Digital PCR-Based T-cell Quantification-Assisted Deconvolution of the Microenvironment Reveals that Activated Macrophages Drive Tumor Inflammation in Uveal Melanoma

Lange, M.J. de; Nell, R.J.; Lalai, R.N.; Versluis, M.; Jordanova, E.S.; Luyten, G.P.M.; ... ; Velden, P.A. van der

Digital PCR-based T Cell Quantification Assisted Deconvolution of the Microenvironment Reveals that Activated Macrophages Drive Tumor Inflammation in Uveal Melanoma

Running title: The Immune Environment in Uveal Melanoma

Mark J de Lange[1*], Rogier J Nell[1*], Rajshri N Lalai[1], Mieke Versluis[1], Ekaterina S. Jordanova[2,3], Gre PM Luyten[1], Martine J Jager[1], Sjoerd H van der Burg[4], Willem H Zoutman[5], Thorbald van Hall[4], Pieter A van der Velden[1]

1. Department of Ophthalmology, LUMC, Leiden, the Netherlands.

2. Department of Pathology, LUMC, Leiden, the Netherlands

3. Center for Gynecologic Oncology Amsterdam, VUmc, The Netherlands

4. Department of Medical Oncology, LUMC, Leiden, the Netherlands.

5. Department of Dermatology, LUMC, Leiden, the Netherlands.

* contributed equally to the manuscript

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest were disclosed.

Department of Ophthalmology, LUMC

P.O. Box 9600, 2300 RC Leiden

The Netherlands

Tel. 31-71-5263938

Fax 31-71-5248222

Corresponding author: Dr. Pieter A. van der Velden, PhD. E-mail: velden@lumc.nl

ABSTRACT

Uveal melanoma (UM) progression can be predicted by gene expression profiles enabling a clear subdivision between tumors with a good (class I) and a poor (class II) prognosis. Poor prognosis UM can be subdivided by expression of immune-related genes, however it is unclear if this subclassification is justified; therefore, T cells in UM specimens were quantified using a digital PCR approach. Absolute T cell quantification revealed that T cell influx is present in all UM associated with a poor prognosis. However, this infiltrate is only accompanied by differential immune-related gene expression profiles in UM with the highest T cell infiltrate. Molecular deconvolution of the immune profile revealed that a large proportion of the T cell-related gene expression signature does not originate from lymphocytes but is derived from other immune cells, especially macrophages. Expression of the lymphocyte homing chemokine CXCL10 by activated macrophages correlated with T cell infiltration and thereby explains the correlation of T cell numbers and macrophages. This was validated by in situ analysis of CXCL10 in UM tissue with high T cell counts. Surprisingly, CXCL10 or any of the other genes in the activated macrophage-cluster, was correlated with reduced survival due to UM metastasis. This effect was independent of the T cell infiltrate which reveals a role for activated macrophages in metastasis formation independent of their role in tumor inflammation.

INTRODUCTION

Uveal Melanoma (UM) is the most common intraocular neoplasm in adults with an incidence of 6 to 8 per million annually in Caucasians (1). UM presents as a genetically homogenous disease in the sense that the vast majority shares the same driver mutations (GNAQ/11) (2,3). Downstream oncogenic signaling pathways include the ERK pathway via PKC and the Hippo pathway via YAP activation (4,5). For years, it has been known that UM that metastasize significantly differ both in their genetic and their phenotypic make up, from the ones that do not metastasize. Early studies showed recurrent genomic abnormalities that allowed clustering of tumors into prognostic classes (6-8). Monosomy of chromosome 3 and gain of chromosome 8q discriminate between good and poor prognosis (9-11). More recently, an advanced approach of classifying UM was revealed. Based on genome wide gene expression analysis, tumors were assigned to prognostic classes (class I and classII) that overlap largely with the genomic classification (12). The former classification is based on hundreds of differentially expressed genes that may also provide insight into the biology of UM. Recently we and others demonstrated that a part of the UM with a poor prognosis are characterized by an extensive immune infiltrate (6,13). Besides T cell markers also macrophage markers were recognized in the expression profiles of UM with metastatic potential. Macrophage activation has been shown to precede T cell infiltration in UM progression (14). To further investigate the mechanisms of UM inflammation, the expression profile was analyzed for the inflammatory compartment. For this purpose absolute T cell counts were integrated with UM expression profiles. The resulting T cell-related genes were compared to a range of immune cells to identify immune cells in the tumour microenvironment. By using profiles of both naïve and activated immune cells we

could infer that activated macrophages are pivotal in T cell infiltration. Combined we show that by using absolute T cell counts, expression profiles of heterogeneous tissue can be effectively dissected into different immune components.

MATERIAL & METHODS

Tumor material was obtained from 64 enucleated eyes of UM patients that had been enucleated at the Leiden University Medical Center, Leiden, The Netherlands, between 1999 and 2008. This study was approved by the Medical Ethics Committee of the Leiden University Medical Center. Tumor material was handled according to the Dutch National Ethical Guidelines ('Code for Proper Secondary Use of Human Tissue'), and the tenets of the Declaration of Helsinki (World Medical Association of Declaration 2013; ethical principles for medical research involving human subjects). None of the tumors had prior treatment and only tumors with a follow-up time of at least 5 years were used. The maximum follow-up was 14 years. The average age at enucleation was 60.6 years (range 13 to 88); 33 patients were male and 31 female. Tumor material was snap frozen using 2-methyl butane, and RNA and DNA was isolated using the RNeasy mini kit and QIAmp DNA minikit, respectively, (both Qiagen, Valencia, USA) from 20 sections of 20µm according to the manufacturer's guidelines.

Gene expression

Gene expression analysis was performed as published before (6). In short, the Illumina HumanHT-12 v4 chip containing 47,000 probes across the whole genome was used. Supervised cluster analysis was used to identify which genes were responsible for the subdivision of the tumors in classes. For differences between subgroups, i.e. I versus II, a correction was made for differences between IIa and IIb classified as a log fold change smaller than -0.5 or greater than 0.5 and a p-value smaller than 0.05. Since gene expression data have been obtained in two batches, a batch effect correction was applied. The R packages used were: 'ber' for batch correction and 'lumi' for unsupervised clustering.

