

# Single cell analysis of the CD8+ T cell response to therapeutic Melan-A peptide vaccination over time

Master's thesis in Medicine

Alexandre Huber

Under the supervision of

Nathalie Rufer, PhD, MD & Philippe Gannon, PhD

## Abstract

### Background

Therapeutic vaccines are currently being optimized to enhance melanoma patients' own immune defenses against cancer cells. The Melan-A<sup>MART-1</sup> tumor antigen is frequently expressed in melanoma and has been widely used in clinical immunotherapy studies. We previously reported that vaccination with the native unmodified Melan-A (defined as EAA hereafter), but not the analog peptide (defined as ELA), despite its stronger binding to HLA-A2, elicited a polyfunctional phenotype (i.e. co-expressing genes associated with memory and effector attributes) by tumor-specific T cells, together with increased tumor cell killing capacities by these cells.

### Objective

The present study aims at characterizing the timing of acquisition of memory/effector properties among individual tumor-reactive CD8 T cell clones in melanoma patients following vaccination with the native/EAA peptide, compared to responses induced with the analog/ELA peptide. Moreover, we assessed the evolution of the tumor-specific T cell clonotype repertoire across time following vaccination in both cohorts of patients and its specific selection at early time points as compared to the specific responses pre-existing before the start of immunotherapy.

### Methods

We generated a large library of cDNAs (n = 1500) isolated from single Melan-A-specific effector-memory (EM) CD28+ (memory-like; EM28+) and CD28- (effector-like; EM28-) T cells at different time points before and after peptide vaccination. Our highly sensitive single cell gene expression approach allowed the direct *ex vivo* characterization of individual tumor-specific T cells in melanoma patients who received either the native/EAA (n = 4) or the analog/ELA (n = 3) peptide vaccine. PCRs specific for the TCR $\beta$ -chain variable region were further performed to determine the TRBV/CDR3-based clonotype of each individual tumor-specific T cells across time after vaccination.

### Results

When compared to pre-vaccine T cell phenotype, tumor-specific T cells underwent a drastic differentiation change into effector-memory following CpG vaccination, regardless of the type of peptide vaccine used (native/EAA or analog/ELA). Acquisition of the EM28+ phenotype shortly after vaccination was specifically associated with the significant increase in effector gene expression and concomitant reduction of a memory gene expression pattern (IL-7R $\alpha$ + CXCR3- CCR5-). Dominant tumor-specific TCR clonotypes were expanded during this process and many of them were maintained stable over time, as they also were found present at late time points after immunotherapy. Using the advantage of the single cell resolution, a difference in CD27 gene expression could be identified between EM T cells with dominant or non-dominant TCR clonotypes.

### Conclusions

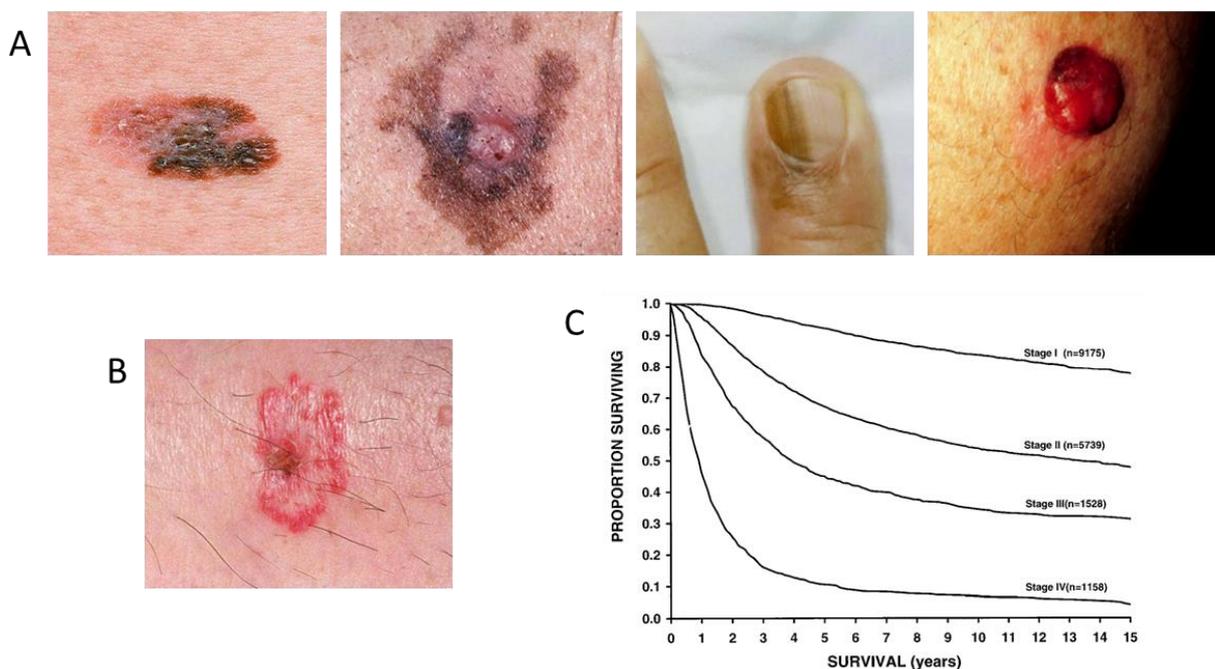
Melan-A peptide/CpG vaccination induced the strong T cell differentiation of tumor-reactive T cells, which was accompanied by increased effector gene expression and expansion of dominant clonotypes, in both native/EAA and analog/ELA based vaccine. Moreover, our data further suggest that tumor-specific T cells may undergo a maturation process during the course of vaccination, which may account for the differences observed in the co-expression of memory and effector genes within native/EAA EM28+ T cells at later time points. Finally, once established, the clonal composition of tumor-specific T cell responses was kept stable along immunotherapy. Collectively, such analyses provide important insights on the *in vivo* impact of natural over analog peptide vaccination on T cell polyfunctionality and clonotype selection induced by each type of vaccination.

## Introduction

### Melanoma

Melanoma is the deadliest type of skin cancer surpassing squamous and basal cell carcinoma despite their higher incidence. It is diagnosed in about 2'000 patients and accounts for nearly 500 deaths each year in Switzerland (1). The main risk factor is the exposure to UV irradiation in predisposed individuals, which causes DNA damage and is reflected by the higher number of mutations found in melanomas relative to other cancer types (2).

Melanoma derives from melanocytes, which contain melanin and gives the skin its pigmentation. Melanomas most frequently present as pigmented lesions of the skin or mucosae, but non-pigmented lesions may also be diagnosed (Fig. 1A-B). These amelanotic melanomas are a diagnostic challenge and tend to be diagnosed at more advanced stages (3). Importantly, melanomas may develop either from previous pigmented lesions such as naevi or normal skin (4).



**Figure 1.** A. Melanoma subtypes. From left to right: superficial spreading melanoma, lentigo maligna melanoma, acral lentiginous melanoma and nodular melanoma (4) B. Amelanotic melanoma lesion (4) C. 15-year melanoma-specific survival curves according to stage of the disease at time of diagnosis (5).

### Evolution and prognosis

Melanoma usually follows a progression pattern from the primary skin lesion to the lymphatic system and to distant organs. The prognosis of the disease is best apprehended using the Tumor-Node-Metastasis (TNM) staging system (3). TNM stages with similar prognoses are usually pooled in four group stages (I-IV). Melanoma limited to the skin (stages I-II) may be classified into various subtypes with different local invasion patterns, either more parallel to the skin plane or with early vertical invasion (4). For all types, the most important prognostic feature is the thickness of the lesion, or Breslow index, which provides an indirect measurement of lymph node metastasis probability (6). The mitotic rate is the second most important independent prognostic factor in early lesions. Other parameters also carry prognostic value, such as ulceration and localization.

Lymph nodes metastases (stage III) are a major negative prognostic factor as reflected by survival curves (Fig. 1C) and may often progress to stage IV with distant hematogenous metastases to the lungs, the bones, the brain and the liver (5,6).

## Treatment

Treatment options for melanoma patients include surgery, standard chemotherapy, targeted therapies and immunotherapy, while radiotherapy plays a more palliative role (7). Targeted therapies are recent developments and include mutated BRAF (vemurafenib, dabrafenib) and MEK inhibitors (cobimetinib, trametinib). They have shown improved progression-free and overall survival over standard chemotherapy, and may be combined for increased efficiency (8–11). However, long term progression-free survival could not be demonstrated with these targeted therapies.

Melanoma immunotherapy is a rapidly evolving field. While it has long been limited to IL-2 and IFN $\alpha$  therapies, which were associated with major adverse effects (12), novel treatments like immune checkpoint inhibitors (CTLA4 and PD-1 inhibitors) and adoptive cell therapies are now available in the clinic (13). These treatments function through the stimulation of the anti-tumoral adaptive cellular immune response. Other immunotherapeutic modalities, such as peptide-based vaccines, the main subject of this report, are currently in development and aim at enhancing the same immune responses in a more targeted manner.

## Cytotoxic T cell immune response

Cytotoxic CD8<sup>+</sup> T cells are the primary effector of a potent anti-tumor immune response. The overarching goals of any immune-based therapy are to favor the activation of CTLs and promote their migration into the tumor bed. Although other cell types are directly implicated in the development of the immune response (CD4<sup>+</sup> T cells, DC...), a large set of pre-clinical and clinical models have shown that cytotoxic CD8 T cells play a primary role in controlling and eliminating tumor cells (14,15).

## Functions

Cytotoxic T lymphocytes (CTLs) are important components of the human adaptive immune system whose primary function is to destroy abnormal cells displaying specific antigens, such as viral epitopes or aberrantly expressed or mutated epitopes in the context of cancer. The target recognition relies on surface antigen presentation by the infected or malignant cells under the form of short peptides (9-10 amino-acids) loaded onto the major histocompatibility complex I (MHC-I). CTL specificity for a given antigen is ensured by direct interaction of their T cell receptor (TCR) and CD8 coreceptor with the peptide·MHC-I complex. Upon antigen recognition, activated CTLs have the capacity to degranulate and to release cytotoxic proteins such as perforins, which form pores in the target cell membrane, and granzymes, which cleave and activate cytoplasmic caspases to induce the apoptotic cascade. CTLs may also produce chemokines (e.g. IFN $\gamma$ , TNF $\alpha$ ), which serve to activate other immune cells (e.g. macrophages), and promote antiviral pathways of somatic cells, such as MHC-I-dependent antigen presentation. IFN $\gamma$  also enhances MHC-II-dependent antigen presentation by macrophages to further stimulate the adaptive immune response.

## T cell receptor

The TCR is a transmembrane heterodimer, composed of an  $\alpha$  or  $\gamma$  subunit associated with a  $\beta$  or  $\delta$  subunit, respectively. The majority of T cells express an  $\alpha/\beta$  complex rather than a  $\gamma/\delta$  complex. The TCR complex consists of the non-covalent association between the TCR $\alpha\beta$  chains with the invariant CD3 proteins and the binding of the TCR to the pMHC complex triggers the activation of these CD3-associated subunits allowing TCR-mediated signal transduction and T cell activation (16).

In order to recognize a very large panel of possible foreign antigens, the adaptive immune system has the unique ability to generate a tremendous diversity of TCRs, allowing peripheral CTLs to express an extremely wide repertoire of TCRs. This diversity, known as gene-segment recombination, is ensured for each subunit by genomic recombination of one gene segment of the V (variable) group, which forms the 5' of the coding sequence, with one gene segment of the D (diversity) group (in TCR $\beta$  and  $\delta$  only) and one gene segment J (joining) group. These sequences code for the antigen recognition domain and are finally recombined with the constant 5' part of the gene, coding for the rest of the extracellular domain and the membrane-spanning region. Recombination of the segments ensures a

large part of the TCR repertoire diversity in combination with other processes such as random nucleotide addition.

### **T cell development and selection**

T cell precursors are produced by the hematopoietic system in the bone marrow. Their maturation is a multistep process, which mainly occurs in the thymus (17). During this process, thymocytes undergo TCR gene rearrangement as well as positive and negative selection, whereby the TCR affinity for self-antigens is the key element in determining the fate of T cells. Indeed, single positive CD4 or CD8 thymocytes expressing TCRs that efficiently interact with MHC-I or -II within a given affinity window will survive, exit the thymus and form the peripheral pool of mature CD8 and CD4 T lymphocytes. Importantly, the process of negative selection, which removes thymocytes that bind strongly to self-antigen/MHC complexes, is not perfect and allows T cells of relative high avidity for self-peptide/MHC to survive. While this selection step limits T cell autoimmunity and related disorders, it also restricts the anti-tumoral T cell response.

### **CTL activation**

Peripheral mature CTLs that have not yet encountered any cognate antigen are called naïve (often defined as CD45RA+/CCR7+ or CD45RA+/CD62L+). These cells patrol the lymph nodes where they may become activated through antigen presentation, most commonly by dendritic cells, although other cells may also function as antigen-presenting cells (APCs) (e.g. macrophages and B cells). Antigen presentation alone is not sufficient to activate CTLs and will even induce anergy, a non-functional state which prevents further activation. Costimulatory molecules such as CD80/86 must be presented to the CD28 receptor present on naïve CTLs at the same time as the antigen to the TCR for proper activation and expansion (16). A third signal is necessary and is mediated by cytokines, such as IL-2. On the other hand, coinhibitory molecules may interfere with that process to promote immune tolerance and are expressed on either APCs or CTLs. A prototypic example is the CTLA-4 receptor which may compete for CD80/86 binding and negatively regulate activation (16).

