported by the Wellcome Sanger core funding (WT098051). The CLEM studies were conducted at the Irving and Cherna Moskowitz Center for Nano and Bio-Nano Imaging at the Weizmann Institute of Science. J.R.G.-P. is funded by (and whole-genome sequencing was paid for using) NIH NIAID K01AI143881. Whole-genome sequencing of bacteria isolates was performed by the MD Anderson Advanced Technology Genomics Core grant, funded by the CCSG NIH NCI grant CA016672. A.A. is supported by the Israel Science Foundation grant no. 1435/16. D.J.A., K.V. and T.D.L. are supported by the Wellcome Sanger core funding (WT098051).

Author contributions S.K., A.N. and Y.S. designed the study and wrote the paper. S.K. and A.N. analysed the data, and performed and supervised experiments. S.K., A.N., M.A., E.B. and P.G. performed the HLA peptidomics experiments. D.N., N.S., G.F. and R.S. analyzsed the bacteria composition of tumours. T.D. and S.L.-Z. performed the CLEM experiments. L.T.G., N.B., N.G.-Z., L.R., O.G. and R.E. helped with the bacteria staining techniques, imaging and imaging analysis. A.N., P.G., G.Y. and A.P. grew the cell lines for HLA peptidomics. K.V., T.D.L. and D.J.A. performed the taxonomic analysis of whole-genome sequencing data from melanoma. Y.B., W.C.S., A.W., M.L.-P. and E.S. performed whole-genome sequencing and analysis of bacteria isolates. J.R.G.-P. and S.N.P. performed bacteria isolation and sequencing. C. Barbolin, K.C., R.L., K.W., L.H. and E.R. performed computational analysis. R.R. helped with the statistical analysis. A.N. and S.T. performed the immunofluorescence experiments. M.L. helped with the TIL rapid expansion protocols. A.A. and Y.L. helped with the HLA peptidomics and mass spectrometry analyses. A.R., S.B.J., C. Bernatchez, C.H., R.A., M.P.L., M.J.B., J.S. and J.A.W. provided patient materials and information. S.L.C.K. and P.K. helped with TIL reactivity assay. G.B. performed HLA typing. All authors contributed to the final version of the paper.

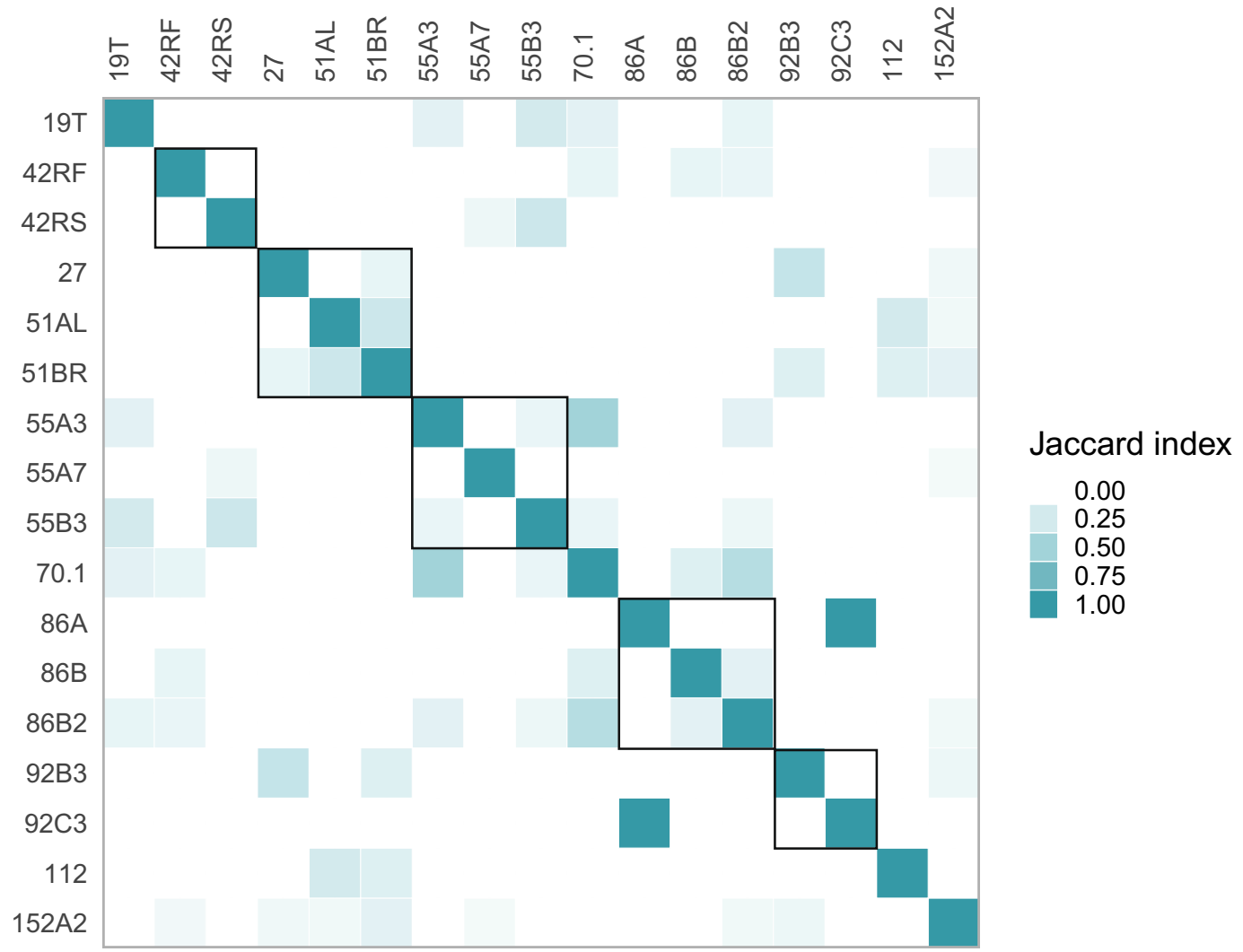
Competing interests J.A.W. is an inventor on a US patent application (PCT/US17/53.717) submitted by the University of Texas MD Anderson Cancer Center that covers methods to enhance immune checkpoint blockade responses by modulating the microbiome. J.A.W. reports compensation for speaker's bureau and honoraria from Imedex, Dava Oncology, Omniprex, Illumina, Gilead, PeerView, Physician Education Resource, MedImmune and Bristol–Myers Squibb. J.A.W. serves as a consultant/advisory board member for Roche/Genentech, Novartis, AstraZeneca, GlaxoSmithKline, Bristol–Myers Squibb, Merck, Biothera Pharmaceuticals and Microbiome DX. J.A.W. also receives research support from GlaxoSmithKline, Roche/Genentech, Bristol–Myers Squibb and Novartis.

Additional information

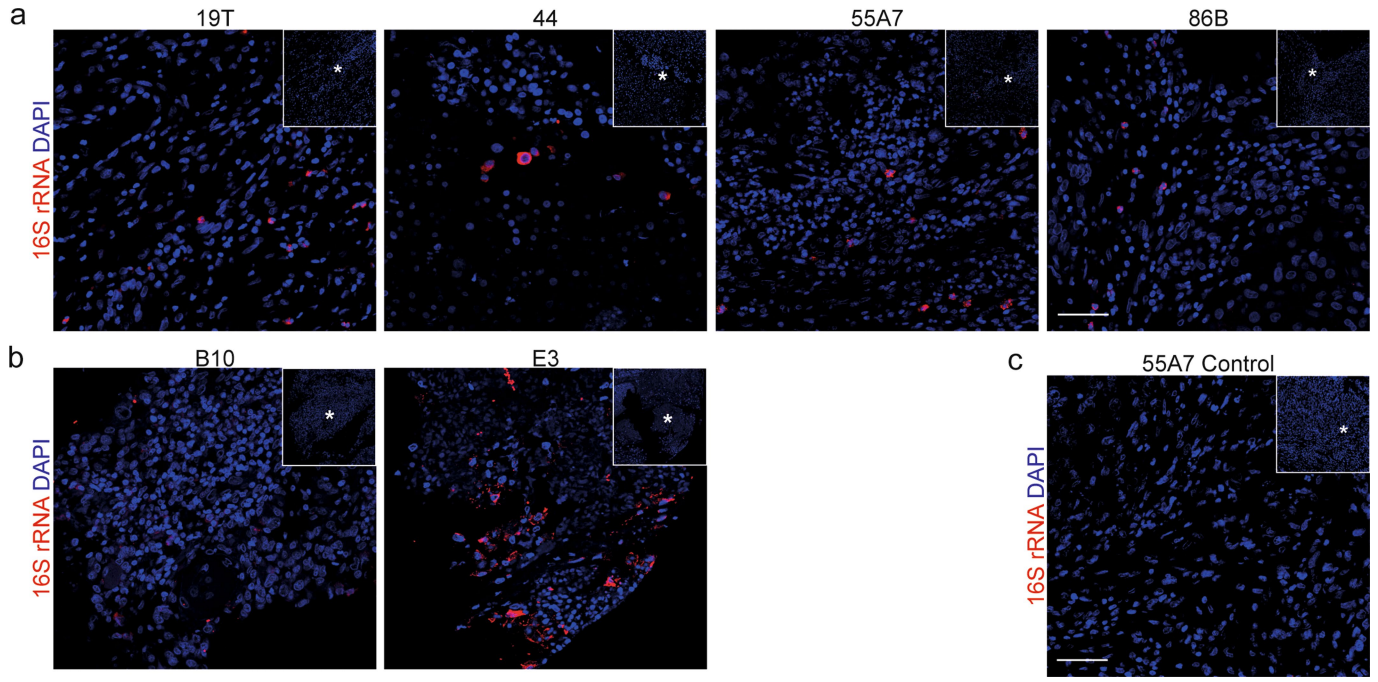
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03368-8.