Genetic T cell quantification

In order to quantify T cells in tumor samples, a dPCR assay was developed directed at a specific locus of the TCR-β gene; located between gene segments Dβ1 and Jβ1.1, and from now on called ΔB. This particular part of the gene is biallelically deleted by T cell receptor rearrangements during T cell maturation. Consequently, peripheral T cells are lacking ΔB compared to other cell types (somatic loss of germline DNA). This genetic dissimilarity can be utilized in a basic copy number variation (CNV) dPCR assay in order to quantify T cells in the presence of other cell types, like UM (tumor) cells. Because a stable genomic reference is essential in this determination, $\frac{[\Delta B]}{[\text{REFERENCE}]}$ was calculated for 3 different reference genes: TTC5 (chr. 14), TERT (chr. 5) and VOPP1 (chr. 7). The average of the two closest ratios was used in the following formula:

$$\text{T-cell fraction} = 1 - \text{average}\ \frac{[\Delta B]}{[\text{REFERENCE}]}$$

Although the target gene segment ΔB is located at a locus not frequently mutated in UM cells, copy number alterations in tumor cells may give rise to a distortion of our calculations. It is possible to correct for this by using the following formula:

$$\text{adjusted T-cell fraction} = \text{average}\ \frac{[\text{VOPP1}]}{[\text{REFERENCE}]} - \text{average}\ \frac{[\Delta B]}{[\text{REFERENCE}]}$$

In those cases where chromosome 7 shows an obvious loss or gain (defined by a concordant copy number alteration > 0.075 seen in $\frac{[\text{VOPP1}]}{[\text{TTC5}]}$ and $\frac{[\text{VOPP1}]}{[\text{TERT}]}$), we chose to determine T cell fractions according to this formula. In those calculations, VOPP1

was not used as reference gene. Calculations per tumor are outlined in Supplementary Table 2 (Table S2).

The ddPCR was performed using ddPCR Supermix for probes (Bio-Rad Laboratories, Hercules, CA, USA) in 20 µL with 50ng of DNA resulting in 0.75 copies per droplet (CPD) of haploid genomes after partitioning into 20,000 droplets.

DNA restriction digestion was performed using HaeIII directly in the ddPCR reaction solution according to the protocol supplied by Bio-Rad. Droplets were generated using an AutoDG System (Bio-Rad) and droplet emulsion was transferred to a 96-well PCR plate for amplification in a T100 Thermal Cycler (Bio-Rad). Cycle parameters were as follows: Enzyme activation for 10 minutes at 95°C; denaturation for 30 seconds at 94°C, annealing and extension for 1 minute at 60°C for 40 cycles; enzyme deactivation for 10 minutes at 98°C; infinite cooling at 12°C. Ramp rate for all cycles was 2°C/sec. Cycled droplets were stored at 4-12°C until reading. Positive and negative droplets were measured as a CNV1 experiment using a QX200 Droplet Reader (Bio-Rad). Primers and probes are proprietary of Bio-Rad except for the primers and probes for TRB, which have been published before (15). In Supplementary Table 3 the amplicon sequences are provided (Table S3).

BIOGPS METHOD

Obtained gene expression profiles from UM samples represent a mixture of cell types, i.e. melanoma cells and stromal cells. We developed a computational approach to dissect which cell types contribute to the expression signatures.
At the basis of our *in silico* analysis lies the selection of genes of interest for which the expression level is positively correlated with increased T cell fraction in the class

II UM samples (n=38). Pearson's correlation test was used and $r > 0.5$ and $p < 0.001$ were considered to be significant.

The publicly available datasets GeneAtlas U133A, gcrma (16) and Primary Cell Atlas (17) on the BioGPS-site were used to obtain cell-type specific gene expression patterns for our selected genes (18-20). Hierarchical cluster analysis and principal component analysis, revealed cell specific expression patterns in our gene selection. The following R packages were used: 'mygene' for obtaining gene information and 'pheatmap', 'dendsort' and 'ggplot2' for visualizing data.

Immunofluorescent staining

Phenotypic characterization of lymphocytes was performed using triple fluorescent immunostaining. A previously developed technique for simultaneous immunofluorescence (IF) staining of different epitopes was applied to 4-µm formalin-fixed, paraffin-embedded tissue sections. In brief, deparaffinized and EDTA antigen retrieval-treated sections were stained by a mixture of the following primary antibodies: anti-CD8 (mouse monoclonal IgG1; Dako-Agilent, Santa Clara, USA), melan A (mouse monoclonal, Novus Biologicals, LLC, Littleton, USA), CXCL10 (rabbit polyclonal, Antibodies-online, Aachen, Germany), CD14 (mouse monoclonal IgG2a, Abcam, Cambridge, UK), CD163 (mouse monoclonal IgG1, Novocatra, Milton Keynes, UK). As secondary antibodies to visualize the lymphocytes, a combination of fluorescent antibody conjugates goat anti-rabbit IgG-Alexa Fluor 546, goat anti-mouse IgG2b-Alexa Fluor 647, goat anti-mouse IgG1-Alexa Fluor 488 (all three from Molecular Probes, Invitrogen, Breda, the Netherlands), and goat anti-rabbit-Alexa Fluor 647, goat anti-mouse IgG2a-Alexa Fluor 546, and goat anti-mouse IgG1-Alexa Fluor 488 (all three from Life Technologies, Grand Island, USA) was used. Antibodies were used as described previously (21). Images were captured with a

confocal laser scanning microscope (LSM510; Carl Zeiss Meditec, Jena, Germany) in a multitrack setting. A microscope objective (PH2 Plan-NEOFluar 25x/0.80 Imm Korr; Zeiss) was used. T cells were manually counted using the LSM 5 Image Examiner software and represented as the number of cells per mm^2 for each slide (average of five 250× images).

Statistical analysis

For gene expression, deconvolution and statistical analyses, the programming language R was used (R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.). Detailed methods for analysis are provided as supplemental methods.

RESULTS

UM subclassification reveals an immune phenotype

With gene expression profiles, UM can be easily subdivided into different prognostic classes (12,22). Class I mainly consists of tumors with a good prognosis while class II represent tumors with a poor prognosis (Figure 1A). Based on the most differentially expressed genes, we recognized two sub-classes (IIa/IIb) in class II UM (6). These subclasses however do not present with a survival difference (Figure 1A). High expression of immune related genes in class IIb UM suggests that these tumours are inflamed, while class IIa UM tumours do not include this phenotype (Figure 1B).

Though gene expression profiles are helpful in exploring the immune infiltrate, they do not provide an immune cell count. The underlying reason may be that cell specific markers such as CD4 and CD8 are regulated during immune reactions. This makes CD4 and CD8 expression useful to describe the T cell populations, rather than for T cell quantification. CD3 expression is marginally regulated during immune responses compared to CD4 and CD8 and therefore more appropriate to enumerate the T cell infiltrate. In order to define the extent of the immune cell infiltration in UM more accurately, we quantified T cells in UM DNA specimens from which the gene expression profiles were also obtained.