### **Dendritic cells**

Dendritic cells are present in virtually all tissues. Their main function is to phagocytose pathogens upon infection, and to migrate to the draining lymph node in order to present potential antigens to both CTLs and helper CD4 T cells through the corresponding MHCs. This complex multistep process is highly regulated to prevent activation in non-pathological situations. Dendritic cells must first receive adequate local signals in order to be properly activated. They may for example sense cytokines produced by tissue and immune cells. They also have the ability of directly recognizing invariant damage- and pathogen-associated molecular patterns (DAMPs and PAMPs) through a set of dedicated receptors. An important example of such receptors is the Toll-like receptor 9 (TLR9), which is sensitive to unmethylated CpG DNA sequences and may get activated upon corresponding microbial infection. Integration of these various types of signals is necessary to stimulate antigen presentation and allow the expression of costimulatory molecules at the surface of dendritic cells. Although this mechanism may be sufficient for CTL activation, an additional step is required to trigger a long-lasting and fully functional CTL response. This regulatory mechanism involves prior MHC-II-dependent antigen presentation to cognate activated CD4 T cells, which in turn license the dendritic cells through their CD40 receptor (18). Licensing is thought to rely on the enhancement of costimulatory molecules and the reduction of inhibitory molecule expression as well as induction of chemokines that may attract naïve CTLs through their CCR5 receptor (19).

### **CTL expansion and differentiation**

Following acute antigen-specific stimulation, CD8+ T cells mainly differentiate into effector cells. These cells have full killing potential and migrate to the sites of infection to help clearing pathogen-infected cells. They are, however, short-lived and will undergo apoptosis once the stimulus is removed. Another population of CTL is generated alongside and is characterized with a prolonged lifespan potential and varying degrees of self-renewal abilities. These cells are called memory cells as they have

the capacity to be reactivated and generate a faster effector response upon further encounter of the same antigenic stimulus, which constitutes the hallmark of an adaptive immune response (20). Expression of many genes have been associated with the memory phenotype and regulate the activation / proliferation potential (CD27, CD28, EOMES), cytokine response (CD127/IL-7R $\alpha$ ) and homing properties (CCR7, CXCR3, CCR5, CD62L) of these cells (21).

Distinct subsets of memory cells have been described with varying degrees of memory and effector traits and may be distinguished by combinations of surface marker expression. Of the three main subsets, memory stem cells have the highest degree of self-renewal potential (22), followed by central memory cells (CD62L<sup>Hi</sup> CD45RA<sup>-</sup> CCR7<sup>+</sup>) and effector memory (CD62L<sup>Lo</sup> CD45RA<sup>-</sup> CCR7<sup>-</sup>) (23). Conversely, effector memory cells express higher levels of effector markers (IFN $\gamma$ , PERP, GZMB, CD94) than central memory cells, which correlate with their higher cytotoxic potential.

Effector memory (EM) cells may be further subdivided in CD28<sup>+</sup> (EM28<sup>+</sup>) and CD28<sup>-</sup> (EM28<sup>-</sup>) cells (21,24). EM28<sup>+</sup> cells express effector proteins at a lower level than EM28<sup>-</sup> cells, but may have increased memory and proliferation potential through the expression of memory genes and the telomerase enzyme, respectively.

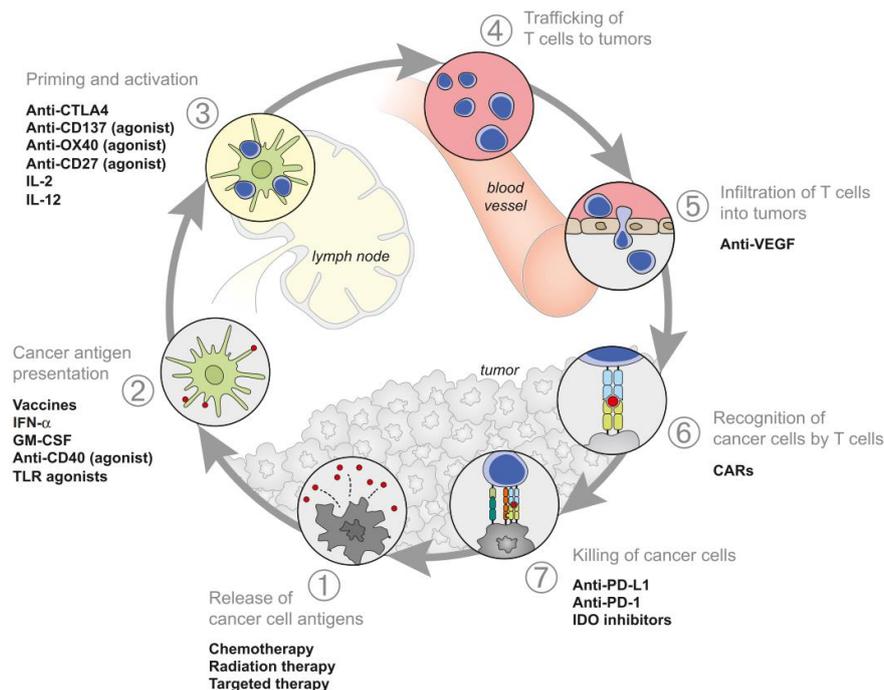
## Cancer immunity

The primary mechanism of cancer is the accumulation and selection of genetic anomalies leading to the loss of growth and proliferation regulation (2). These mutations lead to gene expression dysregulation, downregulating tumor suppressor genes while alleviating the repression of genes that are needed for cell survival and replication (oncogenes). These alterations may be recognized as foreign antigens by the adaptive immune system and serve as targets for immunotherapy. There are at least four classes of cancer antigens. In a decreasing order of specificity, neoantigens resulting from somatic mutations come first, followed by cancer testis antigens (a subset of genes that are normally not expressed in any tissue except male germ cells), differentiation antigens (genes expressed in the tissue from which the cancer arises, e.g. PSA in prostate cancer) and overexpression antigens (e.g. HER2 in breast cancer) (25). It is still not clear which specific antigens are or should be targeted in immunotherapy to achieve an efficient antitumoral response. Some hypotheses have been suggested, for example targeting neoantigens resulting from driver mutations may be more advantageous than targeting non-mutated self-antigens, since the T cell repertoire available for these antigens is not affected by central T cell tolerance (26). However, neoantigens also have the disadvantage of only being shared by a minority of patients with a given cancer type, while differentiation antigens are expected to be much more commonly expressed and shared among patients and tumor types.

## Cancer-immunity cycle

An ideal anti-cancer CTL immune response based on antigens released by the tumor would rely on their efficient presentation by DCs for a proper priming / activation of CTLs. These CTLs would have the capacity to migrate back to the tumor and to specifically lyse cancer cells in an antigen-dependent manner. The killing of cancer cells would finally lead to the further release and presentation of cancer antigens leading to a virtuous cycle known as the cancer-immunity cycle (25). However, although spontaneous CTL responses against cancer antigens have been described, they usually do not provide protection against tumors (25). This poor anti-tumoral response may be due to many mechanisms including failure to recognize tumoral peptides as foreign antigens and evolution of the tumor to escape recognition by CTLs. Two well described examples of therapeutically exploited molecular mechanisms limiting the anti-tumoral immune response are the inhibitory CTLA4 receptor (cf. CTL activation) and the PD-L1/PD-1 system (27). The PD-1 receptor is a negative regulator expressed by activated CTLs. Its ligands, PD-L1/2, can be expressed and presented by tumor cells, thereby causing local immunosuppression within the tumor bed (27). Many other mechanisms have been described by which tumors may escape immune surveillance, in part by creating an immunosuppressive microenvironment through the secretion of specific cytokines/chemokines (e.g. CXCL12), and the recruitment of immunosuppressive cells, such as regulatory T cells, tumor-associated macrophages, myeloid suppressor cells and suppressive dendritic cells (28,29). Immunoediting is another important

concept and has been divided into 3 distinct phases defined as elimination, equilibrium and escape (30). During the elimination phase, the immune system recognizes and eliminates tumor cells. If only partial tumor elimination occurs, a temporary state of equilibrium may develop between the immune cells and the emerging tumor cells (31). With time, particular tumor clones evolve and acquire various capacities allowing such cells to resist and suppress the antitumor immune response, and leading to the escape phase. This model is supported by the observation that immunosuppressed patients are at increased risk of developing tumors, including these that are not of known viral etiology (32). In line with this theory, animal experiments have also shown that carcinogen-induced tumors from immunodeficient mice (Rag2<sup>-/-</sup>) are more immunogenic in syngeneic transplantation models than tumors from wild-type control mice (31).



**Figure 2.** The cancer-immunity cycle is depicted as seven steps (25). Treatments and techniques aiming at enhancing the cycle are listed at the steps that they promote.

## Immunotherapy

Immunotherapy aims at triggering *de novo* anticancer immune responses and/or at enhancing the cancer-immunity cycle (Fig. 2) and may be classified as passive and active immunization modalities.

### Passive immunotherapy

By definition, passive immunotherapy protocols aim at enhancing the immune response by the exogenous administration of molecules with immunomodulatory properties, such as cytokines, antibodies, or of T lymphocytes (also defined as adoptive cell transfer). Administration of IL-2 and IFN $\alpha$  are among the first examples of immune modulators that have been used for the treatment of melanoma patients, allowing enhanced antigen presentation and CTL priming and activation (12,33). More recently, strategies of antibody-mediated blockade of CTLA4 and/or PD-1 have been developed and showed a significantly improvement of the clinical outcome of metastatic melanoma patients (34). For instance, Ipilimumab (anti-CTLA4) was shown to induce an increase in both progression-free and overall survival of advanced stage melanoma patients in randomized phase III trials (35). Nivolumab (anti-PD-1), is currently tested as a single agent in phase III trials (36,37). Moreover, a phase I study suggests that the combination of nivolumab and ipilimumab may further increase the treatment response rate with an acceptable safety profile (38).

Other more recent protocols include adoptive cell therapies where patients' CTLs are collected and expanded *in vitro* before being reinfused (39). This simplistic summary hides many technical parameters that constitute as many hurdles (e.g. the need for immune depletion prior to reinfusion), as opportunities for further development of this technique (e.g. the possibility of T cell reprogramming to express modified TCRs or engineered chimeric antigen receptors). Many of these different protocols are currently in clinical trials (39). Of note, adoptive cell therapy was recently shown to induce durable remission (> 8 years) in a significant proportion of melanoma patients that had relapsed after multiple other treatments (40) and is currently being compared to ipilimumab in a phase III study (clinicaltrials.org ID: NCT02278887).

### Active immunotherapy

Active cancer immunotherapy, or therapeutic vaccination, combines a multitude of protocols aiming at generating and/or boosting tumor-specific immune responses *in vivo* by the direct exposure to antigens. Active immunotherapy has been largely used in the clinic but the benefits of this approach have been, with rare exceptions, rather unsuccessful in generating potent tumor-reactive T cells allowing the control and destruction of tumor cells (41). A major hurdle is the appropriate and successful activation and expansion of those rare *in vivo* tumor-specific T cell precursors of low TCR avidity in order to mount a full protective anti-tumoral immune response. To achieve this ultimate objective, a variety of vaccine formulations (comprising both the antigen form and adjuvant composition) have been developed and were tested in clinical trials, including autologous antigen-loaded dendritic cells, viral vectors, and proteins/peptides with chemical adjuvants (to stimulate native immune pathways) or in combination with helper peptides (to stimulate T helper-dependent cross-presentation) (39,42). Different types of antigens have been used in these approaches from whole tumor cell lysates to short immunogenic peptides. Vaccination with defined proteins / peptides has the disadvantage of being restricted to specific HLA variants for antigenic presentation and might, therefore, be limited to subgroups of patients. It has, however, the theoretical advantage of directing the immunization to controlled specific targets, thereby limiting potential off-target effects.

Sipuleucel-T (Provenge®) is a DC-based vaccine, which harbors the PAP antigen and is the first non-viral cancer vaccine approved by the FDA for asymptomatic and minimally symptomatic metastatic castration-resistant prostate cancer (39). Many other target-formulation pairs are currently being tested in the clinic, including some against melanoma. The most advanced melanoma target is gp100, a melanocyte differentiation marker, with a randomized phase III trial showing increased progression-free and overall survival after immunization with a gp100 peptide emulsified in Montanide ISA-51 (a water-in-oil emulsion) and IL-2 treatment vs. IL-2 alone (43). Other tumor-specific targets include cancer testis antigens, such as NY-ESO-1, and differentiation antigens, such as tyrosinase and Melan-A/MART1 (44), the latter being the subject of this study. Yet, further progress is required to improve the vaccines, with the goal to increase the strength of immune activation.

### Melan-A peptide vaccination study (LUD00-018 study)

#### Design

The LUD00-018 study (clinicaltrials.gov ID: NCT00112229) was an interventional phase I clinical trial, which aimed at determining whether vaccination with tumor antigenic peptides and both CpG7909 and Montanide adjuvants could induce a T-cell based immune response in stage III-IV melanoma patients (45). The design of the trial was built on the prior observation that co-administration of the CpG7909 adjuvant (a single-stranded synthetic DNA composed of deoxycytidyl-deoxyguanosin oligodeoxynucleotides known to trigger TLR9) to the Melan-A peptide and Incomplete Freund's Adjuvant (IFA) vaccine formulation induced much higher Melan-A-specific CTL responses as compared to Melan-A and IFA as the only adjuvant. At that time, these results were obtained with the analog (ELAGIGILTV) HLA-A\*0201-restricted Melan-A<sub>26-35</sub> peptide, modified by an Ala to Leu substitution at position 27, allowing the enhanced generation of tumor-specific CD8 T cell responses (46,47). Thus, the clinical introduction of CpG7909 as adjuvant resulted in such a drastic increase in tumor-specific frequencies, raising the question whether immune CD8 T cell responses against the

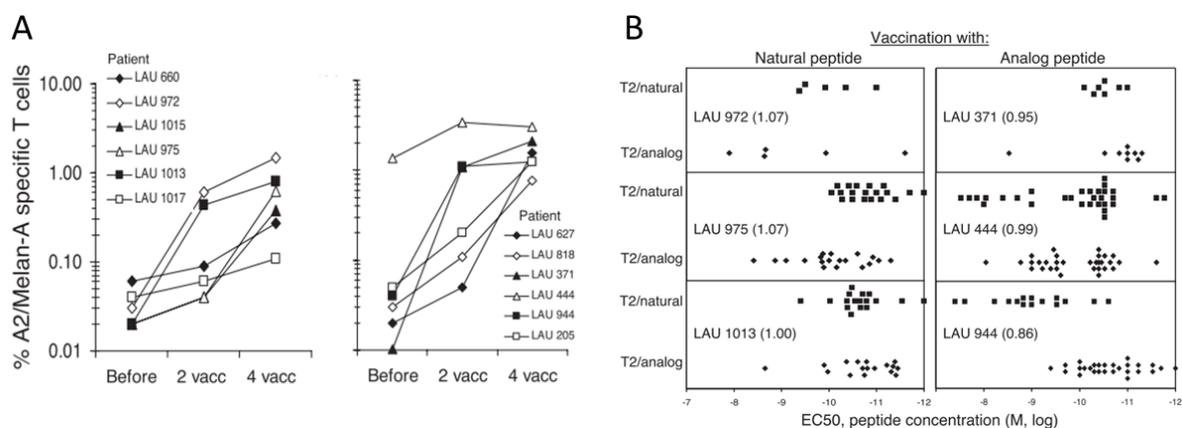
natural (EAAGIGILTV) Melan-A<sub>26-35</sub> peptide could as well be observed directly *ex vivo* and compared to the analog (ELA) T-cell mediated response. The LUD00-018 vaccine formulation was therefore based on the combination of both IFA and CpG7909 as adjuvants together with low doses of the natural (EAAGIGILTV) or analog (ELAGIGILTV) Melan-A peptide, either alone or in combination with the tyrosinase YMD peptide.