Correspondence and requests for materials should be addressed to Y.S.

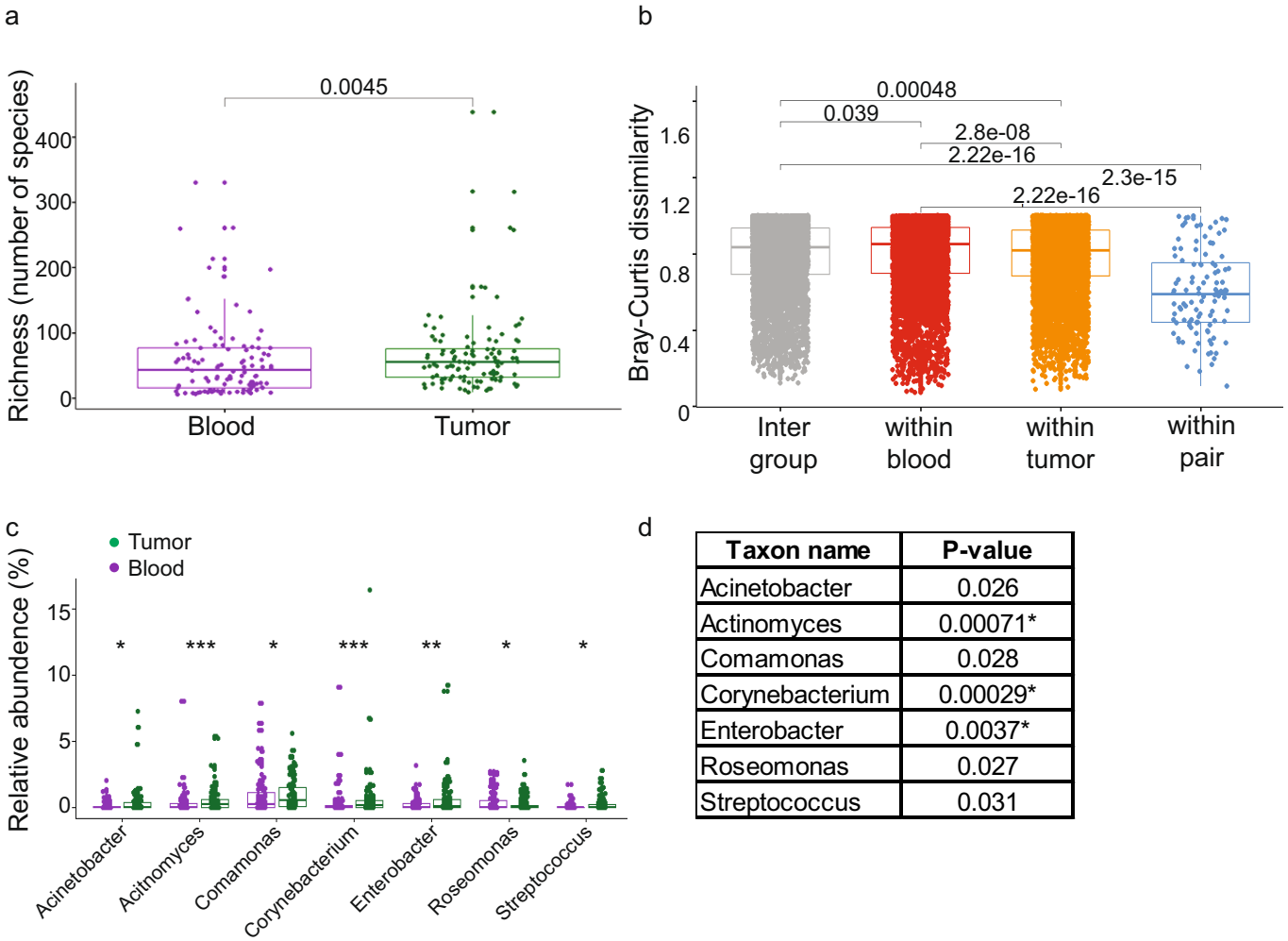
Peer review information *Nature* thanks Angelika Riemer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Similarity of bacterial composition between metastases. A Jaccard index was calculated to determine the similarity between the bacterial composition of the different metastases on the species level. Colour code indicates the Jaccard index. The highest similarity was observed between metastases from the same patient, but metastases of different patients also showed similarity. Black boxes indicate tumour samples taken from the same patient.

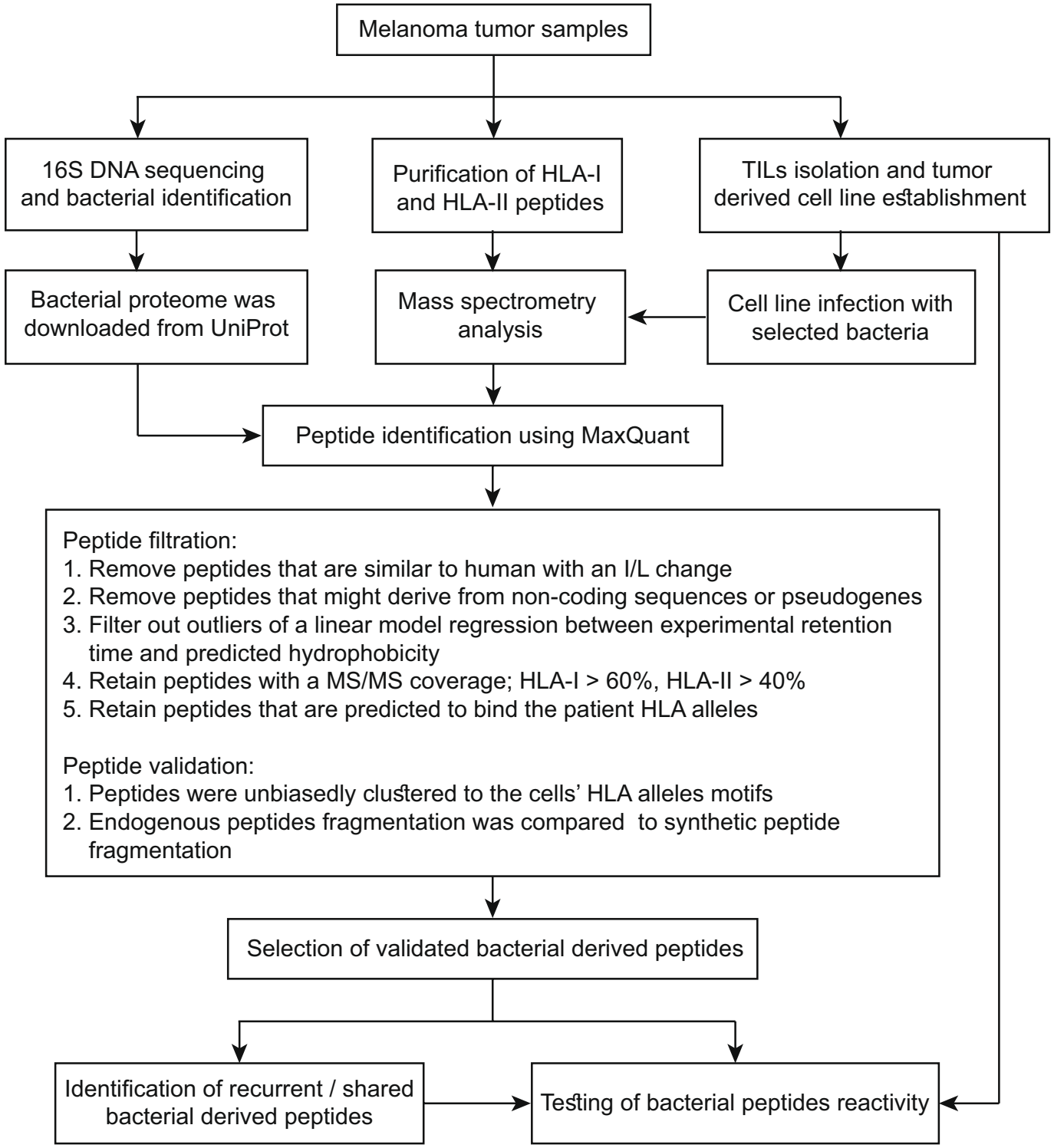


Extended Data Fig. 2 | Visualization of bacterial 16S rRNA in tissue sections from melanoma tumours. a, 16S rRNA fluorescence in situ hybridization (FISH) staining of tissue sections from melanoma tumours using pan-bacteria EUB338 probe (red) and DAPI (blue). **b**, 16S rRNA FISH staining of tissue microarray sections of melanoma tumours (red) and DAPI (blue). Slice name indicates the position in the tissue microarray. Images are presented at 20× magnification. Scale bars, 100 µm. **c**, Representative control using 16S FISH nonspecific control probe. Asterisks mark the region that was selected for higher magnification. Figures are representative of at least three independent experiments.