Digital PCR reveals T cell infiltration in the whole of class II UM

Based on somatic rearrangements of the T cell receptor genes, T cells can be distinguished from other cells at the genomic and RNA level. The number of possible rearrangements is however innumerable and amplification of every possible T cell receptor requires many assays (23). The complexity of the TCR genes therefore

hampers accurate measurement and analysis at genomic level and the gene expression level. Instead of counting each individual rearrangement, we used an alternative amplification method for quantification that depends on a DNA sequence (ΔB). This sequence is deleted in both alleles of lymphocytes during early T cell maturation and therefore the marker of choice for counting T cells (15,24).

With dPCR, T cell numbers were quantified in the tumor mass of 64 UM that were previously analyzed with gene expression arrays. Significantly higher T cell fractions were observed in class IIb tumors compared to class IIa and class I tumors (Figure 2A). However, elevated T cell counts can also be found in class IIa, compared to class I UM. This is in contrast to gene expression analysis of CD3 which did not reveal significant differences between class I and class IIa (Figure 2B). The lack of differential expression of CD3 between class I and class IIa likely marks the reduced sensitivity and specificity of gene expression arrays compared to the DNA based T cell count. Significant correlations between genomic T cell counts and gene expression of CD3 were nevertheless observed (Figure 2C). To assess the degree to which T cells contribute to the gene expression profile we systematically correlated T cell number and gene expression profiles of classifier genes.

Integration of T cell count with the expression profile of UM reveals structure of the immune environment

Instead of investigating the expression profile for immune cell markers, we investigated the degree to which T cells contribute directly or indirectly to the expression profile of class II UM. Therefore we reversed the analysis and performed supervised cluster analysis on the basis of the T cell count. The correlation with T cell count was determined for 1538 (logfold change >0.5) probe sets, which are

differentially expressed between class I and class II. This revealed that 60 genes of the class II classifier are positively correlated with T cell count ($R>0.5$, $p<0.05$) (Figure 3A, Table S1). Among these T cell classifier genes are obligate T cell markers such as CD3 and CD8. Also lymphocyte attracting chemokines like CCL5 are expressed by T cells and found to be present in the T cell classifier (25). Comparison of the expression profiles with 35 expression profiles of naive and activated immune cells indicated that the majority of T cell-correlated genes is actually not expressed by T cells (Figure 3B). The myeloid lineage of immune cells was found to contribute considerably to the T cell classifier. This is illustrated by the correlation between T cell count and macrophage markers such as CD14 ($R=0.484$, $P=0.002$), CD86 ($R=0.70$, $p<0.05$) and CSF1R (R=0,606, $P<0.001$). However, the variability of the correlation between macrophage markers and T cell count, or even absence of a correlation (CD68, $R=0.196$, $p=0.239$), indicates that macrophage polarization is highly dynamic. Moreover, the variable correlation of macrophage marker expression with T cell count indicates that specific subtypes of macrophages are present in inflamed UM. With principal component analysis (PCA), the T cell classifier converged into 4 clusters that represent different cell populations (Figure 3C). The most distinct gene-cluster contained expression profiles of macrophages that had been activated with classic immune activators like LPS and interferon gamma. Another prominent cluster corresponded to the profile of activated CD8-positive T cells, as can be witnessed by high granzyme expression. The lymphocyte attracting chemokine CXCL10 was highly expressed in the activated macrophage gene cluster, and may be functionally related to T cell infiltration in UM (Figure 3B-C). Expression of CXCL10 by macrophages was investigated in UM tissue with two triple staining procedures. Either T cell or macrophage markers were analyzed

alongside a melanocyte marker and CXCL10 expression. This confirmed high CXCL10 expression by macrophages in UM with high T cell counts. While staining with macrophage markers (CD14, CD163) revealed that macrophages are the origin of CXCL10 expression. Though UM cells also occasionally displayed CXCL10 expression, strong staining was uniquely observed in macrophages that express CD163 (Table 1) (Figure 4A-B). Combined, this displays an active immune response in part of the UM and the question emerges whether this results in a clinical response.

Clinical consequence of the immunologic phenotype

Previously we reported that the class IIa and class IIb UM presented with similar prognosis (Figure 1A). Based on that, we claimed that presence or absence of an immune infiltrate did not influence disease outcome in UM (6). With two immune cell populations in UM defined, we evaluated the role of T cell count and activated macrophages in the development of UM metastasis separately. We analyzed this in a molecular UM risk model, to which we added T cell count as well as expression markers for T cell phenotype (CD4, CD8) and CXCL10 expression as marker for activated macrophages. In this model, monosomy 3 and chromosome 8q gain were significantly correlated with survival (Table 2). T cell count did not contribute to survival (Figure 5), and neither did the expression of the T cell markers (CD4, CD8). CXCL10 as marker for activated macrophages did surprisingly contribute to the survival risk of UM in this complex model. Thereby, it was shown to represent an independent risk factor that is not confounded by monosomy 3, gain of 8q or T cell count.

DISCUSSION

Molecularly, UM can be easily divided into different prognostic classes (class I and class II) based on their genome wide gene expression (12,22). Recently, our gene expression analysis on 64 UM revealed an additional subdivision. With supervised cluster analysis of the classifier genes, class II UM was subdivided intoclass IIa and class IIb. Genes that were differentially expressed between these classes, and thus responsible for this subdivision, were functionally annotated to be related to the immune system. Expression of interferon signature genes like CD2, CD3D, CCL5 and CXCL10 reflect an ongoing tumor inflammation (26). Moreover, expression of cytolytic genes (GZMA, GZMK, and NGK7) in the same cluster supported that the T cells are cytotoxic effector cells. Class IIa UM contained little involvement of the immune system as opposed to class IIb UM, in which an inflammatory phenotype was observed (6). This subdivision is reminiscent of the class 3 and class 4 classification that the TCGA consortium recently described (13). It is however the question whether class II subclassification is warranted on molecular merits or that it is solely based on the degree of immune infiltration. The TCGA consortium estimated the leucocyte fraction with methylation profiles that were correlated to histopathologically determined leucocyte fractions (27). With this approach, leucocytes were found to be elevated in class 4, similar to what we observed with expression markers for T cells in class IIb UM (6). Though expression profiles and methylation profiles may be related to cell fractions, they do not represent absolute cell numbers. Expression and methylation profiles rather identify cell fractions that are present in the tumor tissue. Alternatively, absolute quantification of immune cells can be achieved by flow sort analysis of tumor material, but this is difficult and has not been applied to UM. T cell and macrophage counting in tissue with IHC has been