The study included HLA-A2-positive adult volunteers with a histologically confirmed melanoma expressing Melan-A with or without tyrosinase. Patients with clinically significant heart disease, a history of immunodeficiency or autoimmune diseases, coagulation or bleeding disorders as well as patients that were unlikely to complete the study because of critical health conditions were excluded.

The primary outcomes of the study were the immunological efficacy and the clinical safety during one year after the first vaccine dose vs. the pre-vaccine baseline. Efficacy was evaluated by the rates of tetramer-positive CTLs in peripheral blood and *in vitro* functional assays (Elispot IFN $\gamma$  assays) (45). For patients with measurable tumor burden, the tumor response was monitored as a secondary outcome measure.

### Cytotoxic T lymphocyte response characterization

One of the main interests of the LUD00-018-based study was to apply the novel CpG7909 vaccine formulation in melanoma patients and to directly compare vaccination with the natural (EAA) versus the analog (ELA) Melan-A peptide. CpG-mediated vaccination triggered a fast and strong tumor-specific CTL response that was detectable directly *ex vivo*, readily after two rounds of vaccination, when using Melan-A-specific multimers combined to  $\gamma$  flow cytometry as experimental read-out (Fig. 3A). Moreover, compared with natural peptide vaccination, the analog peptide induced T cell frequencies that were approximately two-fold higher. Importantly, it is also the first synthetic vaccine formulation to consistently induce *ex vivo* detectable T cell responses even when using the natural tumor peptide antigen (45).



**Figure 3. A.** Melan-A+ CD8+ T cell-specific frequencies before and after 2 or 4 natural/EAA (left panel) or analog/ELA (right panel) peptide vaccination doses. Of note, patient LAU444 showed high Melan-A-specific frequencies before vaccination due to a spontaneously primed tumor-specific CD8 T cell response (45). **B.** The killing potential of Melan-A+ CD8+ T cell clones isolated from patients vaccinated either with the natural (left panels) or the analog (right panels) peptide was measured in cytotoxic assays with T2-pulsed Melan-A peptide (natural versus analog) of increasing concentrations and were plotted as EC50 values for each indicated patient. This figure was adapted from (45).

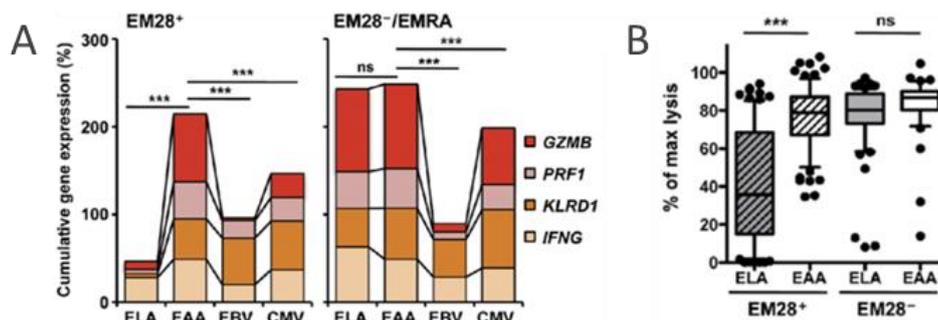
Another striking finding from this study (45) was that anti-tumoral CTL responses following natural peptide vaccination were of better quality, with superior tumor recognition, resulting in stronger protein expression of effector molecules (Granzyme B and Perforin), target cell killing, and cytokine production, as compared to vaccination with the analog peptide (Fig. 3B).

Due to the strong immunogenicity induced by vaccination with peptide and the CpG adjuvant, this vaccine approach also provided the unique opportunity to study T cell-mediated priming and TCR

repertoire selection in humans. Most vaccinated patients showed progressive restrictions in the T cell clonotype diversity along cell differentiation (from EM28+ to EM28-) with preferential expansion of several co-dominant clonotypes of intermediate to high frequencies, irrespectively of whether the natural or the analog peptide was used for vaccination (48). In line with this study, we demonstrated that T cell repertoires generated against natural or analog Melan-A peptides following vaccination in melanoma patients exhibited slightly distinct, but otherwise overlapping and structurally conserved  $\alpha\beta$  TCR features (49). For instance, a strong preference of the TCR V $\alpha$ 2.1 segment usage and a conserved GLG motif within the CDR3 $\beta$  were observed in nearly all patients, regardless of the type of peptide (natural vs analog) vaccine.

Together, these observations revealed that while tumor-specific T cell responses generated following vaccination with natural and analog tumor peptides exhibited similar T cell differentiation and TCR $\alpha\beta$  repertoire selection attributes (48,49), the natural peptide-induced T cells showed an enhanced overall functionality compared to the analog peptide (45,50). These data confirmed that vaccination trials using analog peptides require careful re-evaluation with regards to the risk of activating T cells with imprecise antigen specificity.

We recently performed highly sensitive single cell gene expression profiling allowing the direct *ex vivo* characterization of individual tumor-specific T cells from melanoma patients with the aim to further investigate the observed functional differences between natural and analog Melan-A<sub>26-35</sub> peptide vaccination (51). Importantly, differential gene expression features could only be identified when tumor-specific T cell responses were carefully dissected at the memory (EM28+) and effector (EM28-) subpopulation level. While memory EM28+ T cells derived from patients receiving the analog ELA vaccine typically expressed genes associated to memory/homing function, natural peptide vaccination induced tumor-reactive EM28+ T cells with frequent co-expression of memory/homing- and effector-related genes. Surprisingly, the latter cells exhibited comparable levels of effector genes with those found in the differentiated EM28- tumor-specific cells or protective CMV-specific T cells (Fig. 4A). These gene expression differences nicely correlated with the enhanced killing potential of the EM28+ EAA-specific T cells in *in vitro* assays (Fig. 4B). These observations (51) revealed a previously unknown level of gene expression diversity, suggesting that such broad functional gene expression signatures within antigen-specific T cells may be critical for mounting efficient responses against pathogens or tumors.



**Figure 4.** A. Effector gene expression rates in early- (EM28+) and late- (EM28-) differentiated cells in patients vaccinated with the natural (EAA) or the analog (ELA) peptide and in EBV and CMV-infected control subjects. B. Killing potential of Melan-A+ CD8+ T cell clones isolated from early- (EM28+) and late- (EM28-) differentiated cell populations from EAA versus ELA-vaccinated patients. Adapted from (51).

## Project rationale and hypothesis

The large majority of the results pertaining to the LUD00-018 clinical trial were obtained at relatively late time points after the start of vaccination (> 6 months). In the present study, we sought to assess whether the differences observed in the expression of effector mediators and cell functionality between memory EM28+ T cell following native and analog peptide vaccination were established early following the start of vaccination. To that end, we characterized and compared the expression of effector- and memory-related genes in single memory (EM28+) and effector (EM28-) tumor-specific CD8 T cells from seven patients receiving either the natural/EAA (n = 4) or the analog/ELA (n =3) peptide vaccine at early time points (< 3 months). Using this approach, we further addressed the question whether there was any discernible evolution in the expression of these gene expression signatures across time.

Alongside the functional gene expression analysis, we aimed at characterizing the evolution of the tumor-reactive T cell clonotype repertoire across time. In particular, it remained unclear whether the peptide/CpG7909 vaccination allowed the preferential selection and expansion of Melan-A-specific T cells that were already pre-existing before the start of immunotherapy (also defined as the spontaneously generated/tumor-induced vaccine-boosted model) or whether this vaccine formulation induced the selective activation and expansion of *de novo* rare tumor-specific T cell clones (defined as the vaccine-induced model). Moreover, both natural/EAA and analog/ELA peptide-vaccinated patients showed previously a high prevalence of relatively few co-dominant tumor-specific T cell clonotypes, which were mostly enriched in the differentiated EM28- subset, at later time points (48). We therefore asked whether this dominant T cell repertoire selection was established early after the start of vaccination and whether it occurred simultaneously in both cohorts (ELA vs EAA) of patients. For that purpose, we determined the clonotype of each individual tumor-specific T cell by sequencing the CDR3 regions of TCR $\alpha$  and TCR $\beta$  mRNAs.

Our overall goal was to identify gene expression signatures, which correlated with the increased T cell functionality and long-term persistence of particular tumor-reactive T cell clonotypes in melanoma patients following peptide vaccination.

## Aims

### Aim 1. Characterization of the effector and memory gene expression phenotype of Melan-A-specific CTLs across time and at the single cell level

- a. Comparison of gene expression phenotype between native/EAA and analog/ELA peptide vaccinated patients
- b. Comparison of gene expression phenotype evolution across time after peptide vaccination
- c. Comparison of gene expression phenotype before and after peptide vaccination

### Aim 2. Characterization of the TCR clonotype repertoire across time and at the single cell level

- a. Analysis of TCR clonotype repertoire selection and evolution across time
- b. Comparison between vaccine-induced or tumor-primed/vaccine-boosted models of the Melan-A-specific CTL response
- c. Characterization of the gene expression phenotype of dominant tumor-specific T cell clonotypes

## Material and methods

### Samples and CD8+ T cell sorting

Ficoll-Hypaque (Pharmacia) centrifuged peripheral blood mononuclear cells (PBMC) samples that were originally drawn from each indicated patients at indicated time points before and after vaccination (Table 1) were cryopreserved in RPMI 1640, 40% FCS and 10% DMSO in the vapor phase of liquid nitrogen until further use.

PBMCs were stained using the following combination: FITC anti-CD28 (BD Biosciences), HLA-A\*0201 analog/ELA Melan-A/MART-1-26-35 (A27L) multimers (TCMetrix Sàrl), ECD anti-CD45RA, APC A750 anti-CD3 and Krome Orange anti-CD16 (Beckman Coulter), and BV(421) anti-CCR7 (Biolegends), Alexa700 anti-CD8 (eBioscience) with Vivid Aqua (Life Technologies) for live cell discrimination. Samples were acquired and single cells sorted using a BD FACSAria cytometer (BD Biosciences) and the FACSDiva software (v.1.6, BD Biosciences). Photo Multiplier Tube (PMT) voltages were set using unstained PBMCs and compensations were calculated using single stained PBMCs. The following three tumor-specific CD8 T cell sub-populations were sorted as indicated (Table 1): the total Tetramer+ population with Vivid Aqua/CD16- CD3+ CD8+ Melan-A+ and the effector-memory (EM) CD28+/- subsets with Vivid Aqua/CD16- CD3+ CD8+ Melan-A+ CD45RA- CCR7- CD28+ (defined thereafter as EM28+ cells) and Vivid Aqua/CD16- CD3+ CD8+ Melan-A+ CD45RA- CCR7- CD28- (defined thereafter as EM28- cells).

Peptide	Patient	Time point	Days post-vaccination	n (Tet+)	n (CD28+)	n (CD28-)	Analyzed TCR $\alpha$ / $\beta$ repertoire	
EAA	LAU972	Early	45		80	34	TRBV1-3, 5*, 7, 13-14, 16-17	
			94		59	31		
		Late	374		39	39		
	LAU1013	Early	101		80	37	TRBV1, 3, 7, 9, 13-14, 17	
		Late	261		44	38		
	LAU1015	Pre-vaccine	-5	40			TRBV1, 3, 6-8, 13-14, 16-17	
		Early	93		20	14		
		Late	385		7	49		
	LAU1106	Pre-vaccine	0	41			TRBV1, 3, 7, 13-14, 17	
		Early	91		66	50		
		Late	675		53	52		
	ELA	LAU444	Pre-vaccine	-336		46	40	TRAV2
Early			84		71	43	TRBV1, 3, 7, 13-14, 17	
Late			514		60	61		
LAU618		Late	430		43	36	TRBV3, 13, 17	
LAU1164		Pre-vaccine	-7	17			TRBV1, 3, 7, 13-14, 17	
		Early	38		57	41		
		Late	241		20	22		
				633			22	-

**Table 1.** Number of analyzed cells (n) after quality controls ( $\geq 2$  housekeeping gene expressed and  $\leq 1$  TCR $\alpha$ / $\beta$  pair of genes detected per individual cell sample) according to patient, time points and sub-population. TRBV and TRAV genes whose expression was assayed are listed for each patient. Cells highlighted in green indicate that the corresponding data was previously generated by Gupta *et al.* (51). \* Primers for TRBV5 expression were specific for TRBV5-1 subtypes. EAA, natural peptide; ELA, analog peptide.