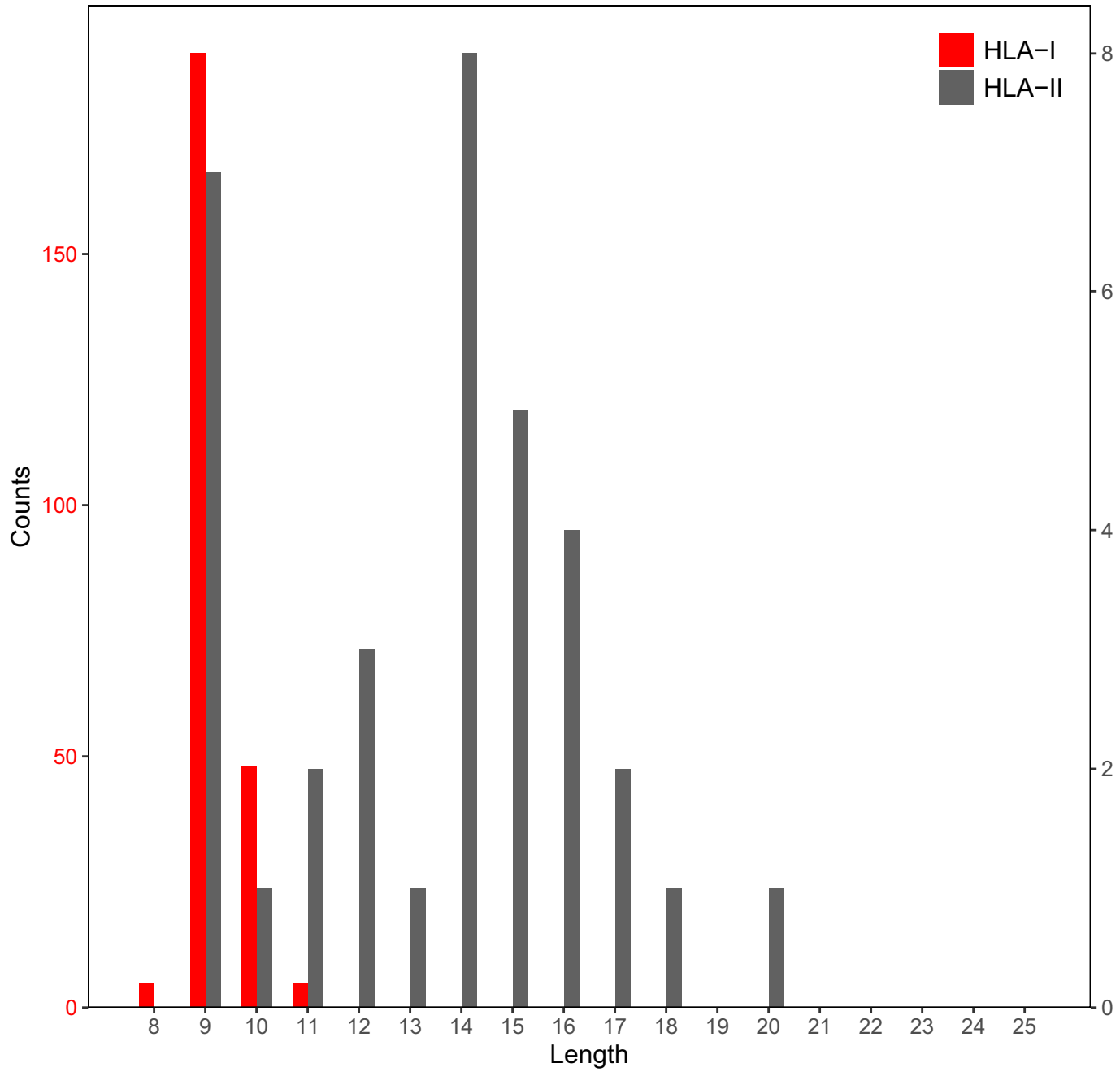


Taxon name	P-value
Acinetobacter	0.026
Actinomyces	0.00071*
Comamonas	0.028
Corynebacterium	0.00029*
Enterobacter	0.0037*
Roseomonas	0.027
Streptococcus	0.031

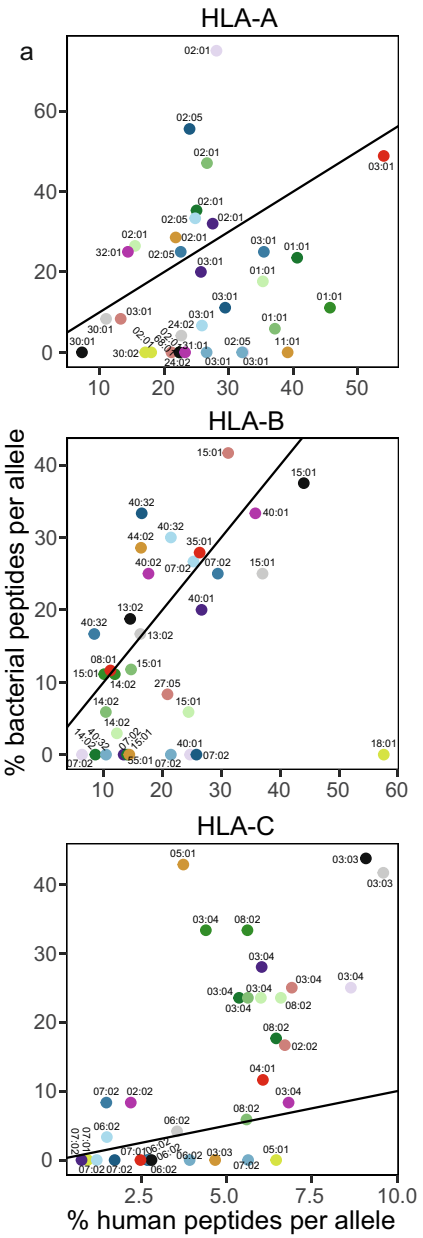
Extended Data Fig. 3 | Taxonomic analysis of 108 whole-genome-sequenced melanoma samples identify a bacterial composition similar to that found in our tumour cohort. a, Alpha diversity, measured as the number of observed species in a tumour (green) or blood (purple) sample. *P* values from paired two-tailed Wilcoxon test between tumour and blood taxonomic diversity. **b**, Microbiome similarity within and between groups. Bray–Curtis dissimilarity measured between each pair of samples, then stratified into four groups. *P* values from two-tailed Wilcoxon test. **c**, Comparison of the relative abundance between tumour samples and associated blood samples. *P* values from paired two-tailed Wilcoxon test between tumour and blood taxonomic abundance. ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$. **d**, List of groups of bacteria that are more abundant in the tumour samples plotted in **c**. *P* values from paired two-tailed Wilcoxon test between tumour and blood taxonomic abundance. *P* values with asterisks survived multiple hypothesis correction (false-discovery rate of 5%). In the box plots, the centre lines represent the medians, the boxes represent the range between the 25th and 75th percentile, and the whiskers represent the range between the smallest and largest data point.