used as an accessible alternative and this confirmed wide ranging T cell and macrophage infiltration in UM (28,29). In the inflamed tumors the T cells and macrophages are spread throughout the tumor and thereby show that immune cell invasion is not limited to a specific histologic structure or tissue (figure 4). Moreover, integration of T cell and macrophage staining with the molecular progression model of UM revealed that macrophage infiltration precedes T cell infiltration in tumor inflammation (6,14). We integrated absolute T cell counts with expression profiles of UM, in order to investigate the biologic mechanisms. We quantified T cells with a DNA based quantitative method that would otherwise require fresh cell homogenates and flow cytometry (15). Integration of RNA expression profiles and DNA based T cell counts in the same tissue revealed that T cell fractions can exceed way over one tenth of the tumor mass. The highest T cell fractions were observed in class IIb UM, and the median T cell fraction was almost twice as high as in class IIa (15.6% and 8.0% respectively). Class I UM on the other hand presents the lowest T cell fraction (5.1%), and combined this suggests an accumulation of T cells during UM progression (Figure 2A). Remarkably, the elevated T cell numbers in class IIa were not reflected by an increase in T cell marker gene expression (Figure 2B). We suppose that dilution of the gene expression profiles of reduced T cell fractions (<10%) in class IIa obscured the contribution of T cells to the complete profile. Alternative explanation could be that CD3 is regulated during immune activation though this is not evident from the reference database that we use(17). In class IIa, 8% T cells were found compared to 5% in class I and it is questionable whether expression array analysis can distinguish this difference. Because of the absolute quantification with digital PCR approach, a gradual increase of T cell infiltrate could now be recognized. Whereas the gene expression analysis initially suggested

immune infiltration in specifically class IIb UM, absolute T cell quantification now showed that T cell influx occurs in the whole of class II UM but is highest in class IIb. Therefore, subclassification of class II UM with expression profiles appears to be based on a quantitative difference in T cell infiltration.

Earlier reports from our research group indicated that the immune system is involved in UM with poor prognosis (28,29). Our analyses revealed an extensive CD8 positive T cell infiltrate in UM and validated immune involvement in class II UM with a poor prognosis. However, both the gene expression based inflammatory phenotype of class IIb UM (Figure 1C), and T cell count in the whole of our UM panel (Table 2), did not correlate with survival. Indeed, class II UM contains more T cell infiltrate than class I UM, but when analyzed in a multivariate statistical model, including the known genetic risk factors, no added risk was revealed for T cell numbers. This also did not depend on T cell differentiation, as both CD4 expression as well as CD8 expression behaved neutrally in our risk model (Table 2).

The question remains how the immune infiltrate in UM should be further interpreted. To investigate this, we combined T cell quantifications with the gene expression profile of each corresponding tumor. The result of this analysis, a list of correlated genes (directly or indirectly related to the T cell immune infiltration), was integrated with publicly available cell-type specific gene expression profiles. This led to the remarkable conclusion that most of these T cell count-related genes are actually expressed by other cells in the immune compartment, mostly monocyte derived.

Deconvolution of the genes that are correlated with T cell count indicated that activated macrophages contribute considerably to the overall UM immune infiltrate as well as activated cytotoxic T cells. The fact that this activated immune infiltrate

does not result in an overt immune response, and consequently an improved prognosis, suggests that immune suppression is involved. Therefore, the eye is an immune privileged organ, making it a unique organ and a favorable location for allograft residence (30). Moreover, the blood-retina-barrier (BRB), which is characterized by tight junctions in the retinal pigmented epithelial layer, actively blocks the influx of immune cells from the surrounding tissue (31). Combined, this possibly reflects a negative selection pressure, as immune reactions in the eye could have detrimental effects on delicate structures, leading to impaired vision (32,33). The presence of activated immune cells in UM is therefore remarkable and may be a consequence of UM development. There is however no correlation to the development of metastases and this may suggest that immune infiltration can be regarded an epiphenomenon of progression that has no effect on survival of patients However, preliminary analysis indicates that activated macrophages, as marked by CXCL10 expression, may be involved in metastasis (Figure 5, Table 2). Interestingly, the role of CXCL10 in UM has been described before and this chemokine showed to be present in UM cells and upregulated in a T cell-rich environment (25,34,35). CXCL10 staining of UM sections in our cohort indicated that CXCL10 is present in some tumor cells but is predominantly found in macrophages (Figure 4B). Although the intensity levels of CXCL10 and macrophage marker gene expression varied in UM with high T cell counts, high levels of CXCL10 expression were restricted to the macrophages. Remarkably, though T cell count and CXCL10 are highly correlated, in survival analysis CXCL10 expression was correlated with a considerably increased risk while T cell count was not correlated to an increased risk. This suggests that besides attracting T cells by expressing CXCL10, macrophages also contribute to UM progression in another way. Possibly by

stimulating UM cell proliferation and extravasation in the same way that skin melanoma cells with ectopic expression of the CXCL10 receptor CXCR3 are stimulated (36,37). It is however unlikely that the same mechanism applies to UM, since CXCR3 was not differentially expressed between the UM classes (Table S1). Possibly other chemokine and chemokine receptor combinations drive tumor growth and progression in UM (38).

With absolute T cell quantification, we managed to take the first step in deconvolution of the immune compartment in UM. Thereby we revealed increasing numbers of activated T cells and activated macrophages in UM with poor prognosis. With CXCL10 expression by macrophages in UM we revealed a possible underlying mechanism of T cell infiltration. Based on survival analysis, we hypothesize that T cell infiltration is an epiphenomenon of a macrophage driven metastatic process. A further deconvolution of the macrophage related expression profile will be the approach to reveal the cells and the involved mechanisms.

Acknowledgements

We would like to thank A.W. Langerak for helpful discussion on T cell genetics and thank D. van Steenderen for counting T cells. N.A. Gruis and T. Larsen are acknowledged for proofreading of this manuscript. R.J. Nell is supported by the European Union's Horizon 2020 research and innovation program under grant agreement No 667787 (UM Cure 2020 project).