### cDNA synthesis and amplification

Single cells were sorted in each well of 96 well plates and directly lysed in 15  $\mu$ l of oligo-dT primed reverse transcription mix (8 U MMLV-reverse transcriptase, 8 U RNAsin, 250 nM 20-mer-oligo-dT, 500  $\mu$ M dNTPs, 3 mM MgCl<sub>2</sub>, 33 mM  $\mu$ g/ml tRNA, 10 mM DTT, 80 mM KCl, 50 mM Tris-Cl pH 8.3, 1.25% Triton X-100). The cDNA was synthesized by incubation at 37 °C for 1 h and the enzymes inactivated

by incubating at 90 °C for 3 min. cDNA samples were stored at -80 °C. The cDNA was precipitated O/N at -80°C with 7.5 µl 7.5 M NH<sub>4</sub>Cl, 3 µl 10 µg/µl glycogen and 45 µl EtOH, pelleted 20 min at 13'000 rpm at 4 °C, washed with cold 70% EtOH, pelleted again and dried at room temperature. The cDNA was polyA-tailed in 5 µl terminal-deoxynucleotide-transferase reactions (2 U TdT, 500 µM dATP in 1x supplied buffer, Promega) for 30 min at 37 °C. The enzyme was inactivated for 3 min at 90 °C.

The cDNA was amplified as described previously with 40 PCR cycles with the Iscove-dT primer (5'-CAT GTC GTC CAG GCC GCT CTG GGA CAA AAT ATG AAT TCT TTT TTT TTT TTT TTT TTT T-3') in the following final conditions: 5 U Taq polymerase, 200 nM Iscove-dT primer, 200 µM dNTPs, 2 mM MgCl<sub>2</sub>, 0.1 µg/µl BSA, 50 mM KCl, 10 mM Tris-Cl pH 8.8, 0.5% Triton X-100. Reactions were denatured for 3 min at 90 °C before enzyme addition and amplification with 5 initial cycles at low annealing temperature (50 s. at 94 °C, 2 min. at 37 °C, 9 min. at 72 °C) followed by 35 cycles at high annealing temperature (50 s. at 94 °C, 90 s. at 60 °C, 8 min. at 72 °C) and a final elongation step of 8 min. at 72 °C. The amplified cDNA was stored at -80 °C.

### Target gene amplification, detection and sequencing

Housekeeping, effector and memory genes were detected by PCR using previously published specific primer pairs (48,49,51). Briefly, 1 µl of amplified cDNA was used as template in 20 µl reactions (0.5 U KAPA HotStart Taq polymerase, 200 nM each primer, 200 µM dNTPs, 1.5 mM MgCl<sub>2</sub>, 0.1 µg/µl BSA, 50 mM KCl, 10 mM Tris-Cl pH 8.8). The enzyme was activated for 2 min at 94 °C and the targets were amplified with 40 cycles (30 s. at 94 °C, 45 s. at 58 or 60 °C as previously published, 60 s. at 72 °C) with a final elongation step of 10 min. PCR products were resolved on 2.5 agarose gels and scored with the help of a positive control using PBMC cDNA as template (0 = no specific product, 0.5 = minor specific product, 1 = major specific product). TCRα/β transcripts were detected with the same protocol with the following modifications: the forward primer concentration was increased to 300 nM and MgCl<sub>2</sub> concentration was decreased to 1.25 mM. The final elongation step of the PCR was lengthened to 45 min at 72 °C. 5 µl of each PCR product was sent for sequencing (Fasteris) using the reverse Cα/β primer used for amplification to determine the sequence of the CDR3 region.

### Data analysis

#### Effector / memory gene expression analyses

Cells were considered as positive for a given gene when its expression score was equal to or greater than 0.5. Out of 1732 analyzed cells, 232 cells expressed less than two detectable housekeeping gene and thus were filtered out from all subsequent analyses. Moreover, 48 cells for which two TCRβ mRNAs were detected within one single sorted cell sample were also filtered out, leaving a total of 1452 cells that were included in further analyses.

Cluster analysis of effector and memory genes was performed using the pooled data of all single cells from the early and late time points. Distance between each gene was calculated with a Pearson correlation coefficient. Dimension reductions were performed by aggregating the expression data of two genes. The aggregation variable was considered as positive for a given cell if one of the two genes was found expressed in that cell.

The distribution bias ( $\Delta$ ) of two compared cell populations (A and B) in a given expression effector/memory gene expression pattern was defined as the difference of the ratios of number of cells ( $n$ ) with the effector/memory marker combination to the total number of cells within each population:

$$\Delta_{AB}(\text{effector/memory}) = \frac{n_A(\text{effector/memory})}{n_A} - \frac{n_B(\text{effector/memory})}{n_B}$$

Associated  $p$  values were calculated using the hypergeometric distribution and reflect the probability of observing number of cells with a defined effector/memory gene expression pattern in population A given the number of cells with the same expression pattern in the total population (A + B):

$$P(n_A(\text{effector/memory}) \mid n_A, n_{A \cup B}(\text{effector/memory}), n_{A \cup B})$$

### Clonotype repertoire analyses

Clonotypes were considered for further analysis only when a positive TCR $\beta$  PCR amplification signal was observed and an interpretable TCR CDR3 $\beta$  sequence could be determined. TCR $\alpha$  chain PCRs were performed in a second step only in the single T cell samples of known and defined TCR $\beta$  CDR3 sequences.

Variable regions of TCR $\beta$  are referred to according to the nomenclature proposed by Arden and colleagues (52). Clonotypes that were found in more than two individual cell samples from the same patient were attributed a number, and are referred to by the following identification tag format: BV[TCR $\beta$  variable region number]-[clonotype number].

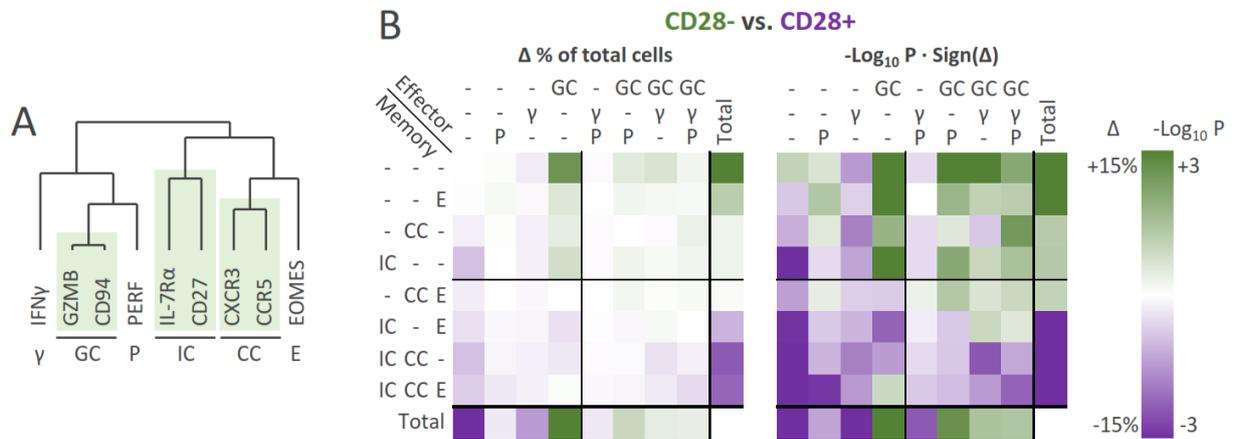
## Results

For this study, we generated a large library of cDNAs ( $n = 1'732$ ) isolated from individual tumor-specific CD8 T cell subpopulations following vaccination of patients with either the natural ( $n = 4$ ) or the analog ( $n = 3$ ) Melan-A<sub>26-35</sub> peptide vaccine together with IFA and CpG7909 (Table 1). For this purpose, we previously developed a strategy consisting of cell lysis and cDNA synthesis in a single-step procedure, followed by a modified PCR protocol that relies on the detection of specific cDNAs after global amplification of expressed mRNAs (51). Samples from melanoma patients (LUD00-018 study) were taken before vaccination (pre-vaccine), and at early ( $\leq 3$  months) and late time points ( $> 6$  months) after vaccination, and were processed as described in the *Material and methods* section. For pre-vaccine time points, the Melan-A-specific T cells showed mostly a naïve-like phenotype (CCR7+ CD45RA+; data not shown) and total live single CD3+ CD8+ Melan-A+ T cells were directly sorted by FACS (defined as Tet+). After vaccination, the so-called effector memory cells (CD45RA- CCR7-) constituted the largest CTL population and were sorted into the following two subpopulations; early- (CD28+) and late-differentiated (CD28-) EM cells. The cDNA of each sorted cell was synthesized and amplified, and all generated single cell samples were tested for the expression of housekeeping genes (B2M, GAPDH and RPL13A). 1'500 cells passed the first quality control by expressing at least two detectable genes. Of these, 48 additional cells were excluded as two TRBV sequences were detected presumably because of contaminating cells during sorting, leaving 1452 cells that were included in the following analyses. The selected single cell samples were further subjected to specific effector-related (GZMB, CD94, IFN $\gamma$  and PERF) and memory-related (IL-7R $\alpha$ , CD27, CXCR3, CCR5 and EOMES) gene expression PCRs as described in the *Material and methods* section.

The effector and memory gene expression data comprised nine binary variables that created a total of 512 possible gene expression combinations. To be able to detect statistically meaningful enrichments for each of these potential expression patterns, we sought to reduce the number of variables, in order to increase the number of single cells per bin. To that end, we performed a cluster analysis that comprised the effector- and memory-associated gene expression data of all patients and all populations across early and late time points to guide the dimensional reduction approach (Fig. 5A). Effector and memory genes segregated in two distinct clusters as expected given their known differential gene expression patterns found in EM28+ vs. EM28- cells (24; Fig. 6B & C). We selected three pairs of genes that showed the least distance from each other and aggregated their data into the following single variables; GZMB/CD94 (defined thereafter as GC), IL-7R $\alpha$ /CD27 (defined as IC), and CXCR3/CCR5 (defined as CC). The aggregation variables were assigned the value of 1 if at least one of the two genes was expressed or 0 if none of them were expressed. The dimensional reduction strategy allowed us to limit the number of variables (including the gene expression paired groups) to a total of six, with three variables defining the effector gene-associated expression (IFN $\gamma$ , GZMB/CD94 and PERF) and three variables defining the memory gene-related expression (IL-7R $\alpha$ /CD27, CXCR3/CCR5 and EOMES). The number of potential gene expression combinations was thereby reduced to 64, increasing the average number of cells per bin to about 23, an acceptable number to detect biases in the gene expression distribution (data not shown).

To test our analytical approach and the quality of our new set of single cell data, we first asked whether effector and memory gene expression differences could be observed between the well-characterized EM28+ and EM28- tumor-reactive CD8 T cell subsets. For that purpose, we pooled all native/EAA and analog/ELA derived cell samples from post-vaccination time points and calculated the relative bias of distribution between cells of the two pooled groups (EM28+ vs EM28-) for each expression pattern bin and plotted the data as a heat map (Fig. 5B; left panel). The distribution biases of cells with the various effector or memory expression patterns were also plotted in the "total" row and column, respectively. The  $p$  values associated to the distribution biases of each bin were calculated using the hypergeometric distribution, and their absolute log<sub>10</sub> values were signed according to the direction of the bias (+15% vs - 15%) and were plotted as a second heat map (Fig. 5B; right panel).

As expected, specific EM28+ CD8 T cells were overrepresented in bins with a high number of memory-associated genes (IL-7R $\alpha$ /CD27+ with at least one other memory-associated gene; Fig. 5B). In sharp contrast, the EM28- tumor-specific CD8 T cells were mostly accumulated in GZMB/CD94+ bins with limited numbers of expressed memory-associated genes (Fig. 5B). Together, these results are in line with previous reports from our group (24,48), showing the high quality of our new set of data and validating the analytical approach used in this work.



**Figure 5.** Comparison of effector and memory gene expression between early- and late-differentiated EM cells. **A.** Genes were sorted by cluster analysis according to their expression pattern. The results are displayed as a hierarchical clustering tree. The three groups of paired-genes whose expressions were aggregated to reduce the dimensions of the data are highlighted in green (GZMB/CD94 IL-7R $\alpha$ /CD27 and CXCR3/CCR5). Abbreviations are listed under each corresponding gene or paired-gene group. **B.** EM28+ (n = 699) and EM28- (n = 569) tumor-specific CD8 T cell of all post-vaccine time points were distributed according to effector (columns) and memory (rows) gene expression patterns defined in panel A. Left panel: bias of CD28- vs. CD28+ populations relative to the total cell number of each group. Right panel: p values associated with the bias of CD28- vs. CD28+ distribution. GC: GZMB+ and/or CD94+;  $\gamma$ : IFN $\gamma$ +; P: PERF+; IC: IL-7R $\alpha$ + and/or CD27+; CC: CXCR3+ and/or CCR5+; E: EOMES+.

### Aim 1a - Comparison of gene expression phenotype between native/EAA and analog/ELA peptide vaccinated patients

We previously showed a significantly lower killing activity associated to a lower expression of effector genes in EM28+ tumor-specific CD8 T cells generated following analog (ELA) peptide vaccination compared to the equivalent native (EAA)-specific T cells showing enhanced co-expression of effector and memory-associated gene patterns (51). These results were based on samples acquired at relatively late time points (> 6 months) after the start of the vaccination. Moreover, at that time, we had only included two patients from each peptide vaccination cohort. Therefore, the work described here provides an extension of this study with the in-depth characterization of a total of four native/EAA and three analog/ELA peptide vaccinated patients at the single cell level. We first focused on the late time point and assessed whether this new complementary set of gene expression data from individual native/EAA and analog/ELA-specific CD8 T cell subsets at late time points correlated to our prior report (51).

While no significant EM cell distribution bias among memory gene expression patterns could be observed between the native/EAA versus the analog/ELA peptide vaccinated cohort of patients (Fig. 6A, lower left panel), EM28+ tumor-specific T cells isolated following EAA peptide vaccination accumulated in GZMB/CD94+ expression pattern bins as compared to the equivalent analog/ELA cells (Fig 6A, lower left panel). Consistently, cumulative gene expression of effector genes were significantly increased in the EM28+ T cells from native/EAA peptide vaccination cohort, in comparison to their ELA counterparts (Fig. 6B).