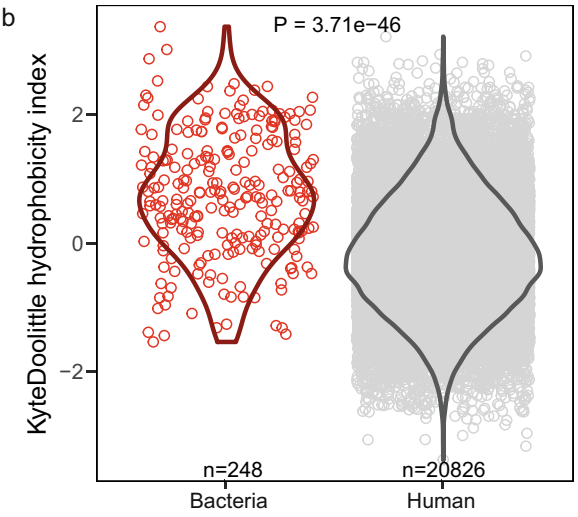


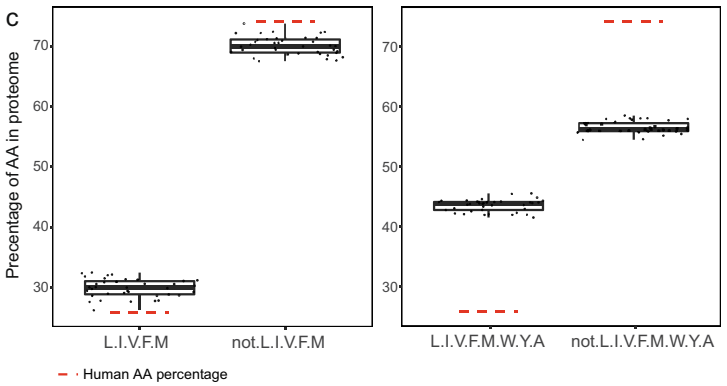
Extended Data Fig. 4 | Pipeline for the identification of bacteria-derived HLA peptides. Tumour samples were 16S-sequenced to identify their bacterial composition and analysed using HLA peptidomics. Searching the data according to the bacteria resulted in the identification of bacteria-derived peptides. Peptides were filtered according to their identification quality and ability to bind to the HLA alleles of the patient. Selected bacterial peptides were then tested for reactivity. Identification of bacterial peptide presentation was validated by a HLA peptidomics analysis of cell lines cocultured with bacteria.



Extended Data Fig. 5 | Length distribution of bacteria-derived peptides. The length distribution of bacteria-derived peptides is similar to the expected length of HLA-I and HLA-II peptides.







d

aa	Human proteome	Bacterial proteome (mean)	p-value	FDR
L,I,V,F,M	25.86	29.91	1.39E-18	2.78E-18
not L,I,V,F,M	74.14	70.09	1.39E-18	2.78E-18
L,I,V,F,M,W,Y,A	36.63	43.53	5.07E-33	2.03E-32
not L,I,V,F,M,W,Y,A	63.37	56.47	5.07E-33	2.03E-32

Patients and metastases

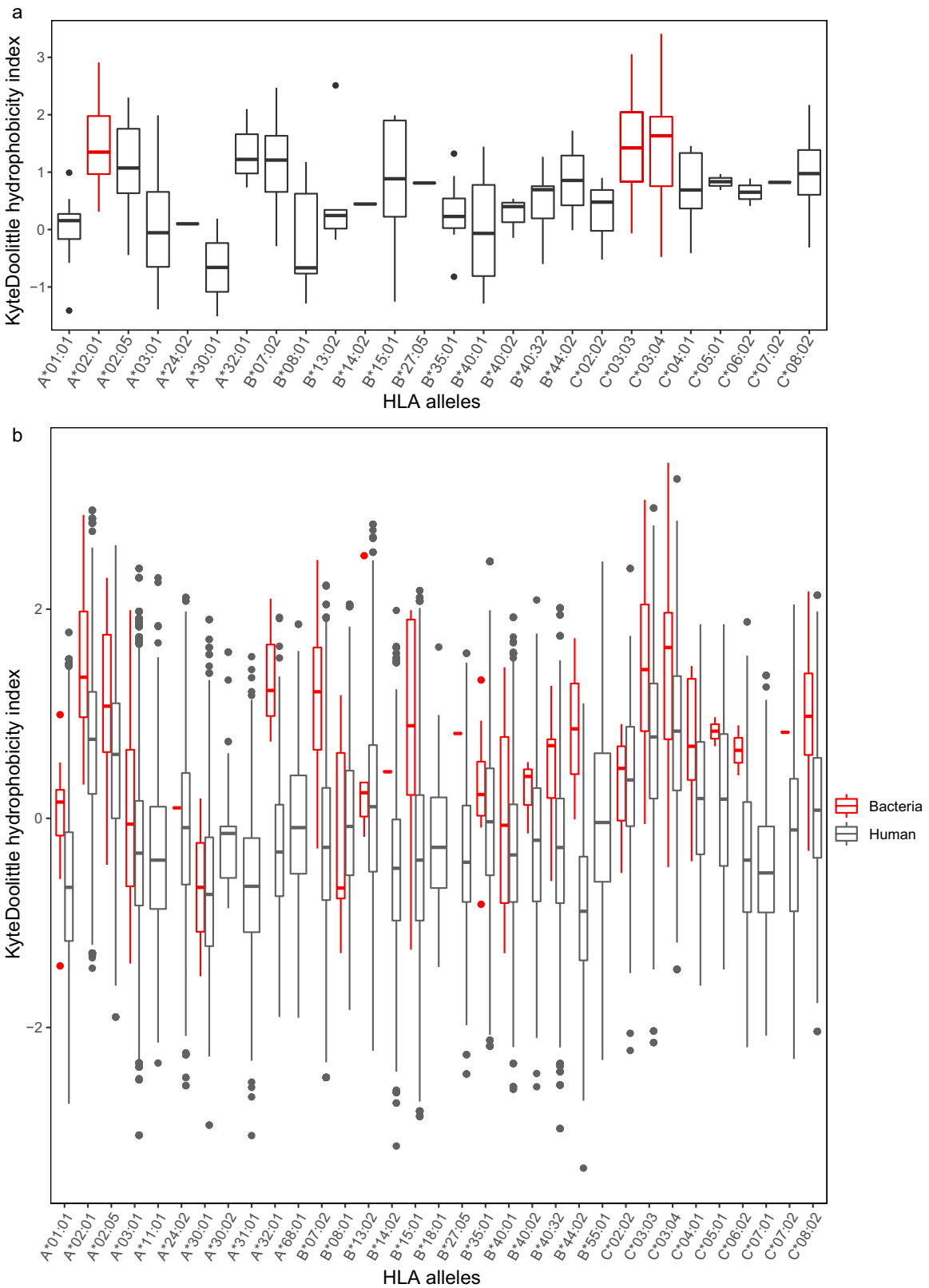
19	42	51
19T	42RF 42RS	27 51AL 51BR-1 51BR-2