TABLES

Table 1, histological T cell and CXCL10 counts

UM	Class	T cell counts/mm^2	CXCL10 Macrophages	CXCL10 Tumor cells
05-005	IIb	5.61	Low	Low
05-020	I	18.00	Intermediate	Intermediate
05-046	IIa	5.55	Intermediate	High
05-061	IIa	8.31	Low	Intermediate
06-008	IIb	75.07	High	Low
06-009	IIa	2.77	Low	High
06-014	IIb	456.71	High	High
06-041	IIb	0.46	Low	Low
07-007	I	217.13	High	High
08-029	IIb	34.77	High	Low

Table 2, survival analysis of UM

Variables in the equation	B	p-value	Exp(B)
T cell fraction	,008	,726	1,008
Chromosome 6p copy number	,239	,529	1,269
Chromosome 3 copy number	-1,929	,001	,145
Chromosome 8q copy number	,321	,018	1,379
CD8 expression	-,282	,209	,754
CD4 expression	-,032	,969	,968
CXCL10 expression	,578	,017	1,782

1. Singh AD, Turell ME, Topham AK. Uveal melanoma: trends in incidence, treatment, and survival. Ophthalmology **2011**;118:1881-5
2. Van Raamsdonk CD, Bezrookove V, Green G, Bauer J, Gaugler L, O'Brien JM*, et al.* Frequent somatic mutations of GNAQ in uveal melanoma and blue naevi. Nature **2009**;457:599-602
3. Van Raamsdonk CD, Griewank KG, Crosby MB, Garrido MC, Vemula S, Wiesner T*, et al.* Mutations in GNA11 in uveal melanoma. NEnglJMed **2010**;363:2191-9
4. Feng X, Degese MS, Iglesias-Bartolome R, Vaque JP, Molinolo AA, Rodrigues M*, et al.* Hippo-independent activation of YAP by the GNAQ uveal melanoma oncogene through a trio-regulated rho GTPase signaling circuitry. Cancer cell **2014**;25:831-45
5. Sagoo MS, Harbour JW, Stebbing J, Bowcock AM. Combined PKC and MEK inhibition for treating metastatic uveal melanoma. Oncogene **2014**;33:4722-3
6. de Lange MJ, van Pelt SI, Versluis M, Jordanova ES, Kroes WG, Ruivenkamp C*, et al.* Heterogeneity revealed by integrated genomic analysis uncovers a molecular switch in malignant uveal melanoma. Oncotarget **2015**:10
7. McNamara M, Felix C, Davison EV, Fenton M, Kennedy SM. Assessment of chromosome 3 copy number in ocular melanoma using fluorescence in situ hybridization. Cancer GenetCytogenet **1997**;98:4-8
8. Speicher MR, Prescher G, du MS, Jauch A, Horsthemke B, Bornfeld N*, et al.* Chromosomal gains and losses in uveal melanomas detected by comparative genomic hybridization. Cancer Res **1994**;54:3817-23
9. Prescher G, Bornfeld N, Hirche H, Horsthemke B, Jockel KH, Becher R. Prognostic implications of monosomy 3 in uveal melanoma. Lancet **1996**;347:1222-5
10. Patel KA, Edmondson ND, Talbot F, Parsons MA, Rennie IG, Sisley K. Prediction of prognosis in patients with uveal melanoma using fluorescence in situ hybridisation. BrJOphthalmol **2001**;85:1440-4
11. Scholes AG, Damato BE, Nunn J, Hiscott P, Grierson I, Field JK. Monosomy 3 in uveal melanoma: correlation with clinical and histologic predictors of survival. Invest OphthalmolVisSci **2003**;44:1008-11
12. Onken MD, Worley LA, Ehlers JP, Harbour JW. Gene expression profiling in uveal melanoma reveals two molecular classes and predicts metastatic death. Cancer Res **2004**;64:7205-9
13. Robertson AG, Shih J, Yau C, Gibb EA, Oba J, Mungall KL*, et al.* Integrative Analysis Identifies Four Molecular and Clinical Subsets in Uveal Melanoma. Cancer Cell **2017**;32:204-20.e15
14. Gezgin G, Dogrusoz M, van Essen TH, Kroes WG, Luyten GP, van der Velden PA*, et al.* Genetic evolution of uveal melanoma guides the development of an inflammatory microenvironment. Cancer immunology, immunotherapy : CII **2017**
15. Zoutman WH, Nell RJ, Versluis M, van Steenderen D, Lalai RN, Out-Luiting JJ*, et al.* Accurate Quantification of T Cells by Measuring Loss of Germline T-Cell Receptor Loci with Generic Single Duplex Droplet Digital PCR Assays. The Journal of molecular diagnostics : JMD **2017**;19:236-43
16. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D*, et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. Proceedings of the National Academy of Sciences of the United States of America **2004**;101:6062-7
17. Mabbott NA, Baillie JK, Brown H, Freeman TC, Hume DA. An expression atlas of human primary cells: inference of gene function from coexpression networks. BMC Genomics **2013**;14:632
18. Wu C, Jin X, Tsueng G, Afrasiabi C, Su AI. BioGPS: building your own mash-up of gene annotations and expression profiles. Nucleic acids research **2016**;44:D313-6
19. Wu C, Macleod I, Su AI. BioGPS and MyGene.info: organizing online, gene-centric information. Nucleic acids research **2013**;41:D561-5

20. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S*, et al.* BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome biology **2009**;10:R130
21. Heeren AM, Punt S, Bleeker MC, Gaarenstroom KN, van der Velden J, Kenter GG*, et al.* Prognostic effect of different PD-L1 expression patterns in squamous cell carcinoma and adenocarcinoma of the cervix. Mod Pathol **2016**;29:753-63
22. Tschentscher F, Husing J, Holter T, Kruse E, Dresen IG, Jockel KH*, et al.* Tumor classification based on gene expression profiling shows that uveal melanomas with and without monosomy 3 represent two distinct entities. Cancer Res **2003**;63:2578-84
23. Robins HS, Ericson NG, Guenthoer J, O'Briant KC, Tewari M, Drescher CW*, et al.* Digital genomic quantification of tumor-infiltrating lymphocytes. Science translational medicine **2013**;5:214ra169
24. Dik WA, Pike-Overzet K, Weerkamp F, de Ridder D, de Haas EF, Baert MR*, et al.* New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. The Journal of experimental medicine **2005**;201:1715-23
25. Jehs T, Faber C, Juel HB, Bronkhorst IH, Jager MJ, Nissen MH. Inflammation-induced chemokine expression in uveal melanoma cell lines stimulates monocyte chemotaxis. Investigative ophthalmology & visual science **2014**;55:5169-75
26. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR*, et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest **2017**;127:2930-40
27. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T*, et al.* Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol **2012**;30:413-21
28. Bronkhorst IH, Ly LV, Jordanova ES, Vrolijk H, Versluis M, Luyten GP*, et al.* Detection of M2 macrophages in uveal melanoma and relation with survival. Invest Ophthalmol Vis Sci **2010**
29. Bronkhorst IH, Vu TH, Jordanova ES, Luyten GP, Burg SH, Jager MJ. Different subsets of tumor-infiltrating lymphocytes correlate with macrophage influx and monosomy 3 in uveal melanoma. Invest Ophthalmol Vis Sci **2012**;53:5370-8
30. Medawar PB. Immunity to homologous grafted skin; the fate of skin homografts transplanted to the brain, to subcutaneous tissue, and to the anterior chamber of the eye. British journal of experimental pathology **1948**;29:58-69
31. Campbell M, Humphries P. The blood-retina barrier: tight junctions and barrier modulation. Advances in experimental medicine and biology **2012**;763:70-84
32. Taylor AW. Ocular Immune Privilege and Transplantation. Frontiers in immunology **2016**;7:37
33. Streilein JW. Ocular immune privilege: therapeutic opportunities from an experiment of nature. Nature reviews Immunology **2003**;3:879-89
34. Triozzi PL, Aldrich W, Singh A. Effects of interleukin-1 receptor antagonist on tumor stroma in experimental uveal melanoma. Investigative ophthalmology & visual science **2011**;52:5529-35
35. Borden EC, Jacobs B, Hollovary E, Rybicki L, Elson P, Olencki T*, et al.* Gene regulatory and clinical effects of interferon beta in patients with metastatic melanoma: a phase II trial. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research **2011**;31:433-40
36. Correa D, Somoza RA, Lin P, Schiemann WP, Caplan AI. Mesenchymal stem cells regulate melanoma cancer cells extravasation to bone and liver at their perivascular niche. International journal of cancer **2016**;138:417-27
37. Robledo MM, Bartolomé RA, Longo N, Rodríguez-Frade JM, Mellado M, Longo I*, et al.* Expression of functional chemokine receptors CXCR3 and CXCR4 on human melanoma cells. Journal of Biological Chemistry **2001**;276:45098-105

38. Dobner BC, Riechardt AI, Joussen AM, Englert S, Bechrakis NE. Expression of haematogenous and lymphogenous chemokine receptors and their ligands on uveal melanoma in association with liver metastasis. Acta Ophthalmol **2012**;90:e638-44

Figure legends

Figure 1

Classification of UM using gene expression analysis. (A)Unsupervised hierarchical clustering of genome wide expression levels, divides UM in two prognostic classes (class I and class II). By supervised clustering of the classifier genes, class II is even further subdivided into class IIa and class IIb. Subdivision in class IIa and classIIb does not result in a survival difference (6). (B) The genes that define class IIb reveal an immunologic signature.

Figure 2

Quantification of T cells in 64 UM divided over the three gene expression classes. (A) Significantly higher T cell fractions are found in class IIa and class IIb compared to class I and class IIa respectively. (B) CD3D expression (Y-axis) in the classes, with a significant expression difference between class IIa and class IIb but no significant expression difference between class I and class IIa. (C) Correlation of CD3D expression (Y-axis) with T cell count in the three expression classes.

Figure 3

Deconvolution of the immune phenotype of UM with T cell count. (A) Analysis of the UM classifier genes for T cell-related genes consisted of two additional steps; a correlation between T cell count and gene expression level and analysis of cell type distribution of differentially expressed genes. (B) Relative expression levels of the 60 T cell-correlated genes in a range of immune cells. (C) T cell-related genes dispersed in 4 cell types after clustering for cell type distribution. In blue the

monocyte/macrophage cluster, in green the activated macrophages, in pink the activated cytotoxic T cells and in orange an undefined population of immune cells.

Figure 4

Multiplex immunofluorescent staining of UM samples with a high (06-014: 42%) and low (06-009: 0%) calculated T cell fraction. CXCL10 co-localization with T cells and macrophages in UM. (A) CXCL10 (red) does not co-localize with T cells (CD8: green) and hardly ever with melanoma cells (melanA: blue). Zoomed in picture insert indicates melanoma cells that express CXCL10. (B) CXCL10 (blue) co-localizes with macrophages (CD163: green, CD14: red) of varying polarization.

Figure 5

Survival analysis of UM containing high numbers of T cells did not show a benefit for T cell infiltrate. Class I (green) presents with a good prognosis while class II UM are correlated with a poor prognosis. Subdivision of class II in low and high T cell infiltrated UM does not make a difference in survival.