Collectively, comparison of gene expression phenotype between the native and the analog peptide vaccine revealed only small cumulative gene expression (Fig. 6B-C) or gene expression distribution (Fig. 6A) differences in tumor-specific T cells from different subsets and at separated time points. The only exception was the strong enhanced expression of effector genes associated to memory gene co-expression observed within the native/EAA-specific EM28+ T cell subset at the late time point (Fig. 6A-C). Our new set of data therefore confirmed this previously observed phenotypic difference (51).

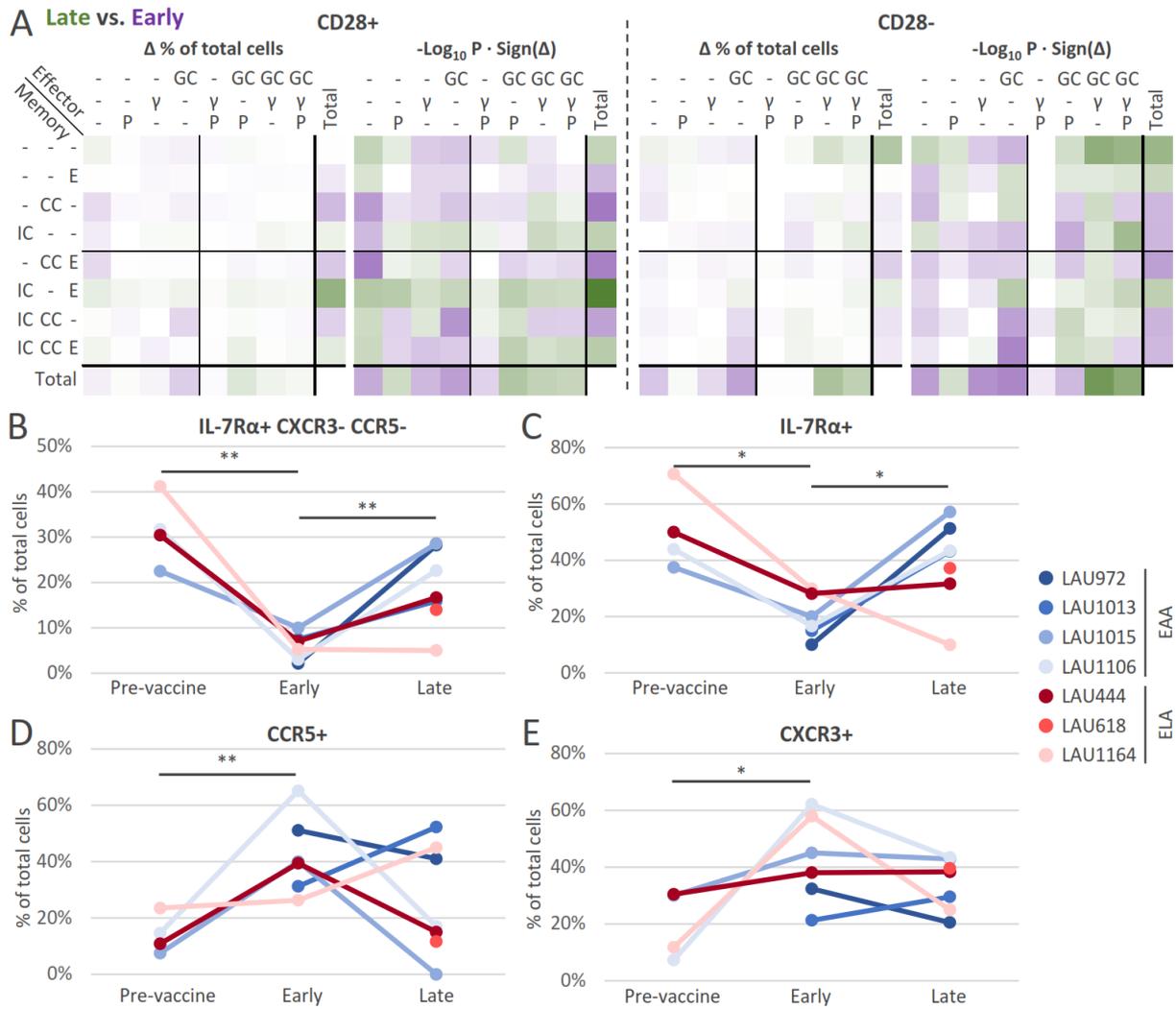
### Aim 1b - Comparison of gene expression phenotype evolution across time after peptide vaccination

We next addressed the question whether the Melan-A-specific T cells could evolve phenotypically in terms of gene expression distribution during the course of peptide vaccination, by comparing early versus late time points. Since no major differences except for the EM28+ cells at the late time point were observed between single cell samples from native and analog peptide vaccine, we decided to pool all native and analog peptide vaccine derived samples and to consider only the effector versus memory gene expression data within EM28+ and EM28- subpopulations across time (Fig. 7).

According to the effector/memory gene expression patterns, we observed an enrichment of effector gene signatures (GZMB/CD94+ IFN $\gamma$ +) in EM28- T cells at the late time point, when compared to the early time point (Fig. 7A right panel). A similar trend was as well found within the EM28+ T cell subset. These gene expression changes were close to statistical significance when the evolution of these populations were considered patient by patient, but did not meet the 0.05 cutoff ( $p = 0.056$ ; data not shown). The only obvious difference of EM28+ gene expression distribution was observed in the IL-7R $\alpha$ /CD27+ CXCR3/CCR5- EOMES+ expression pattern, which showed a 2.44 fold enrichment at the late time point ( $p < 0.0001$ ; Fig. 7A left panel). However, when we considered the contribution of each marker separately, CD27 and EOMES expression only slightly differed with time after vaccination (data not shown). Thus, excluding CD27 and EOMES from the analysis (and focusing on IL-7R $\alpha$ + CXCR3- CCR5- expression pattern) led to a 3.8-fold gene expression increase within EM28+ T cells from early to late time points (data not shown).

Since these observations were based on pooled EM28+ cells from both native/EAA and analog/ELA samples at the studied time points, we next evaluated the expression of the IL-7R $\alpha$ +CXCR3- CCR5- pattern as well as each gene separately (IL-7R $\alpha$ , CCR5 versus CXCR3) within the individual patients across time (Fig. 7B-E). All patients showed a statistically significant enrichment of the IL-7R $\alpha$ +CXCR3- CCR5- gene signature with time after vaccination ( $p < 0.01$ ; Fig. 7B). IL-7R $\alpha$  expression contributed the most to this increase ( $p < 0.05$ ; Fig. 7C), while differences in both CCR5 and CXCR3 expression showed no statistically significant variations between early and late time points (Fig. 7D and E).

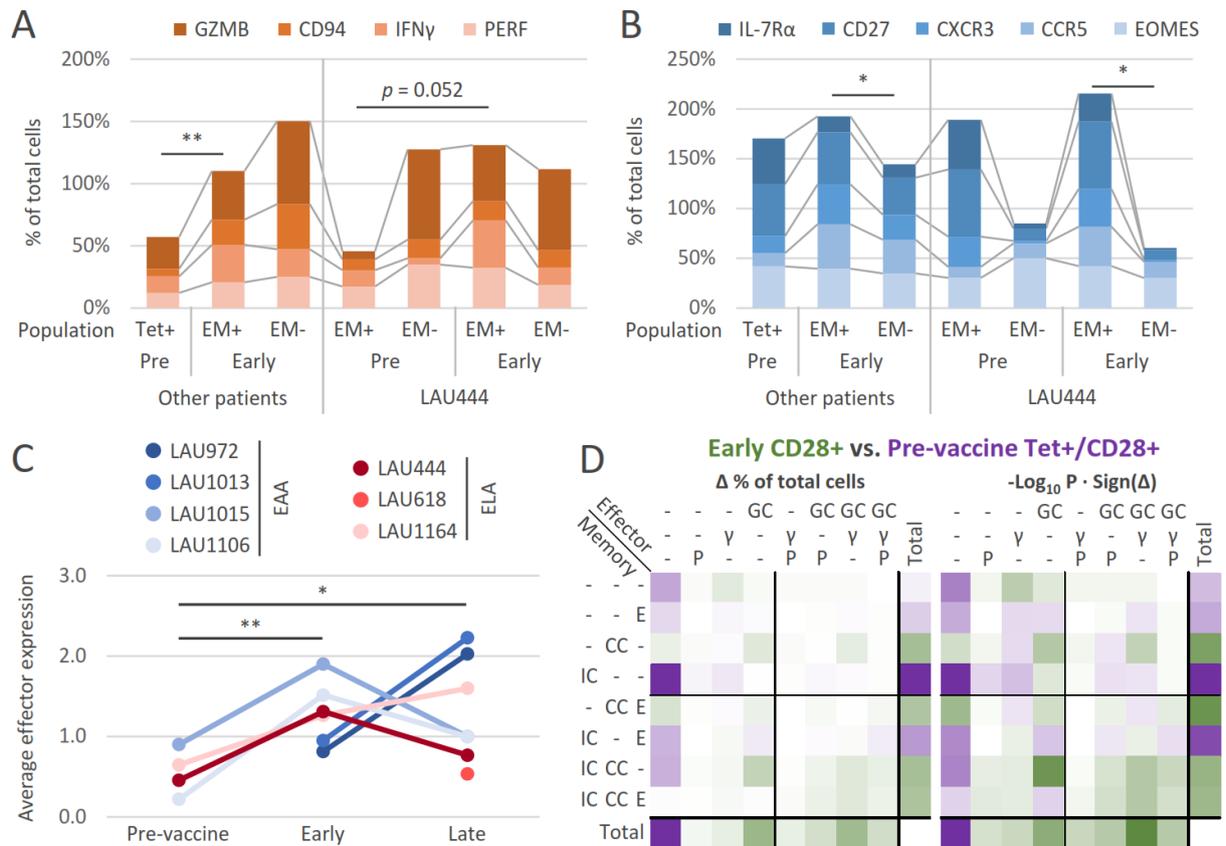
In summary, Melan-A-specific EM T cells showed an evolution in both memory and effector gene expression phenotype over time. The change in memory gene expression was mostly found within EM28+ cells with IL-7R $\alpha$  upregulation, while upregulation of effector gene expression was mostly observed in EM28- cells.



**Figure 7.** Comparison of effector and memory gene expression in tumor-specific T cell subsets between early (CD28+ n = 433; CD28- n = 250) and late (CD28+ n = 266; CD28- n = 319) time points. **A.** Distribution of EM cells at early vs. late time points according to effector (columns) and memory (rows) gene expression patterns in CD28+/- populations. Left panels of each subset: bias of cell distribution at late vs. early time points relative to the total cell number of each group. Right panel: *p* values associated with the bias of cell distribution at late vs. early time points. GC: GZMB+ or CD94+;  $\gamma$ : IFN $\gamma$ +; P: PERF+; IC: IL-7R $\alpha$ + or CD27+; CC: CXCR3+ or CCR5+; E: EOMES+. **B-E.** Evolution of EM28+ cell populations with the indicated memory gene expression patterns. Expression levels of each indicated gene and gene combination (IL-7R $\alpha$ + CXCR3- CCR5-) were plotted according to patients (EAA: blue shades; ELA: red shades) and time points. Associated *p* values were calculated with two-tailed unpaired *t* tests: \* *p* < 0.05; \*\* *p* < 0.01.

## Aim 1c - Comparison of effector and memory gene expression phenotype before and after peptide vaccination

We next asked whether we could observe differences in effector and memory gene expression patterns between tumor-specific CD8 T cells isolated from pre-vaccine and early time points. Melan-A-specific CTLs were rare at the pre-vaccine time point and the vast majority of them displayed a naïve-like phenotype (CD45RA<sup>+</sup> CCR7<sup>+</sup> CD28<sup>+</sup>; data not shown). For these reasons, we sorted all Melan-A+ CTLs for further analysis (defined as Tet<sup>+</sup> population). Patient LAU444 was an exception with the high spontaneous and pre-existing differentiated Melan-A-specific T cells, whereby both EM28<sup>+</sup> and EM28<sup>-</sup> cells could be isolated before vaccination (Fig. 8A and 8B; Table 1; 53).



**Figure 8.** Comparison of effector and memory gene expression in tumor-specific T cell subsets between pre-vaccine and early time points+. **A.** Cumulative expression of effector genes to patients (LAU444 or all others) and time points (pre-vaccine vs. early) and the subpopulation (Tet<sup>+</sup>, EM28<sup>+</sup> or EM28<sup>-</sup>). **B.** Cumulative expression of memory genes according to patients (LAU444 or all others) and time points (pre-vaccine vs. early) and the subpopulation (Tet<sup>+</sup>, EM28<sup>+</sup> or EM28<sup>-</sup>). **A-B.** Associated  $p$  values were calculated with two-tailed paired  $t$  tests: \*  $p < 0.05$ ; \*\*  $p < 0.01$ . **C.** Average number of expressed effector genes in Tet<sup>+</sup> or early-differentiated (CD28<sup>+</sup>) EM cells according to the type of peptide vaccine (EAA: blue shades; ELA: red shades) and each patient across time (pre-vaccine, early, late time points). Associated  $P$  values were calculated with two-tailed unpaired  $t$  tests: \*\*  $P < 0.01$ . **D.** Distribution of EM cells at pre-vaccine vs. early time points according to effector (columns) and memory (rows) gene expression patterns. Tet<sup>+</sup>/CD28<sup>+</sup> EM cells at pre-vaccine time points ( $n = 144$ ) were compared with EM28<sup>+</sup> cells at early time points ( $n = 433$ ). Left panel: bias of cell distribution at early vs. pre-vaccine time points relative to the total cell number of each group. Right panel:  $P$  values associated with the bias of cell distribution at early vs. pre-vaccine time points. GC: GZMB<sup>+</sup> or CD94<sup>+</sup>;  $\gamma$ : IFN $\gamma$ <sup>+</sup>; P: PERF<sup>+</sup>; IC: IL-7R $\alpha$ <sup>+</sup> or CD27<sup>+</sup>; CC: CXCR3<sup>+</sup> or CCR5<sup>+</sup>; E: EOMES<sup>+</sup>.