55	70	86
55A3 55A7-1 55A7-2 55B3	70.1	86B 86B2

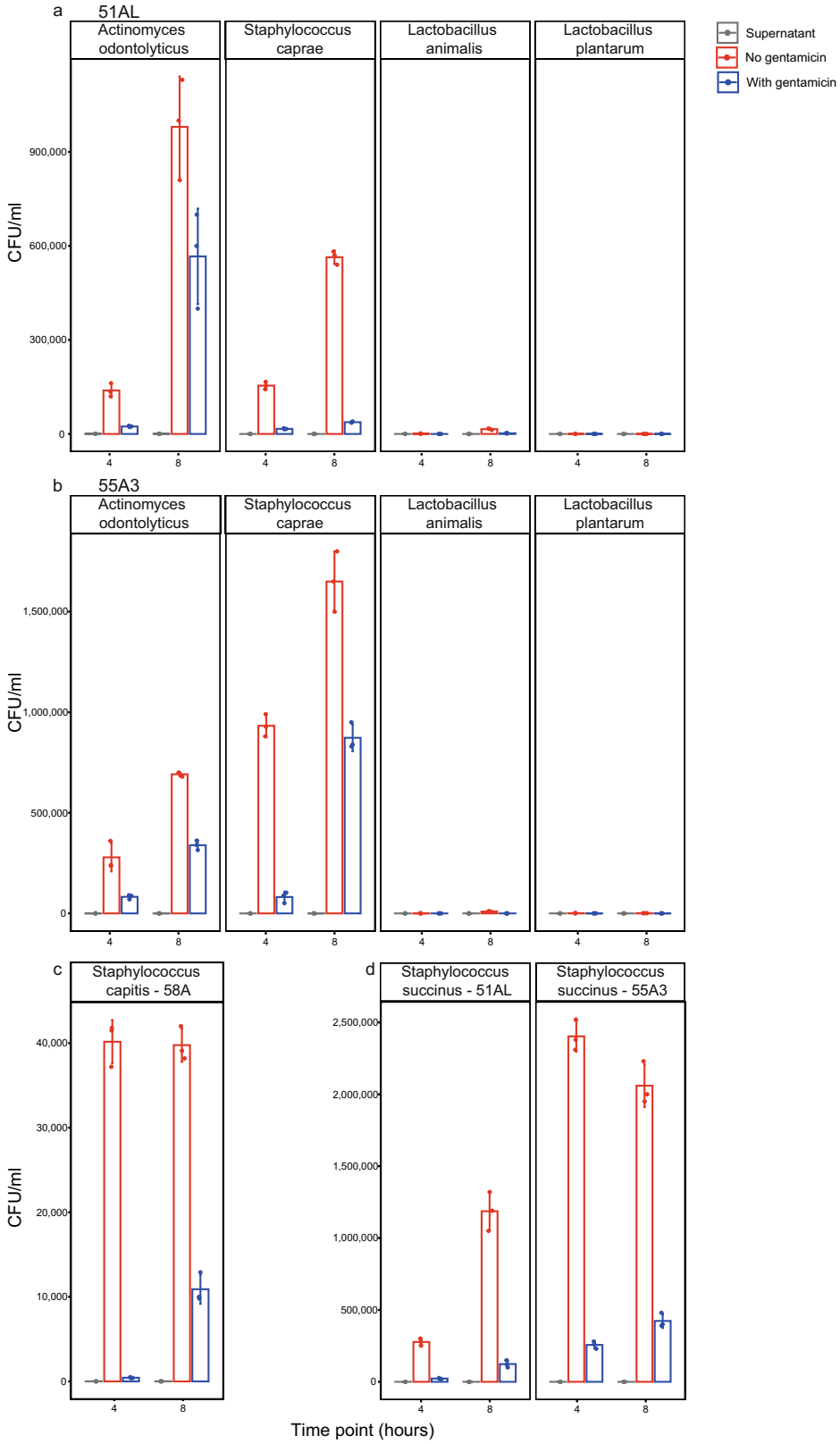
92	112	152
92B3	112	152A2

Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Bacterial proteomes contain a higher amount of hydrophobic amino acids compared to the human proteome. a, For each metastasis, the percentages of bacterial and human peptides that match each *HLA-A*, *HLA-B* and *HLA-C* allele of the patient is indicated. The allele with the best per cent rank binding prediction by NetMHCpan was assigned to each peptide. **b**, Kyte–Doolittle hydrophobicity index was calculated for bacterial and human peptides. The hydrophobicity of HLA-I bacterial peptides is higher than that of human-derived peptides (indicated *P* value is from an unpaired two-sample Wilcoxon test). **c**, The percentage of hydrophobic and nonhydrophobic amino acids was calculated for bacterial proteomes and the human proteome. Two groupings were used for selecting hydrophobic amino acids: L, I, V, F and M, or L, I, V, F, M, W, Y and A. The percentage of hydrophobic and nonhydrophobic amino acids from bacterial proteomes is plotted in the box plot. In the box plots, the centre lines represent the medians, the boxes represent the range between the 25th and 75th percentile, and the whiskers represent the range between the smallest and largest data point. The percentages representing the human proteome are marked by a red dashed line. **d**, Two-sided Student's *t*-test comparing the percentage of hydrophobic and non-hydrophobic amino acids between bacterial proteomes and the human proteome. The *P* values and false-discovery rates are indicated in the table.

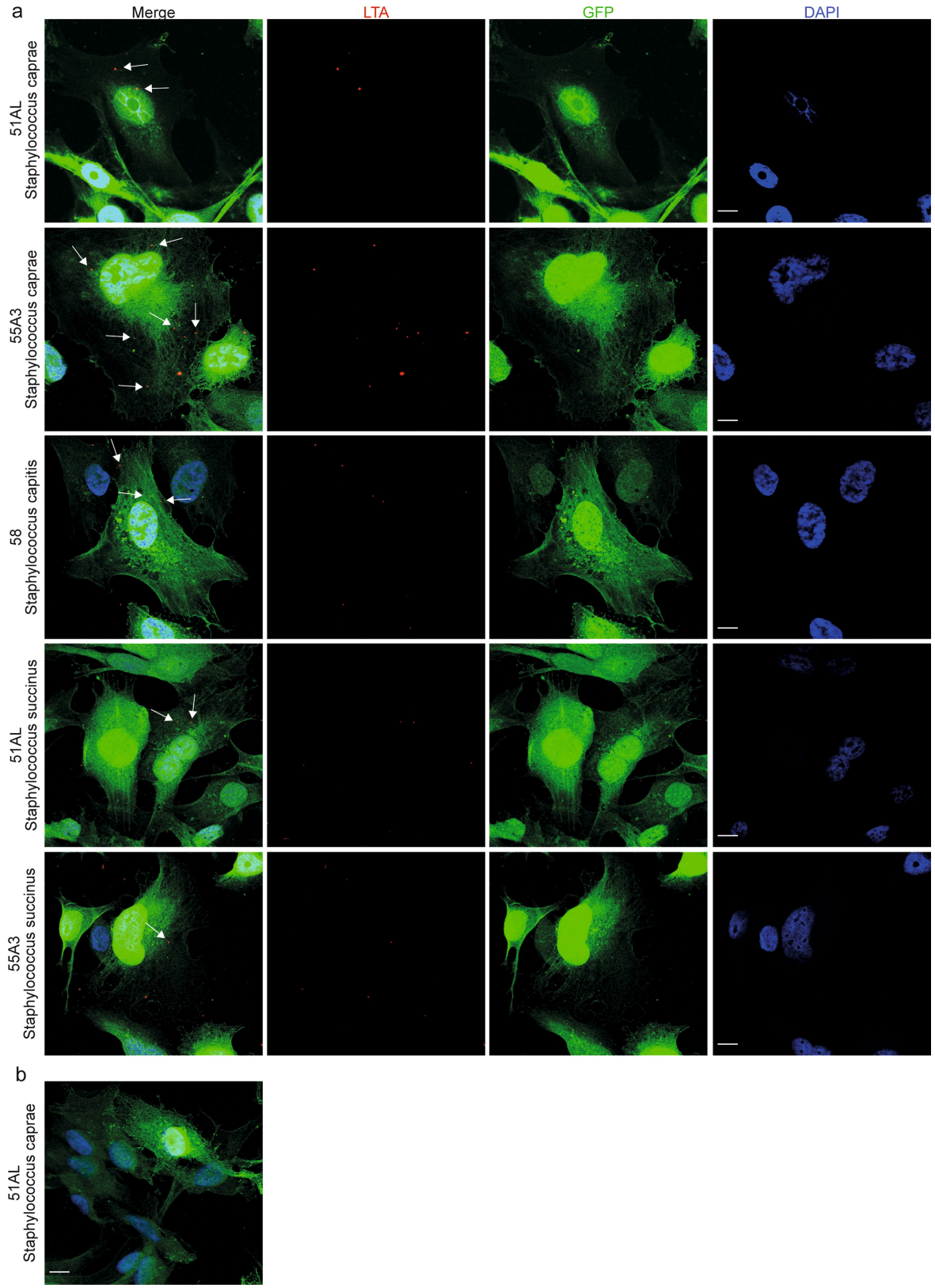


Extended Data Fig. 7 | Hydrophobicity of bacterial and human peptides per allele. Kyte–Doolittle hydrophobicity index was calculated for bacterial and human peptides and plotted in a box plot for each HLA allele. In the box plots, the centre lines represent the medians, the boxes represent the range between the 25th and 75th percentile, and the whiskers represent the range between the smallest and largest data point. **a**, The hydrophobicity of the bacterial peptides that bind to the HLA-C*03:04, HLA-C*03:03 and—to a lesser extent—HLA-A*02:01 was higher compared to the hydrophobicity of other alleles (marked in red). Additional alleles also show this trend, but they were derived from a lower number of tumour samples and therefore are not indicated. **b**, Bacterial peptides are marked in red and human peptides are marked in grey.