Figure 1

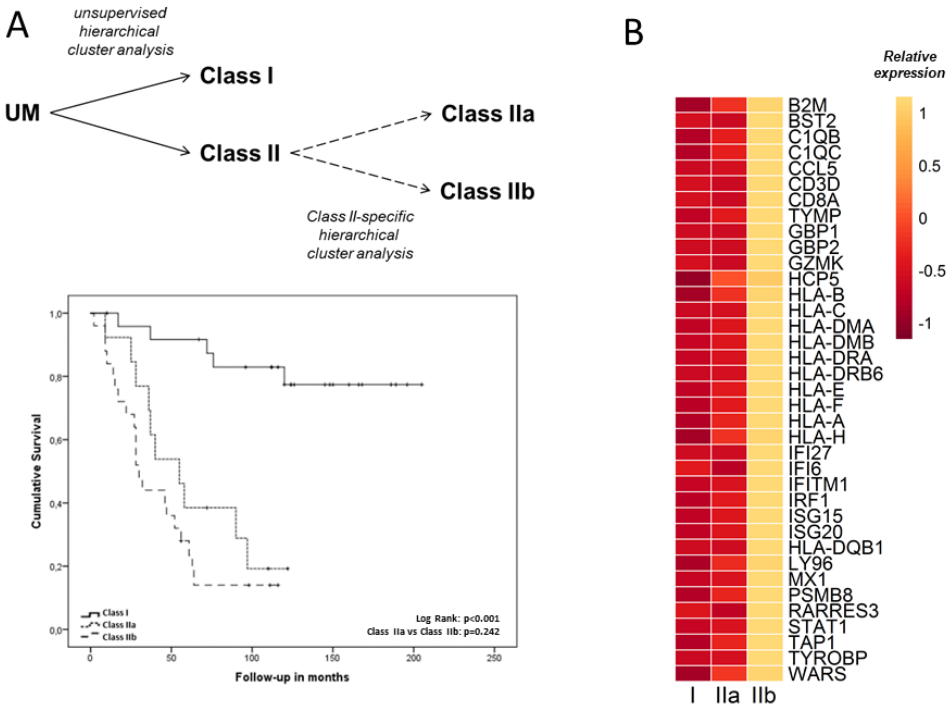

A
unsupervised hierarchical cluster analysis
UM
Class I
Class II
Class IIa
Class IIb
Class II-specific hierarchical cluster analysis
Cumulative Survival
Follow-up in months
Class I
Class IIa
Class IIb
Log Rank: p<0.001
Class IIa vs Class IIb: p=0.242
B
Relative expression
1
0.5
0
-0.5
-1
B2M
BST2
C1QB
C1QC
CCL5
CD3D
CD8A
TYMP
GBP1
GBP2
GZMK
HCP5
HLA-B
HLA-C
HLA-DMA
HLA-DMB
HLA-DRA
HLA-DRB6
HLA-E
HLA-F
HLA-A
HLA-H
IFI27
IFI6
IFITM1
IRF1
ISG15
ISG20
HLA-DQB1
LY96
MX1
PSMB8
RARRES3
STAT1
TAP1
TYROBP
WARS
I IIa IIb

Figure 2

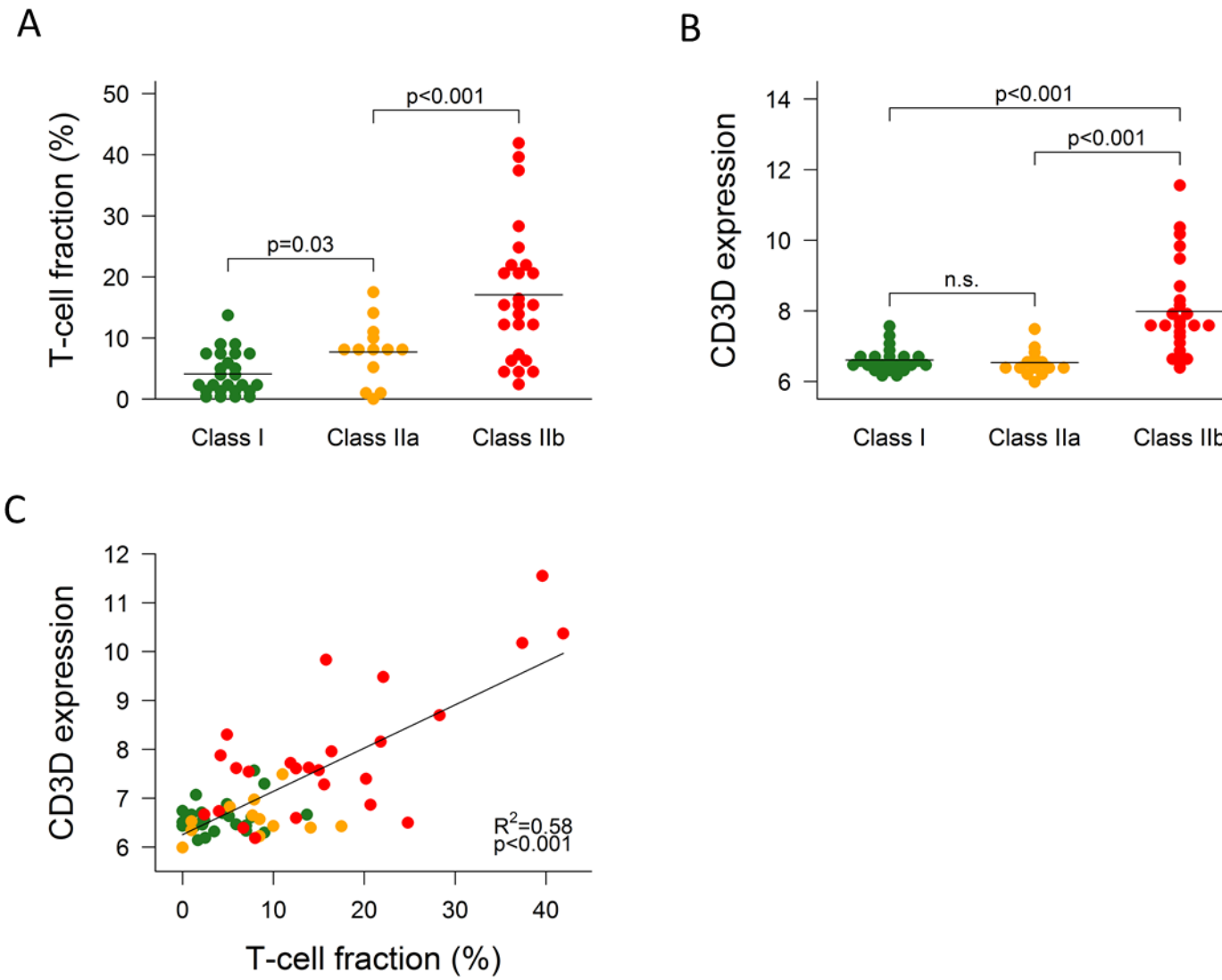

A
T-cell fraction (%)
50
40
30
20
10
0
p<0.001
p=0.03
Class I
Class IIa
Class IIb
B
CD3D expression
14
12
10
8
6
p<0.001
p<0.001
n.s.
Class I
Class IIa
Class IIb
C
CD3D expression
12
11
10
9
8
7
6
0
10
20
30
40
T-cell fraction (%)
R2=0.58
p<0.001