In line with their naïve-like surface phenotype (CD45RA<sup>+</sup> CCR7<sup>+</sup>), tumor-specific T cells expressed mostly memory/homing associated genes, with comparable cumulative levels to the EM28<sup>+</sup> T cells following early time point vaccination (Fig. 8B). Given their naïve-like phenotype and previous reports (24), we expected that the Tet<sup>+</sup> cells would also resemble to the “memory-like” EM28<sup>+</sup> T cells in terms of reduced effector gene expression levels. While this was the case, the pre-vaccinated EM28<sup>+</sup> T cells isolated from patient LAU444 also showed low expression levels of effector genes, which were similarly enhanced in early time point EM28<sup>+</sup> T cells (Fig. 8A). Thus, Melan-A/CpG vaccination significantly triggered the expression of various effector genes within Melan-A-specific T cells, including patient LAU444. Similar results were obtained when the average total number of expressed effector genes was calculated for each individual patient (Fig. 8C). Overall this up-regulation in the average effector gene expression was highly significant ( $p < 0.01$ ; Fig. 8C).

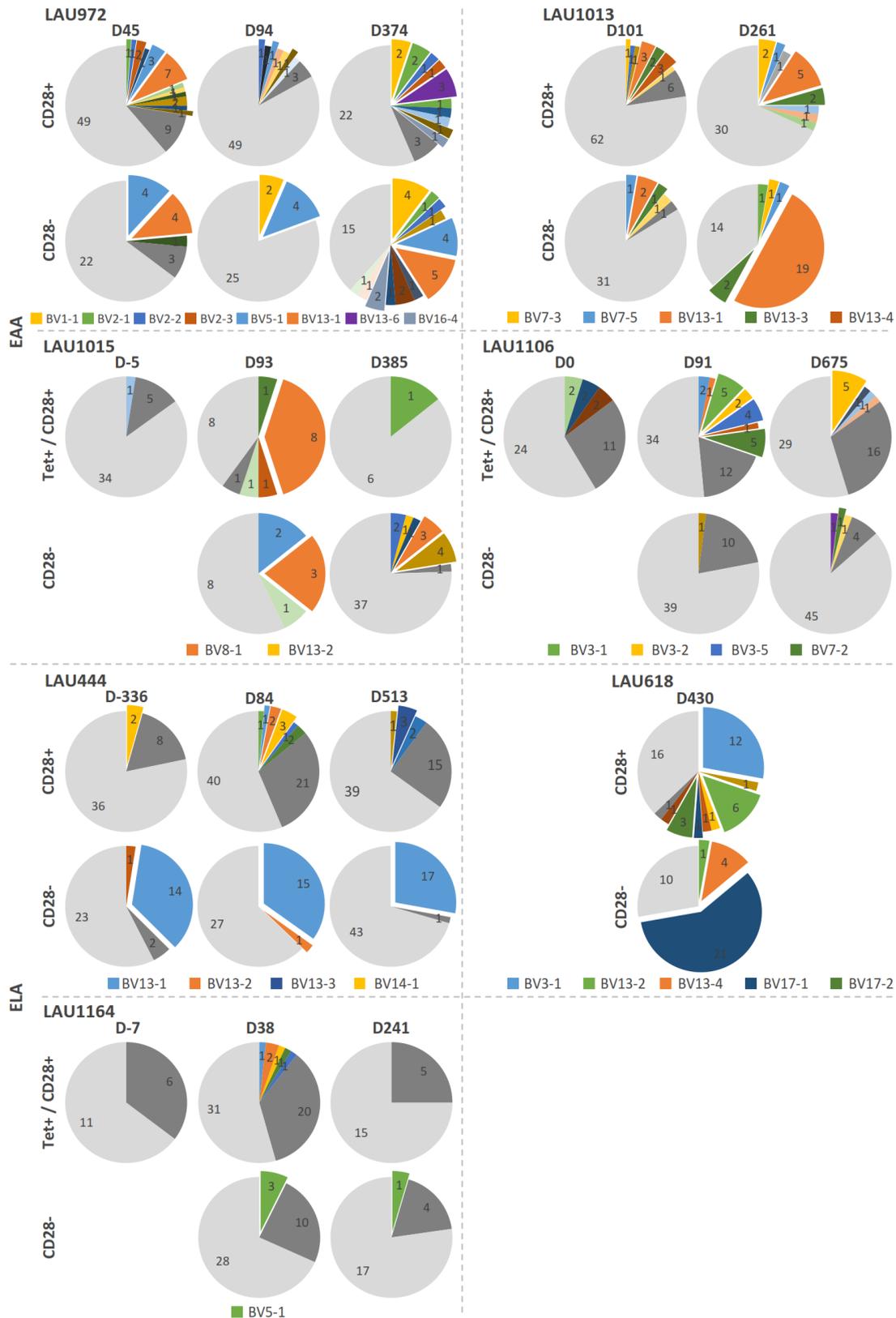
We next investigated whether there was any significant effector/memory gene expression bias between the early time point EM28<sup>+</sup> T cells and the Tet<sup>+</sup> T cells isolated before vaccination using our dimensional reduction approach (Fig. 8D). As expected, pre-vaccine Tet<sup>+</sup> were highly enriched in cells that did not express any effector-associated genes when compared to early time point EM28<sup>+</sup>. Interestingly, a significant fraction of these single cell samples also showed the preferential expression of the memory IL7-R $\alpha$ /CD27<sup>+</sup> CXCR3/CCR5<sup>-</sup> gene expression pattern (Fig. 8D; Fig. 7B). These biases were mirrored by a nearly symmetrical depletion in the IL-7R $\alpha$ /CD27<sup>-</sup> CXCR3/CCR5<sup>+</sup> gene expression signature. When the contribution of each gene marker was considered on its individual basis, IL-7R $\alpha$  expression showed a 2.53-fold down-regulation between pre-vaccine and early post-vaccine T cells ( $p < 0.05$ ; Fig. 7C), contrasting to the concomitant 3-fold-increase found for CCR5 ( $p < 0.01$ ; Fig. 7D) and to a lesser extent for CXCR3 gene expression levels ( $p < 0.05$ ; Fig. 7E) in EM28<sup>+</sup> T cells at early time points.

Together, these results show that peptide/CpG vaccination induced strong T cell differentiation of Melan-A-specific T cells, regardless of the type of peptide vaccine used (native/EAA or analog/ELA). Strikingly, the acquisition of the EM28<sup>+</sup> T cell phenotype following vaccination was associated with the significant increase in effector gene expression concomitant to the reduced CD127/IL-7R $\alpha$  memory gene levels.

## Aim 2a - Analysis of TCR clonotype repertoire selection and evolution across time

The second main goal of this study was to characterize the TCR repertoire in the same individual cells whose effector and memory gene expression were characterized. We targeted our analysis toward specific TCR $\alpha$ / $\beta$  families that we knew from previous studies to account for the majority of the dominant Melan-A-specific repertoire (48,49). Our method was based on PCR amplification using variable region-specific forward primers and reverse primers in their constant portions followed by sequencing of the CDR3 regions. Thereby, we systematically probed for TRBV1, 3, 7, 13, 14 and 17 transcripts in all newly included patient/time point. Additional TRBV families that were previously shown to play a prominent role in the Melan-A-specific repertoire of a minority of patients were also included as indicated in Table 1.

We previously reported the progressive restriction in the TRBV/CDR3 diversity along cell differentiation from EM28<sup>+</sup> to EM28<sup>-</sup>, with the preferential selection and expansion of few co-dominant T cell clonotypes (48,51). This restriction in the EM28<sup>-</sup> T cell repertoire was not specific to Melan-A peptide vaccination as it could also be observed in chronic antiviral CTL immune responses (54,55). Such selected tumor- and virus-specific T cell clonotypes were called “dominant” if they accounted for more than 1% of the population. Here, we further asked whether the selection of such dominant clonotypes also occurred in our single cell database and whether the course of vaccination could influence this selection and expansion. The 1% cut-off was here not applicable given the relatively low number of cells with identified clonotypes and its frequency variation from one patient to another. We therefore decided to use an arbitrary absolute threshold of 3 cells to define a clonotype as dominant. These clonotypes and their proportion found within each EM28<sup>+</sup> and EM28<sup>-</sup> subset are listed in Table 2 and are highlighted in Figure 9.



**Figure 9.** Clonotype repertoire analysis in EM28+ and EM28- T cell subsets in pre-vaccine, and early and late time points following peptide vaccination. Clonotype repertoire compositions are plotted as pie charts for each patient, subsets (Tet+, EM28+ or EM28-) and time points (pre-vaccine vs early vs late time points). Corresponding number of EM cells are indicated for each clonotype. Cells for which no TCR $\alpha/\beta$  transcript could be detected are depicted in light gray. Clonotypes that were detected only once in this study and were not previously detected in prior repertoire analyses of the same patients (48,49,51) were grouped together and depicted in dark gray. Dominant clonotype IDs are indicated in the color legends.

Consistent to these studies, the differentiated EM28- T cell subset was composed of several dominant clonotypes that were highly specific for each patient, and this TRBV/CDR3 preferential selection occurred in both type of peptide vaccine (native/EAA and analog/ELA) (Fig. 9; Table 2). Again, the proportion of dominant T cell clonotypes was globally reduced in the early-differentiated EM28+ T cell subset that was comprised of a more diverse TRBV repertoire, irrespective of whether native or analog peptide was used for vaccination.

Strikingly, many of the most prevalent clonotypes were present at both early and late time points following vaccination (Fig. 9): Without including LAU444, which showed a vaccination-independent Melan-A-specific CTL response, 64 of the 122 cells with identified clonotypes at the early time point displayed 11 clonotypes that were also detected at the late time point. This indicates that once established the clonal composition of the tumor-specific T cell response is kept stable along immunotherapy. Yet, we also identified a few prevalent clonotypes either at early or at late time point (e.g. BV13-2 for LAU1015 and BV13-3 and BV14-1 for LAU444), suggesting some level of plasticity in the repertoire selection.

Moreover, we calculated the prevalence of dominant T cell clonotypes within the total population of cells for each patient and type of peptide vaccine (ELA vs EAA), subset (EM28+ vs EM28-), and time point (early vs late) (Fig. 10A). As expected, differences were observed in the number of dominant clonotypes found between the less-differentiated EM28+ and the more-differentiated EM28- T cell subsets, with most patients displaying a higher proportion of dominant EM28- cells. However, this difference was only significant at the late time point ( $p < 0.05$ ) since LAU1106 and LAU1015 presented an increased frequency of EM28+ dominant T cell clonotypes at the early time point (Fig. 10A). It remains unclear whether this result is due to (i) a smaller repertoire restriction at early time point, (ii) the relative low number of individual cells with identified clonotypes in EM28+ ( $n = 109$ ) and EM28- ( $n = 144$ ) T cells, biasing the statistical analysis or (iii) the low TRBV/CDR3 clonotype identification rate, making up between about 15% to 70% of total specific T cells and depending on patient, subset and time point.

Previous TCR clonotyping (48,51) showed that 50 to 95% and 20 to 55% of EM28- and EM28+ tumor-specific T cells, respectively, had undergone preferential expansion. Thus, the rate of TCR TRBV/CDR3 clonotype identification was lower in our current study, and this was particularly true when we directly compared the data generated here to the previous one (i.e. EM28- T cells from LAU972 and LAU1013 at early versus late time point). Three main explanations may account for this discrepancy: (i) our TRBV targeted analysis ignored important TCR $\beta$  families that may possibly be present at early time point, (ii) the dominant clonotype frequencies were underestimated due to our new cut-off at 3% instead of 1%, and/or (iii) the sensitivity of our PCR-based technique was reduced. Consistent with the last argument, we were able to detect cells that had been previously tested negative for TRBV13 expression, when clonotypic PCRs were performed on the known TCR TRAV2/CDR3 clonotype sequence of BV13-1 clonotype of patient LAU444 (data not shown).

## Aim 2b - Comparison between vaccine-induced or tumor-primed/vaccine-boosted models of the Melan-A-specific CTL response

We next asked whether the dominant tumor-specific T cell clonotypes that were found at early and late time points following peptide vaccination were already present before the start of immunotherapy. Except for patient LAU444, who presented a pre-existing tumor-primed and vaccine-boosted dominant Melan-A-specific CD8 T cell clonotype defined as BVTR13-1/ELGTASY of high frequencies (53), we were unable to detect overlapping clonotypes between the pre-vaccine Tet+ cells and the early/ late time point EM cells (Fig. 9). The relatively low number of pre-vaccine cells ( $n = 184$ ) combined to the relative reduced clonotype identification efficiency may in part explain this absence of common clonotype. An alternative explanation may come from the fact that pre-vaccine Melan-A-specific T cells are mostly composed of naïve-like cells (Tet+) and display a higher clonal diversity, contrasting with the post-vaccinated specific cells undergoing drastic differentiation and clonotype selection (53 & unpublished observations). In such a scenario, one may not expect to find dominant

clonotypes, as those may be relative rare frequencies within the Melan-A-specific CD8 T cell compartment before vaccination. In line with this hypothesis, we could only identify the pre-vaccine dominant BV13-1/ELGTASY from patient LAU444, when focusing on the differentiated EM28- subset (Fig. 9).