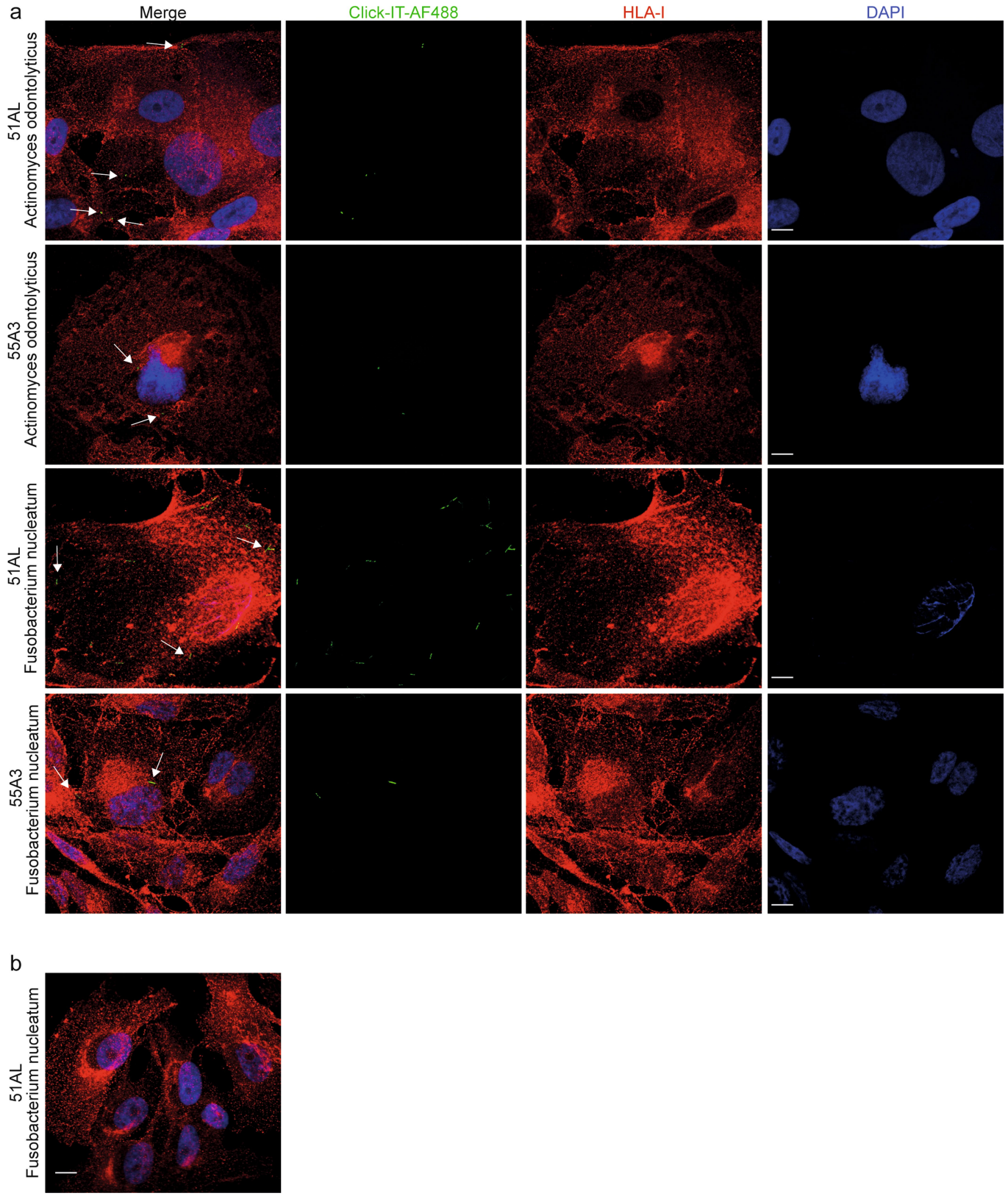


Extended Data Fig. 8 | See next page for caption.

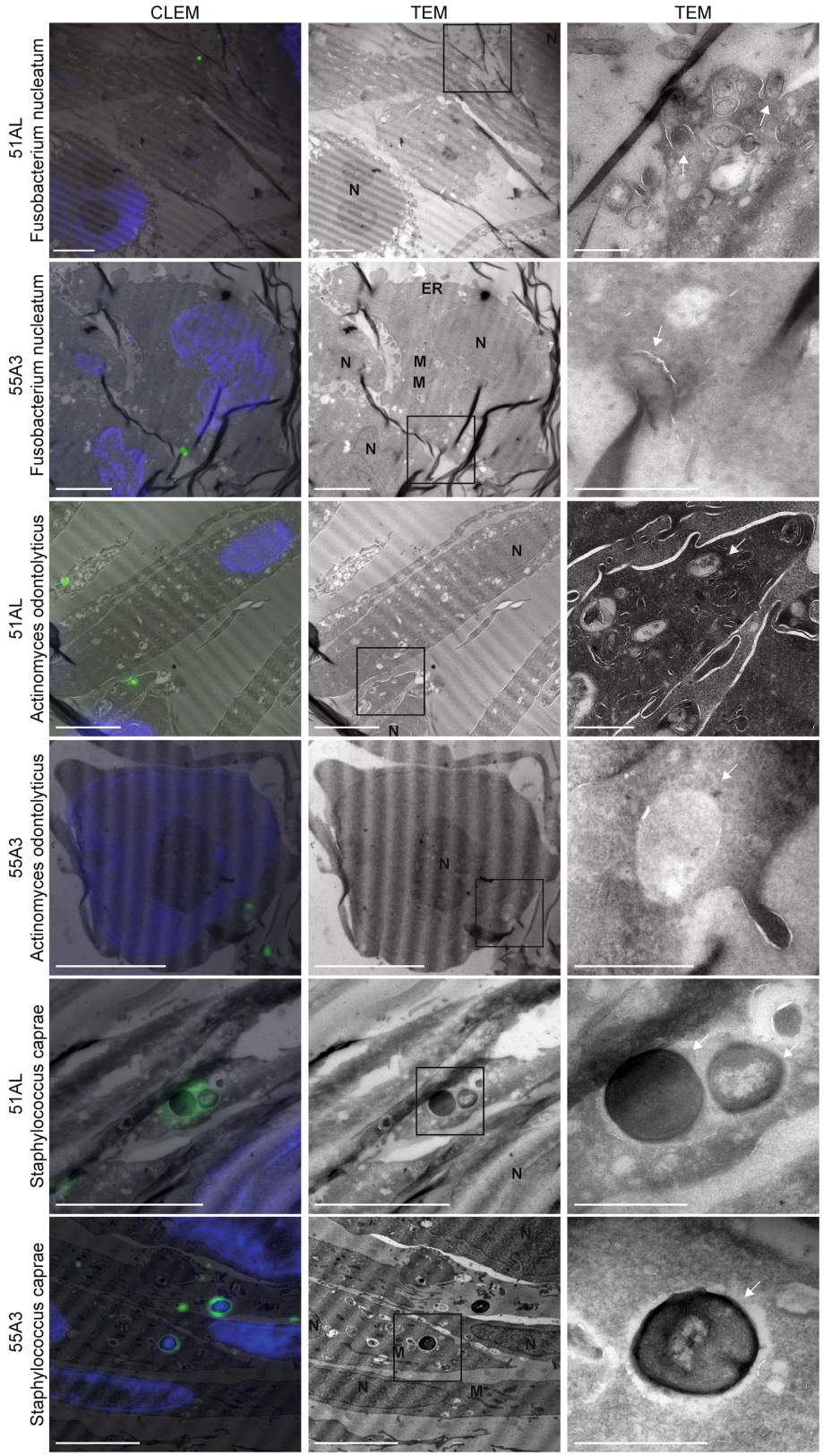
Extended Data Fig. 8 | Gentamicin assay demonstrating the entry of bacteria into melanoma cells. a, **b**, Colony-forming units (CFU) of *A. odontolyticus* and *S. caprae* after coculture with 51AL and 55A3 melanoma cells. Less-invasive bacteria (*L. animalis* and *L. plantarum*) were used as a control, and show lower CFU. **c**, **d**, CFU of *S. capitis* (**c**) (isolated from tumour 58) after coculture with 58A melanoma cells, or *S. succinus* (**d**) (isolated from tumour Mel261) after coculture with 51AL or 55A3 cells. Cells were cultured with the indicated bacteria for 4 and 8 h. 'Supernatant' refers to the CFU of medium taken from samples incubated with gentamicin after coculture with the bacteria (grey). 'No gentamicin' refers to samples not treated with gentamicin after the coculture (red). 'With gentamicin' refers to samples treated with gentamicin for 1 h after the coculture (blue). Bars represent the average of s.e. between biological replicates ($n = 3$). *P* values from Student's *t*-test between the supernatant sample and the without gentamicin or with gentamicin samples; *P* values between *S. caprae* or *A. odontolyticus* with gentamicin to *L. animalis* or *L. plantarum* with gentamicin control samples are from a one-way analysis of variance followed by Tukey's test, and are presented in Supplementary Table 4.