### Aim 2c - Characterization of the gene expression phenotype of dominant tumor-specific T cell clonotypes

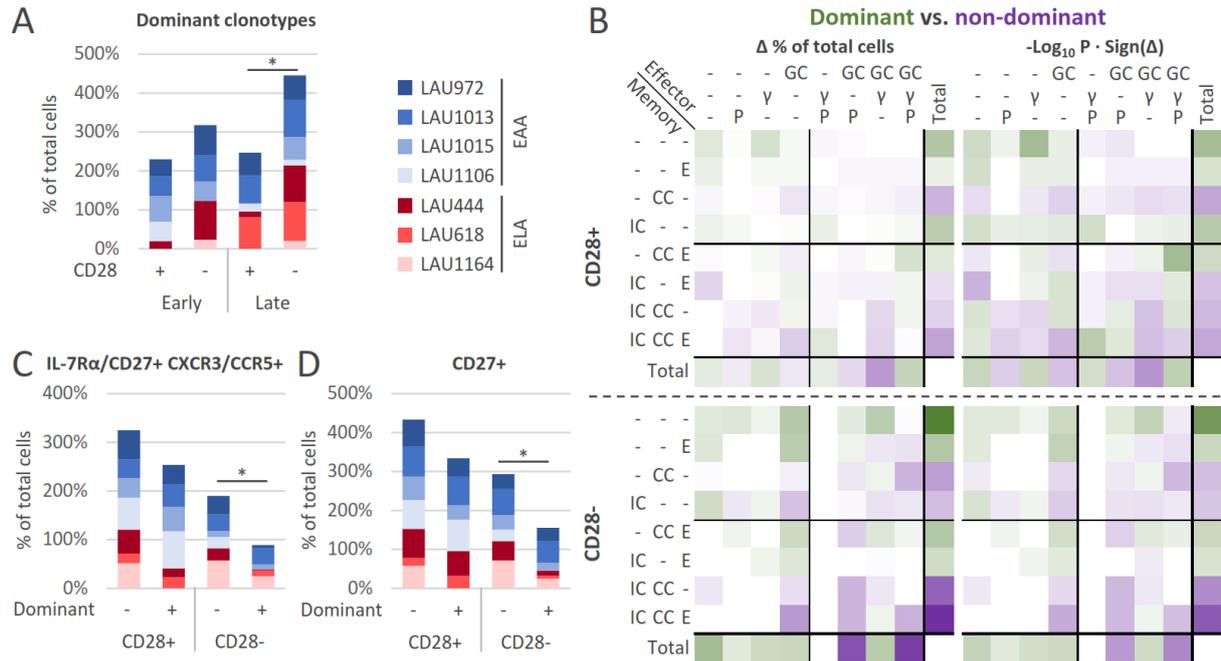
In our previous studies, only little if any phenotypic difference could be observed between dominant and non-dominant tumor-specific T cell clonotypes, both in terms of effector or memory gene expression or target-specific killing (48). These conclusions were based on single *ex vivo* tumor-specific CD8 T cells or on in vitro generated T cell clones derived from the late time point. We next asked whether our data confirmed these previous observations, and whether this was also the case at the early time point. Given the relative low clonotype identification rates, we first pooled the single cells from all patients and from both post-vaccination time points, only separating the EM28+ from the EM28- subpopulation and compared the memory and effector gene expression distribution within dominant vs. non-dominant clonotypes (Fig. 10B).

Patient	TRBV	ID number	CDR3 sequence	TRBJ	CD28+	CD28-	
LAU444	13	1	CAS-SELGTASYEQ-YFG	2.7	1	46	
		2	CAS-SSGQGNTGEL-FFG		2	1	
		3	CAS-SYGQNQPQ-HFG		3	0	
	14	1	CAS-SFGDNQPQ-HFG		5	0	
LAU618	3	1	CAS-SPPGLSGNIQ-YFG	2.4	12	0	
		2	CAS-SPGTLADTQ-YFG	2.3	6	1	
		4	CAS-SAGYGQPQ-HFG	1.5	0	4	
	17	1	CAS-SPGALNTEA-FFG	1.1	1	21	
		2	CAS-SIGPGLGQPQ-HFG	1.5	3	0	
LAU972	1	1	CAS-SVAHVDEQ-FFG	2.1	2	6	
		2	1	CSA-SETGVGQPQ-HFG	1.5	3	1
			2	CSA-SQGLTEA-FFG	1.1	3	1
			3	CSA-REGGILTDTQ-YFG	2.3	3	0
	5	1	CAS-SLGQGDQPQ-HFG	1.5	4	12	
	13	1	CAS-SETGGTEA-FFG	1.1	7	9	
		6	CAS-SELGTASYEQ-YFG	2.7	3	0	
		16	3	CAS-SQGGLGQPQ-HFG	1.5	3	0
LAU1013	7	3	CAS-SQVMVGAVDGY-TFG	1.2	3	1	
		5	CAS-ATAGQSNYGY-TFG		1	2	
	13	1	CAS-SRDSALWISTDTQ-YFG	2.3	8	21	
		3	CAS-SYSRDNEQ-FFG	2.1	4	3	
		4	CAS-SYGGLGQPQ-HFG	1.5	3	0	
LAU1015	8	1	CAS-SLGDVDE-YFG	2.7	8	6	
	13	2	CAS-SYLGMGQPQ-HFG	1.5	0	4	
LAU1106	3	1	CAS-SFQGLGQPQ-HFG	1.5	5	0	
		2	CAS-RAPGLANNEQ-FFG	2.1	7	0	
		5	CAS-SLGLAGNSEQ-FFG	2.1	4	0	
	7	2	CAS-SQGDWGGSQPQ-HFG		5	1	
LAU1164	5	1	XXS-SQGITGGPQ-HFG	1.5	0	4	

**Table 2.** CDR3 sequences of dominant clonotypes by patient and TRBV gene.

No significant difference could be observed in EM28+ cells (Fig. 10B, upper panels). Non-dominant EM28- T cells showed an enrichment in high memory gene expression patterns (IL-7R $\alpha$ /CD27+ CXCR3/CCR5+), contrasting to the dominant EM28- T cells that accumulated low memory gene expression bins (IL-7R $\alpha$ /CD27- CXCR3/CCR5-; Fig. 10B, lower panels). Small accumulation biases were

also observed among effector gene expression patterns, but were of little significance as revealed by the associated P values.



**Figure 10.** Comparison of effector and memory gene expression between late-differentiated (CD28-) EM cells with dominant vs. non-dominant clonotypes. **A.** The cumulative prevalence of dominant clonotypes of each patient was plotted as stack histograms across time in EM28+ and EM28- cells. The prevalence was calculated relatively to the total number of cells with identified clonotypes. Associated  $p$  values were calculated with two-tailed paired t tests: \*  $p < 0.05$ . **B.** Distribution of early- and late-differentiated EM cells with dominant (CD28+  $n = 107$ ; CD28-  $n = 130$ ) vs. non-dominant (CD28+  $n = 162$ ; CD28-  $n = 56$ ) clonotypes according to effector (columns) and memory (rows) gene expression patterns. Data of pre-vaccine time points and cells with undetermined clonotypes were excluded from the analysis. Left panels of each subset: distribution bias of cells from indicated compared groups relative to the total cell number of each group. Right panels of each subset:  $p$  values associated with the distribution bias of cells from indicated compared groups. GC: GZMB+ or CD94+;  $\gamma$ : IFN $\gamma$ +; P: PERF+; IC: IL-7R $\alpha$ + or CD27+; CC: CXCR3+ or CCR5+; E: EOMES+. **C-D.** The cumulative prevalence of indicated memory gene expression patterns of all patients was plotted as stack histograms according to EM cell population (CD28+/-) and clonotype (dominant/non-dominant). Associated  $p$  values were calculated with two-tailed paired t tests: \*  $p < 0.05$ .

Since the number of EM28- cells with identified clonotypes varied greatly from one patient to another, and, with it, their relative contribution to this pooled analysis, we also addressed whether these biases were observable at early and/or late time point by calculating the total rates of memory gene expression for each subpopulation separately (Fig. 10C). The prevalence of IL-7R $\alpha$ /CD27+ CXCR3/CCR5+ was significantly reduced in EM28- cells with dominant clonotypes relatively to their non-dominant counterparts ( $p < 0.05$ ; Fig. 10C). This difference was mainly due to the significant CD27 downregulation in dominant EM28- cells ( $p < 0.05$ ; Fig. 10D). These results are interesting but should be interpreted with caution given the low number of cells included in these analyses.

## Discussion

The first main aim of our study was to describe the effector and memory phenotype of Melan-A-specific CTLs at the single cell level over time and in both arms of the LUD00-0018 study. As previously described (51), EM28<sup>+</sup> cells of EAA patients showed a higher effector gene expression at the late time point compared to similar cells from ELA patients, confirming the quality of our new set of data (Fig. 5C). Moreover, our results revealed a phenotypic evolution of the Melan-A-specific CTLs upon vaccination and between early and late time points after vaccination. Notably, these changes consisted of an upregulation of all effector gene expression upon vaccination. Using a novel data analysis approach, we were also able to identify a memory gene expression pattern (IL-7R $\alpha$ <sup>+</sup> CXCR3<sup>-</sup> CCR5<sup>-</sup>), which was downregulated upon vaccination and became upregulated again at later time points.

The second aim of our study was to characterize the TCR repertoire of Melan-A-specific CTLs across time. Our analyses show that Melan-A vaccination stimulated the expansion of previously undetectable clonotypes, which, for the most part, subsequently show persistence over time. We also confirmed that the repertoire of EM28<sup>-</sup> cells was restricted to fewer more prevalent clonotypes compared to their EM28<sup>+</sup> counterpart. Focusing on the dominant clonotypes, we were able for the first time to identify differences in memory gene expression between these and non-dominant control clonotypes.

### Effector and memory phenotype evolution upon vaccination

By comparing Melan-A-specific CTLs isolated from pre-vaccine and early post-vaccine time points, we were able to observe two main differences. First, all tested effector genes were found to be upregulated upon vaccination and remained high at the late time point (Fig. 8A & C). This effect was observed in all analyzed patients (Fig. 8C). For three patients, the pre-vaccine prevalence of Melan-A-specific CTLs was low and all of these cells had to be sorted for analysis regardless of their phenotype (e.g. naïve, EM28<sup>+</sup>, EM28<sup>-</sup>). This bias could partially explain the difference in gene expression. However, the same difference was observed in the fourth patient, LAU444, from which both EM28<sup>+</sup> and EM28<sup>-</sup> cells could be isolated before vaccination as this patient had developed a spontaneous Melan-A-specific CTL response (Fig. 8A & C). Hence, we can conclude that our vaccination protocol is able to boost effector gene expression even in the case of a previously established Melan-A-specific response. Second, we were able to distinguish a memory gene expression pattern (IL-7R $\alpha$ <sup>+</sup> CXCR3<sup>-</sup> CCR5<sup>-</sup>), which was concomitantly downregulated at early time point post vaccination. This analysis may suffer from the same bias as mentioned above for the effector gene regulation. Yet again, LAU444 showed a similar regulation pattern of these memory gene transcripts (although to a lesser degree for CXCR3), suggesting that they were indeed influenced by the vaccine and not only by a spontaneously generated/tumor-induced stimulation and differentiation of the tumor-specific CD8 T cells.

Interestingly, when we compared the EM28<sup>+</sup> cells from the late and early time points, the same memory, IL-7R $\alpha$ <sup>+</sup> CXCR3<sup>-</sup> CCR5<sup>-</sup>, gene expression pattern was upregulated again over time after vaccination, (Fig. 7A-B). This difference was shared by all patients except one and constituted one of the most obvious changes that we could detect for either EM28<sup>+</sup> or EM28<sup>-</sup> cells between the two time points. Among the three genes involved, IL-7R $\alpha$  was the only one showing significant regulation on its own. These data are consistent with IL-7R $\alpha$  protein expression assays performed by FACS in our lab, which showed an initial decrease in IL-7R $\alpha$  expression in all patients upon vaccination and a slighter re-upregulation over time, especially in EAA patients (Gannon et al. Manuscript in preparation).

The significance of the regulation of these memory genes is not clear. The genes that we monitored, which are commonly labeled as “memory-associated” genes, actually have heterogeneous functions. IL-7R $\alpha$  is a specific subunit of the receptor sensing IL-7, which functions to enhance CTL survival and proliferation. The fact that it became upregulated again upon continuous vaccination is interesting as it could be the marker of a longer lasting memory response to the target antigen.

Overall the data regarding EM cell differentiations showed that our vaccination protocol not only induced higher frequencies of Melan-A-specific CTLs but also had an effect on their phenotype and might thereby increase their effector function and their survival.

## Repertoire analysis

Our repertoire analysis yielded interesting results. We were able to observe the persistence of a majority of clonotypes through the early and late time points (Fig. 9). These results showed that vaccination was able to induce early on a Melan-A-specific repertoire that could remarkably persist over time. On the other hand, short of half the clonotypes observed at one time point were not detected at the other, thereby revealing that some repertoire plasticity was still possible after three months of repeated vaccine doses. The highest degree of repertoire stability was actually observed in the spontaneous Melan-A-specific response of LAU444, where the dominance of the BV13-1 clonotype over the EM28- cell repertoire remained virtually intact from nearly one year before the start of vaccination to more than a year and a half after (Fig. 9). This observation suggested that the repertoire became physiologically more difficult to alter after it had reached a steady state for a long time.

The relative stability of the repertoire after vaccination strongly contrasted with the fact that no common clonotype could be found between pre-vaccine Tet+ and early time point EM28+ cells in patients other than LAU444 (Fig. 9). This is interesting as it suggested that the vaccination stimulated relatively rare clones among the Melan-A-specific CTL population at the beginning.

Using the advantage of the single cell resolution of our analysis to its fullest, we were able to distinguish a downregulation in CD27 expression between EM cells with dominant clonotypes. This difference was mainly observed, and was only significant, in EM28- cells, which displayed a higher prevalence of dominant clonotypes and a lesser level of CD27 transcript (Fig. 10A & D). The significance of this differential expression was not clear. CD27 is a marker of lesser differentiation and its downregulation could be linked to the enhanced proliferation necessary for a CTL clone to achieve dominance.

## Data quality, strengths and limitations

Our data nicely reproduced previously observed differences, such as the lower memory gene expression and repertoire restriction to fewer and more prevalent clonotypes in EM28- cells compared to their EM28+ counterparts (Fig. 5B & 6C), as well as the higher effector gene expression of EM28+ cells from EAA patients compared to ELA-derived controls. These observations were as many indicators of the quality of our data set.

The main strength of the present study lies in its very high resolution and dimensionality. These characteristics allowed us to identify specific gene expression patterns that were regulated over time (e.g. IL-7R $\alpha$ + CXCR3+ CCR5+; Fig. 7A-E) or in specific subpopulations. A nice example of the power of this approach is shown in Fig. 10C-D: simultaneously determining the clonotype of each single cell as well as its gene expression phenotype enabled us to cross these data and allowed the identification of gene expression differences among cells with dominant or non-dominant clonotypes.

Nonetheless, there remains some technical limitation to this study. The cDNA amplification method that we used is a multistep process that must be performed on each single cell separately. Each target gene must also be amplified separately, which again increased the number of tubes / plates to handle. These disadvantages hindered the throughput, and thereby limited the total number of cells and the number of patients that might be analyzed. Also, as the generated gene expression data of binary nature, a relatively high number of cells were required in order to get meaningful data from this stochastic output. These limitations were especially sensible in analyses focusing on smaller subsets of cells. For example, the number of cells with dominant clonotypes (Fig. 10B) precluded a more detailed analysis of their phenotype over time as splitting the cells between the two time points would have reduced their number too much. Another example is the number of cells with identified TCR clonotype at the pre-vaccine time points. Our study was not designed to detect rare clonotypes to begin with and

a much higher number of cells would presumably be necessary to detect common clonotypes between the pre-vaccine cells and the later time points.

Technical improvements are therefore needed to enhance the throughput, quality and sensitivity of such large scale analyses. If we were to imagine an ideal technique for the characterization of the CTL immune response at the single cell, the capacity to process massive numbers of samples in parallel and the simultaneous quantitative measurement of gene expression and TCR gene sequencing would probably be its foremost features.

Recently, various methods have been published, describing single cell cDNA high throughput sequencing (48,56–58). These RNA-Seq methods virtually enable to read the whole transcriptome of any given cell. In addition, high-throughput sequencing techniques may be adapted to analyze a large number of samples in parallel. This is made possible by genetic barcoding of sequencing libraries, which allows many samples to be pooled and sequenced together. Sequences may be deconvoluted at the data analysis stage and reattributed to each sample by the simultaneous sequencing of the barcodes. The first single cell RNA-Seq approach to have been described is very similar to our cDNA amplification protocol. It is based on the combination of poly-dA tailing of the cDNA followed by PCR amplification using oligo-dT containing primers (57). However, this method precludes the introduction of barcodes in the first steps of the library synthesis and would therefore require a separate sample processing. Two other methods were published more recently and are compatible with genetic barcoding at a step as early as reverse transcription (56,58).

Target enrichment is a method to reduce the diversity of a library to focus on sequences of interest. A large scale example is exome sequencing which is most often performed by in solution hybridization (59). Lower throughput PCR-based techniques have also been described (59). Targeted sequencing enables a further increase in number of samples that may be processed in parallel while retaining a high number of reads per target sequence. Such a strategy would be required to process relatively high number of samples (e.g. > 1'000 single cells as in the present study). It has indeed been applied recently to single CD8+ T cells and involved target-specific reverse transcription and multiplex nested PCRs to amplify the target cDNAs (60). This method was successfully applied to assay for the expression of more than a hundred genes in parallel including the TCR $\alpha/\beta$  CDR3 sequences. The main drawback of this technique is that the gene targeting was not dissociable from library synthesis, which precludes the analysis of other genes expression from the same starting material. Also, sample barcoding could only be performed in the final steps of library generation, implying that earlier amplification steps have to be done in separate reactions for each cell, thus reducing the throughput and the reproducibility of the data. To our knowledge, no ideal technique has been described yet, but combinations of allowing sample barcoding at the level of reverse transcription and orthogonal target selection approaches seem like an interesting potential answer to the single cell analytical challenge.

## Acknowledgements

I would like to thank Philippe Gannon & Marina Iancu for generating the amplified single cell cDNA samples that were the base of this study. Philippe Gannon also spent countless time at explaining to me the techniques and, more importantly, the immunological background of the project, time for which I am most grateful. Last but not least, I would like to thank Dr. Nathalie Rufer for her supervision and great support for this nice master project.

## References

1. Le cancer en Suisse. Neuchâtel; 2011.
2. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz L a, Kinzler KW. Cancer genome landscapes. *Science*. 2013; 339(6127): 1546–58.
3. Buzaid AC, Gershenwald JE. Tumor node metastasis (TNM) staging system and other prognostic factors in cutaneous melanoma. *UpToDate*; 14. p. 1–18.
4. Swetter S, Geller AC. Skin examination and clinical features of melanoma. *UpToDate*. *UpToDate*; 14. p. 1–54.
5. Balch CM, Buzaid AC, Soong SJ, Atkins MB, Cascinelli N, Coit DG, et al. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 01. p. 3635–48.
6. Balch CM, Gershenwald JE, Soong S-J, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol*. 2009; 27(36): 6199–206.
7. Sosman JA. Overview of the management of advanced cutaneous melanoma. *UpToDate*. 14. p. 1–10.
8. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med*. 2011; 364(26): 2507–16.
9. Flaherty KT, Robert C, Hersey P, Nathan P, Garbe C, Milhem M, et al. Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma. *N Engl J Med*. 2012; 367(2): 107–14.
10. Robert C, Karaszewska B, Schachter J, Rutowski P, Mackiewicz A, Stroiakovski D, et al. Improved Overall Survival in Melanoma with Combined Dabrafenib and Trametinib. *N Engl J Med*. 2015; 372(1): 30–9.
11. Larkin J, Ascierto P a., Dréno B, Atkinson V, Liskay G, Maio M, et al. Combined Vemurafenib and Cobimetinib in BRAF-Mutated Melanoma. *N Engl J Med*. 2014; 371(20): 1867–76.
12. Sosman JA, Atkins MB. Adjuvant immunotherapy for melanoma. *UpToDate*. 14. p. 1–17.
13. Ross ME. Immunotherapy of advanced melanoma with immune checkpoint inhibition. *UpToDate*. 2014;
14. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol*. 2002; 3(11): 991–8.
15. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 11. p. 646–74.
16. Smith-Garvin JE, Koretzky G a, Jordan MS. T cell activation. *Annu Rev Immunol*. 2009; 27: 591–619.
17. Starr TK, Jameson SC, Hogquist K a. Positive and negative selection of T cells. *Annu Rev Immunol*. 2003; 21: 139–76.
18. Behrens G, Li M, Smith CM, Belz GT, Mintern J, Carbone FR, et al. Helper T cells, dendritic cells and CTL Immunity. *Immunol Cell Biol*. 2004; 82(1): 84–90.
19. Kurts C, Robinson BWS, Knolle P a. Cross-priming in health and disease. *Nat Rev Immunol*. Nature Publishing Group; 2010; 10(6): 403–14.
20. Williams M a, Bevan MJ. Effector and memory CTL differentiation. *Annu Rev Immunol*. 2007; 25: 171–92.
21. Mahnke YD, Devèvre E, Baumgaertner P, Matter M, Rufer N, Romero P, et al. Human melanoma-specific CD8+ T-cells from metastases are capable of antigen-specific degranulation and cytotoxicity directly ex vivo. *Oncoimmunology*. 2012; 1(4): 467–76.
22. Gattinoni L, Klebanoff C a., Restifo NP. Paths to stemness: building the ultimate antitumour T cell. *Nat Rev Cancer*. Nature Publishing Group; 2012; 12(10): 671–84.

23. Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annu Rev Immunol.* 2004; 22: 745–63.
24. Romero P, Zippelius a., Kurth I, Pittet MJ, Touvrey C, Iancu EM, et al. Four Functionally Distinct Populations of Human Effector-Memory CD8+ T Lymphocytes. *J Immunol.* 2007; 178: 4112–9.
25. Chen DS, Mellman I. Oncology Meets Immunology : The Cancer-Immunity Cycle. *Immunity.* 2013;(step 3): 1–10.
26. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science (80- ).* 2015; 348(6230): 69–74.
27. Motz GT, Coukos G. Deciphering and Reversing Tumor Immune Suppression. *Immunity.* Elsevier Inc.; 2013; 39(1): 61–73.
28. Zou W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat Rev Cancer.* 2005; 5(4): 263–74.
29. Muller AJ, Scherle P a. Targeting the mechanisms of tumoral immune tolerance with small-molecule inhibitors. *Nat Rev Cancer.* 2006; 6(8): 613–25.
30. Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoediting. *Nat Rev Immunol.* 2006; 6(11): 836–48.
31. Swann JB, Smyth MJ. Immune surveillance of tumors. *J Clin Invest.* 2007; 117(5): 1137–46.
32. Vajdic CM, Van Leeuwen MT. Cancer incidence and risk factors after solid organ transplantation. *Int J Cancer.* 2009; 125(8): 1747–54.
33. Ross ME. Interleukin-2 and other immunotherapies for advanced melanoma. *UpToDate.* 14. p. 1–14.
34. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer.* Nature Publishing Group; 2012; 12(4): 252–64.
35. Hodi FS, O’Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med.* 2010; 363(8): 711–23.
36. Robert C, Ascierto PA, Maio M, Hernberg M, Schmidt H, Waxman I, et al. A phase III, randomized, double-blind study of nivolumab (anti-PD-1; BMS-936558; ONO-4538) versus dacarbazine in patients (pts) with previously untreated, unresectable, or metastatic melanoma (MEL). *ASCO Meet Abstr.* 2013; 31(15\_suppl): TPS9106 – .
37. Chmielowski B, Hamid O, Minor DR, D’Angelo SP, Pennock GK, Grossmann K, et al. A phase III open-label study of nivolumab (anti-PD-1; BMS-936558; ONO-4538) versus investigator’s choice in advanced melanoma patients progressing post anti-CTLA-4 therapy. *ASCO Meet Abstr.* 2013; 31(15\_suppl): TPS9105 – .
38. Wolchok JD, Kluger H, Callahan MK, Postow M a, Rizvi N a, Lesokhin AM, et al. Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med.* 2013; 369(2): 122–33.
39. Lizée G, Overwijk WW, Radvanyi L, Gao J, Sharma P, Hwu P. Harnessing the Power of the Immune System to Target Cancer. *Annu Rev Med.* 2013; 64(1): 71–90.
40. Rosenberg S a, Yang JC, Sherry RM, Kammula US, Marybeth S, Phan GQ, et al. Durable Complete Responses in Heavily Pretreated Patients with Metastatic Melanoma Using T Cell Transfer Immunotherapy. *Clin Cancer Res.* 2011; 17(13): 4550–7.
41. Rosenberg S a, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. *Nat Med.* 2004; 10(9): 909–15.
42. Yamada A, Sasada T, Noguchi M, Itoh K. Next-generation peptide vaccines for advanced cancer. *Cancer Sci.* 2013; 104(1): 15–21.
43. Schwartzentruber DJ, Lawson DH, Richards JM, Conry RM, Miller DM, Treisman J, et al. gp100 peptide vaccine and Interleukin-2 in patients with advanced melanoma. *N Engl J Med.* 2011; 364: 2119–27.

44. Wang E, Phan GQ, Marincola FM. T-cell-directed cancer vaccines: the melanoma model. *Expert Opin Biol Ther.* 2001; 1: 277–90.
45. Speiser DE, Baumgaertner P, Voelter V, Devereux E, Barbey C, Rufer N, et al. Unmodified self antigen triggers human CD8 T cells with stronger tumor reactivity than altered antigen. *Proc Natl Acad Sci U S A.* 2008; 105(10): 3849–54.
46. Valmori D, Fonteneau J, Marañón C, Gervois N, Liénard D, Rimoldi D, et al. Enhanced Generation of Specific Tumor-Reactive CTL In Vitro by Selected Melan-A/MART-1 Immunodominant Peptide Analogues. *J Immunol.* 2014;
47. Ayyoub M, Zippelius A, Pittet MJ, Rimoldi D, Valmori D, Cerottini JC, et al. Activation of human melanoma reactive CD8+ T cells by vaccination with an immunogenic peptide analog derived from Melan-A/melanoma antigen recognized by T cells-1. *Clin Cancer Res.* 2003; 9: 669–77.
48. Speiser DE, Wieckowski S, Gupta B, Iancu EM, Baumgaertner P, Baitsch L, et al. Single cell analysis reveals similar functional competence of dominant and nondominant CD8 T-cell clonotypes. *Proc Natl Acad Sci U S A.* 2011; 108(37): 15318–23.
49. Wieckowski S, Baumgaertner P, Corthesy P, Voelter V, Romero P, Speiser DE, et al. Fine structural variations of alpha-beta TCRs selected by vaccination with natural versus altered self-antigen in melanoma patients. *Journal of immunology.* 2009.
50. Baumgaertner P, Jandus C, Rivals JP, Derr?? L, L??vgren T, Baitsch L, et al. Vaccination-induced functional competence of circulating human tumor-specific CD8 T-cells. *Int J Cancer.* 2012; 130(11): 2607–17.
51. Gupta B, Iancu EM, Gannon PO, Wieckowski S, Baitsch L, Speiser DE, et al. Simultaneous coexpression of memory-related and effector-related genes by individual human CD8 T cells depends on antigen specificity and differentiation. *J Immunother.* 2012; 35(6): 488–501.
52. Arden B, Clark S, Kabelitz D, Mak T. Human T-cell receptor variable gene segment families. *Immunogenetics.* Springer-Verlag; 1995; 42(6): 455–500.
53. Speiser DE, Baumgaertner P, Barbey C, Rubio-Godoy V, Moulin A, Corthesy P, et al. A novel approach to characterize clonality and differentiation of human melanoma-specific T cell responses: spontaneous priming and efficient boosting by vaccination. *J Immunol.* 2006; 177(2): 1338–48.
54. Iancu EM, Corthesy P, Baumgaertner P, Devereux E, Voelter V, Romero P, et al. Clonotype selection and composition of human CD8 T cells specific for persistent herpes viruses varies with differentiation but is stable over time. *J Immunol.* 2009; 183(1): 319–31.
55. Iancu EM, Gannon PO, Laurent J, Gupta B, Romero P, Michielin O, et al. Persistence of EBV Antigen-Specific CD8 T Cell Clonotypes during Homeostatic Immune Reconstitution in Cancer Patients. *PLoS One.* 2013; 8(10): 1–14.
56. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 2011; 21(7): 1160–7.
57. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009; 6(5): 377–82.
58. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep. The Authors;* 2012; 2(3): 666–73.
59. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, et al. Target-enrichment strategies for next-generation sequencing. *Nat Methods.* 2010; 7(2): 111–8.
60. Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol.* 2014;(June).