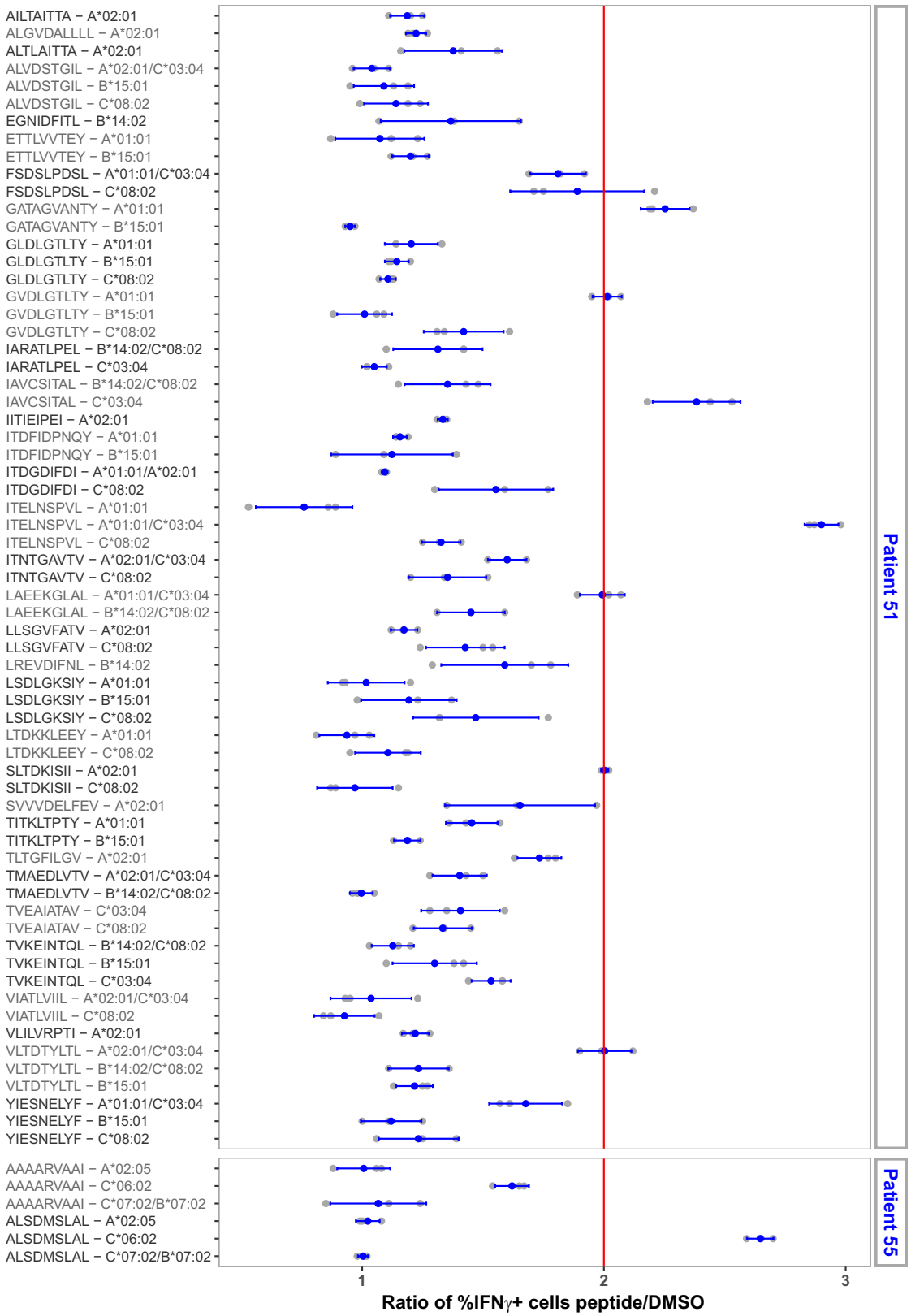
Extended Data Fig. 9 | Immunofluorescence staining of melanoma cells cocultured with aerobically grown bacteria, demonstrating the ability of the bacteria to enter the cells. a, Melanoma cells expressing GFP (green) were cocultured with the aerobically grown bacteria *S. caprae*, *S. capitis* or *S. succinus* stained with antibacterial antibody lipoteichoic acid (LTA) (red); cell nuclei were stained with DAPI (blue). White arrows indicate the location of bacteria that entered the melanoma cells. **b**, A representative image of 51AL cells expressing GFP (green) cocultured with *S. caprae* and stained without a primary LTA antibody (red), to exclude nonspecific staining. Images are presented at 63× magnification. Scale bars, 10 µm. Figures are representative of at least three independent experiments.



Extended Data Fig. 10 | Immunofluorescence staining of melanoma cells cocultured with anaerobically grown bacteria, demonstrating the ability of the bacteria to enter the cells. a, Melanoma cells stained an anti-HLA antibody (red) were cocultured with anaerobically grown bacteria *F. nucleatum* or *A. odontolyticus*. These bacteria were labelled with click chemistry (green). Cell nuclei were stained with DAPI (blue). White arrows indicate the location of bacteria that entered the melanoma cells. **b**, A representative image of 51AL cells stained with the anti-HLA antibody (red) cocultured with *F. nucleatum* that were not grown with D-GalNAz and labelled with Alexa Fluor F488 (green), to exclude nonspecific staining. Images are presented at 63× magnification. Scale bars, 10 µm. Figures are representative of at least three independent experiments.



Extended Data Fig. 11 | CLEM images showing entry of bacteria into melanoma cells. *Fusobacterium nucleatum* was grown with D-GalNAz and then labelled with DIBO–Alexa Fluor 488. *Actinomyces odontolyticus* and *S. caprae* were incubated with an anti-LTA antibody, and then with an anti-mouse secondary antibody labelled with Alexa Fluor 488. The 51AL and 55A3 cell lines were coincubated with the bacteria for 8 h. Ultra-thin sections were analysed by fluorescence microscopy to identify the bacteria (green and blue labelling are for bacteria and nucleus, respectively), followed by transmission electron microscopy (TEM) of the same cells for high-resolution morphology. Left and middle panels show CLEM and TEM images, respectively. Scale bars, 5 µm. The right panel shows high-magnification TEM image of the area in the black box in the corresponding middle panel. Scale bars, 1 µm. Bacteria that entered the melanoma cell are indicated with a white arrow. N, nucleus; M, mitochondrion; ER, endoplasmic reticulum. Figures are representative of at least three independent experiments.



Extended Data Fig. 12 | TIL reactivity towards bacteria-derived antigens. Flow cytometry analysis of IFNγ-secreting TILs after coculture of 51AL and 55A3 TILs with B cells loaded with the indicated bacterial peptide. Each peptide was loaded on a different B cell that exhibited the HLA alleles to which the peptide was predicted to bind. TILs were stained with anti-IFNγ and anti-CD45 bifunctional antibody, which binds secreted IFNγ. The value indicates the ratio of IFNγ-secreting cells with the peptides to those with the DMSO control. Grey dots indicate the results of $n = 3$ biological replicates, and blue dots represent the average of replicates. Bars represent s.e. between replicates.

nature research

Corresponding author(s): Prof. Yardena Samuels

Last updated by author(s): February 4th 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	SMURF package in R, V1 ggplot2 package in R, v3.3.0 ggtree package in R, v2.2.4 BWA-MEM software Kraken v.1.0 phyloseq package in R, v1.3.2.0 vegan version 2.5-4 package in R CD-HIT software, v4.6.6 MaxQuant version 1.5.0.25 NetMHCpan version 4.0 NetMHCIIpan version 3.2 GibbsCluster 2.0 server Seq2Logo 2.0 Peptides package in R, v2.4.1 R package OrgMassSpecR, v0.5.3 Trimmomatic v0.38-WGS data for bacteria (Bolger, A. M. et al., 2014) , SPAdes v3.14 -Quality filter of WGS paired-end de novo assembled (Bankevich A. et al., 2012), QUAST v5.0 -Genome assembly quality(Gurevich, A. et al., 2013), CheckM v1.0.18 -Genome assembly quality(Parks, D. H. et al., 2015), BWA-MEM software v0.7(Li, H.et al,2013), metabat2 v2.14 -Contig grouping (Kang, D. et al, 2015), GTDB release 04-RS89-Taxonomic classifications (Parks, D. H. et al., 2020), GTDB-tk v0.3.2- Taxonomic classification(Chaumeil, P. et al, 2019), fastANI v1.1-Taxonomic classification(Jain, C. et al,2018), dRep- clustering (Olm, M. et al,2017), Prokka v1.140- protein sequence annotations(Seemann, T. et al, 2014). FlowJo software v7

ImageJ 1.53C
Imaris 9.1 (Bitplane) software

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw MS files as well as human and bacteria proteomes and the MaxQuant version used for analyzing the data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD022150. Raw sequence data of isolated bacteria WGS were deposited in the NCBI Sequence Read Archive (SRA), under BioProject accession number PRJNA669827. Source Data 1 is related to Figure 2a, and includes all raw files of identified peptides for HLA peptidomics experiments. Source Data 2 is related to Figure 3e, and includes all the data recived from bacteria WGS (16S rRNA sequences, genome assembly and Protein sequences).
Other publicly available detests used:
The Greengenes database (May 2013 version) for 16S sequencing reference
High-throughput DNA sequencing of 108 pairs of tumor and blood samples (Hayward, N. K. et al., 2017)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determine according to biological material availability
Data exclusions	Data was not excluded
Replication	All experiments required were done in triplicates and repreduced
Randomization	For all high throughput data analysis-randomization for samples into experiment groups is not relevant, as the design of the study requires explorative analysis for identifying discriminative features between already established experimental groups.
Blinding	For all high throughput data analysis-blinding to investigators is not relevant to this study, as the design of the study requires explorative analysis for identifying discriminative features.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-Lipoteichoic acid antibody, Cat #HM5018, Hycult Biotech, Clone 55 , LOT 25823M0119-A anti-LPS, Cat# HM6011, Clone WN1 222-5

anti-mouse secondary antibody labeled with Alexa flour 488, Cat #115-545-146, Jackson Immunoresearch, Clone , LOT 117790
anti-mouse secondary antibody labeled with Alexa flour 546, Cat #A-11030, Invitrogen, AB_141370
AF647-conjugated CD45, Cat# 7368537, lot B239311s, clone 2D1,BioLegend
APC/Cy7-conjugated CD19 , Cat# 302218, lot B279663, clone HIB19,BioLegend
APC-conjugated CD19, Cat# 302212, lot B182157, clone HIB19,BioLegend
APC-conjugated HLA-A,B,C ,Cat# 311409, lot B213470, clone W6/32,BioLegend
APC-conjugated HLA-DR,DP,DQ ,Cat# 361714, lot B289409, clone Tu39,BioLegend
BV421-conjugated CD69,Cat# 310930, lot B284312, clone FN50,BioLegend
anti-GAPDH, Cat# MAB374, lot 3043558, clone 6C5,Millipore
anti-CD45, Cat# 13917, lot 6, clone D9M8I,Cell Signaling
anti-HLA-DP1, Cat# ab157210, lot GR3173627-19, clone EPR11226,Abcam
APC/Cy7-conjugated CD14 Cat# 301820, lot B280758, clone M5F2,BioLegend
Anti-pan-HLA–I antibody, Cat #H1650 ,Sigma, Clone W6/32, LOT 22130603
IFNg capture antibody, as part of kit, Cat #130-054-202, Miltenyi Biotec, LOT 5190207203

Validation

All antibodies used in the study were validated by the manufacturer, data is available at the manufacturer's website, as indicated bellow:
anti-Lipoteichoic acid antibody, Cat #HM5018, Hycult Biotech, Clone 55 , LOT 25823M0119-A: https://www.hycultbiotech.com/hm5018-1
anti-LPS, Cat# HM6011, Clone WN1 222-5, https://www.hycultbiotech.com/hm6011
anti-mouse secondary antibody labeled with Alexa flour 488, Cat #115-545-146, Jackson Immunoresearch, Clone , LOT 117790, https://www.jacksonimmuno.com/catalog/products/115-545-146
anti-mouse secondary antibody labeled with Alexa flour 546, Cat #A-11030, Invitrogen, AB_141370, https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11030
AF647-conjugated CD45, Cat# 7368537, lot B239311s, clone 2D1,BioLegend, https://www.biolegend.com/en-us/products/apc-anti-human-cd45-antibody-12397
APC/Cy7-conjugated CD19 , Cat# 302218, lot B279663, clone HIB19,BioLegend, https://www.biolegend.com/en-us/products/apc-cyanine7-anti-human-cd19-antibody-1910
APC-conjugated CD19, Cat# 302212, lot B182157, clone HIB19,BioLegend, https://www.biolegend.com/en-us/products/apc-anti-human-cd19-antibody-715
APC-conjugated HLA-A,B,C ,Cat# 311409, lot B213470, clone W6/32,BioLegend, https://www.biolegend.com/en-us/products/apc-anti-human-hla-a-b-c-antibody-1870
APC-conjugated HLA-DR,DP,DQ ,Cat# 361714, lot B289409, clone Tu39,BioLegend, https://www.biolegend.com/en-us/products/apc-anti-human-hla-dr-dp-dq-antibody16319
BV421-conjugated CD69,Cat# 310930, lot B284312, clone FN50,BioLegend, https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-human-cd69-antibody-7141
anti-GAPDH, Cat# MAB374, lot 3043558, clone 6C5,Millipore, https://www.merckmillipore.com/INTL/en/product/Anti-Glyceraldehyde-3-Phosphate-Dehydrogenase-Antibody-clone-6C5,MM_NF-MAB374
anti-CD45, Cat# 13917, lot 6, clone D9M8I,Cell Signaling, https://www.cellsignal.com/products/primary-antibodies/cd45-intracellular-domain-d9m8i-xp-rabbit-mab/13917
anti-HLA-DP1, Cat# ab157210, lot GR3173627-19, clone EPR11226,Abcam, https://www.abcam.com/hla-class-i-antibody-w632-low-endotoxin-azide-free-ab23755.html

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

51AL,55A3,58A were produced by MD Anderson
EBV-transformed B-cells were purchased from the IHWG Cell and DNA Bank
Hybridoma cells HB95 and HB145 were purchased from the ATCC
LCL 721.221 cell line was purchased from the ATCC
THP-1 cells were a kind gift from Prof. Steffen Jung (originally from ATCC)

Authentication

Commercial cells were authenticate by the supplier
51AL, 55A3, 58A- The Cell line authentication test was performed at the Genomics Center of Biomedical Core Facility, Technion. The test was performed using the Promega GenePrint 24 System in order to determine short tandem repeat (STR) profile of 23 loci plus Amelogenin for gender determination (X or XY). In addition, the male-specific DYS391 locus is included to identify null Y allele results for Amelogenin. DNA sample from the kit (2800M Control DNA) was included in the analysis and served as positive control for the PCR step. No DNA template was also included as negative control. The results were analyzed using the 3500xl Genetic Analyzer (Life Technologies) and GeneMapper IDX software. Allelic ladder was included in the run.

Mycoplasma contamination

All cell lines were tested regularly and were found negative for mycoplasma contamination (EZ-PCR Mycoplasma Kit, Biological Industries).

Commonly misidentified lines
(See ICLAC register)

no commonly misidentified cell lines were used in the study

Human research participants

Policy information about studies involving human research participants

Population characteristics	All patients included in the analysis were diagnosed with metastatic melanoma. Tumor samples from patients 19, 42, 51, 55, 58, 70, 86, 92, 112, 152 and 422 were received from eleven different patients treated at the University of Texas MD Anderson Cancer Center, all of who had signed an informed consent for the collection and analysis of their tumor samples. The protocol for tumor samples was approved by the MD Anderson Institutional Review Board (protocol numbers 2012-0846, LAB00-063 and 2004-0069, NCT00338377). Synchronous metastatic tumors were resected via surgery at the same time. Sample and patient information are provided in Extended Data Table 1. From each patient, we collected between 1-3 different metastases. Tumor sample HG38 was received from patient biopsies removed at the University of Zurich Hospital. The patient had signed a patient release form, which has been approved by an ethics committee and assigned the numbers EK647 and EK800. Tumor 261MS was received from patient biopsies removed at the Ella Lemelbaum Institute for Immuno-Oncology (Sheba medical center, Israel), under Institutional Review Board number 6786-20-SMC.
Recruitment	Patient were recruited by University of Texas MD Anderson Cancer Center, all of who had signed an informed consent for the collection and analysis of their tumor samples. Tumor sample HG38 was received from patient biopsies removed at the University of Zurich Hospital. The patient had signed a patient release form, which has been approved by an ethics committee and assigned the numbers EK647 and EK800. Tumor 261MS was received from patient biopsies removed at the Ella Lemelbaum Institute for Immuno-Oncology (Sheba medical center, Israel), under Institutional Review Board number 6786-20-SMC.
Ethics oversight	The protocol for tumor samples was approved by the MD Anderson Institutional Review Board (protocol numbers 2012-0846, LAB00-063 and 2004-0069, NCT00338377). Tumor sample HG38 was received from University of Zurich Hospital under an ethics committee and assigned the numbers EK647 and EK800. Tumor 261MS was received from the Ella Lemelbaum Institute for Immuno-Oncology (Sheba medical center, Israel), under Institutional Review Board number 6786-20-SMC.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	TILs after co-culture with EBV transformed B-cells loaded with peptides, were labeled with an IFNg capture antibody (130-054-202, Miltenyi Biotec) and then incubated for 45 minutes at 37°C. Then cells were stained with the detection antibody. For HLA-I and HLA-II staning all cells were washed X2 and staind for 30 minutes on ice with the antibodies.
Instrument	BD LSR II flow cytometer (BD Biosciences)
Software	FlowJo software
Cell population abundance	Known cell amounts, of the different cell types, in known ratio and was constant during experiments
Gating strategy	In all assays, cells were first gated for live, single cells, and then gated according to the requested staining.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.