Figure 3A

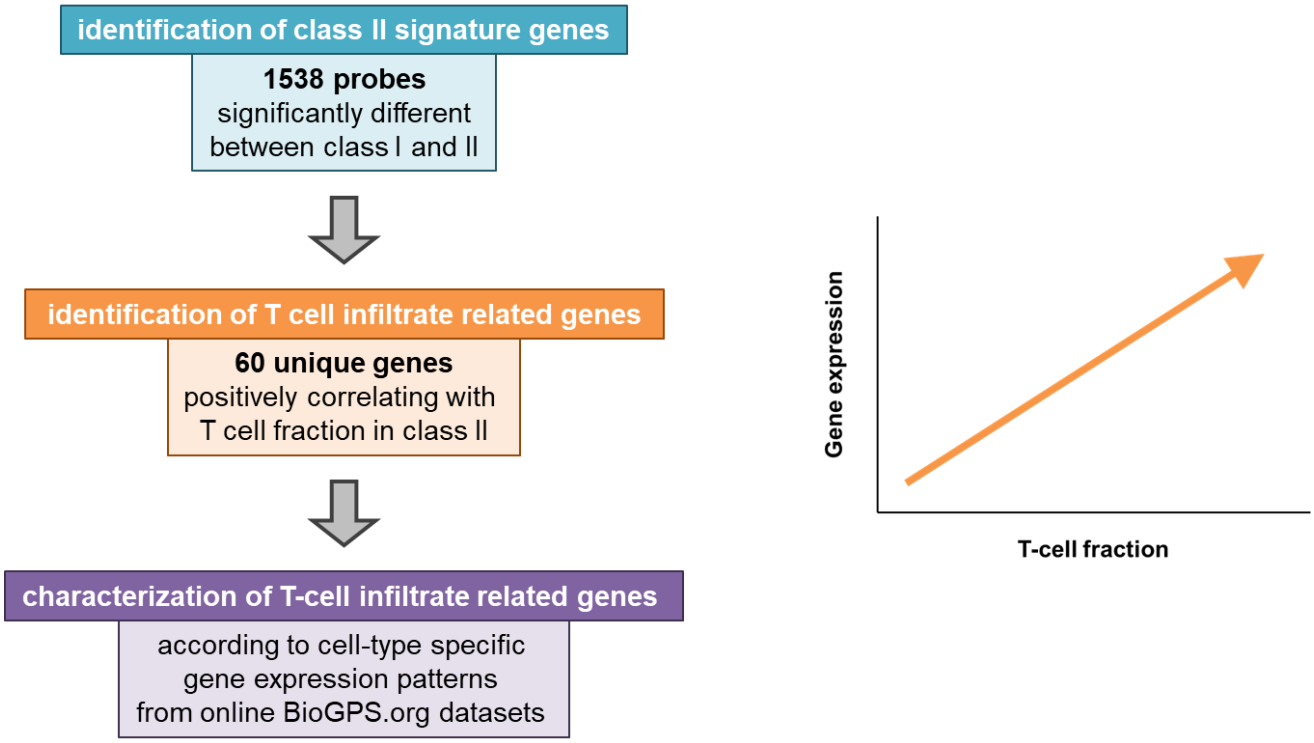
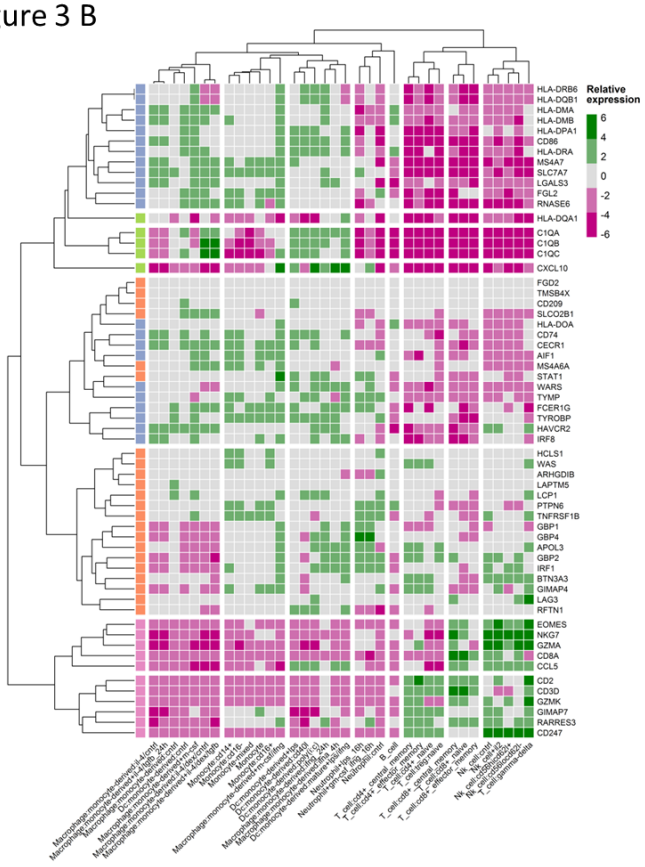

identification of class II signature genes
1538 probes
significantly different
between class I and II
identification of T cell infiltrate related genes
60 unique genes
positively correlating with
T cell fraction in class II
characterization of T-cell infiltrate related genes
according to cell-type specific
gene expression patterns
from online BioGPS.org datasets
Gene expression
T-cell fraction

Figure 3 B


Relative expression
6
4
2
0
-2
-4
-6
HLA-DRB6
HLA-DQB1
HLA-DMA
HLA-DMB
HLA-DPA1
CD86
HLA-DRA
MS4A7
SLC7A7
LGALS3
FGL2
RNASE6
HLA-DQA1
C1QA
C1QB
C1QC
CXCL10
FGD2
TMSB4X
CD209
SLCO2B1
HLA-DOA
CD74
CECR1
AIF1
MS4A6A
STAT1
WARS
TYMP
FCER1G
TYROBP
HAVCR2
IRF8
HCLS1
WAS
ARHGDIB
LAPTM5
LCP1
PTPN6
TNFRSF1B
GBP1
GBP4
APOL3
GBP2
IRF1
BTN3A3
GIMAP4
LAG3
RFTN1
EOMES
NKG7
GZMA
CD8A
CCL5
CD2
CD3D
GZMK
GIMAP7
RARRES3
CD247

3C

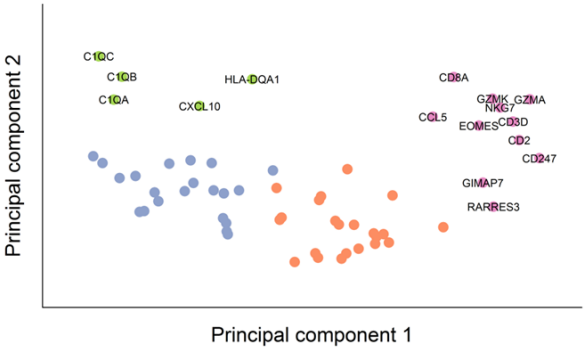

C1QC
C1QB
C1QA
HLA-DQA1
CXCL10
CD8A
GZMK
GZMA
NKG7
CCL5
EOMES
CD3D
CD2
CD247
GIMAP7
RARRES3
Principal component 2
Principal component 1

Figure 4

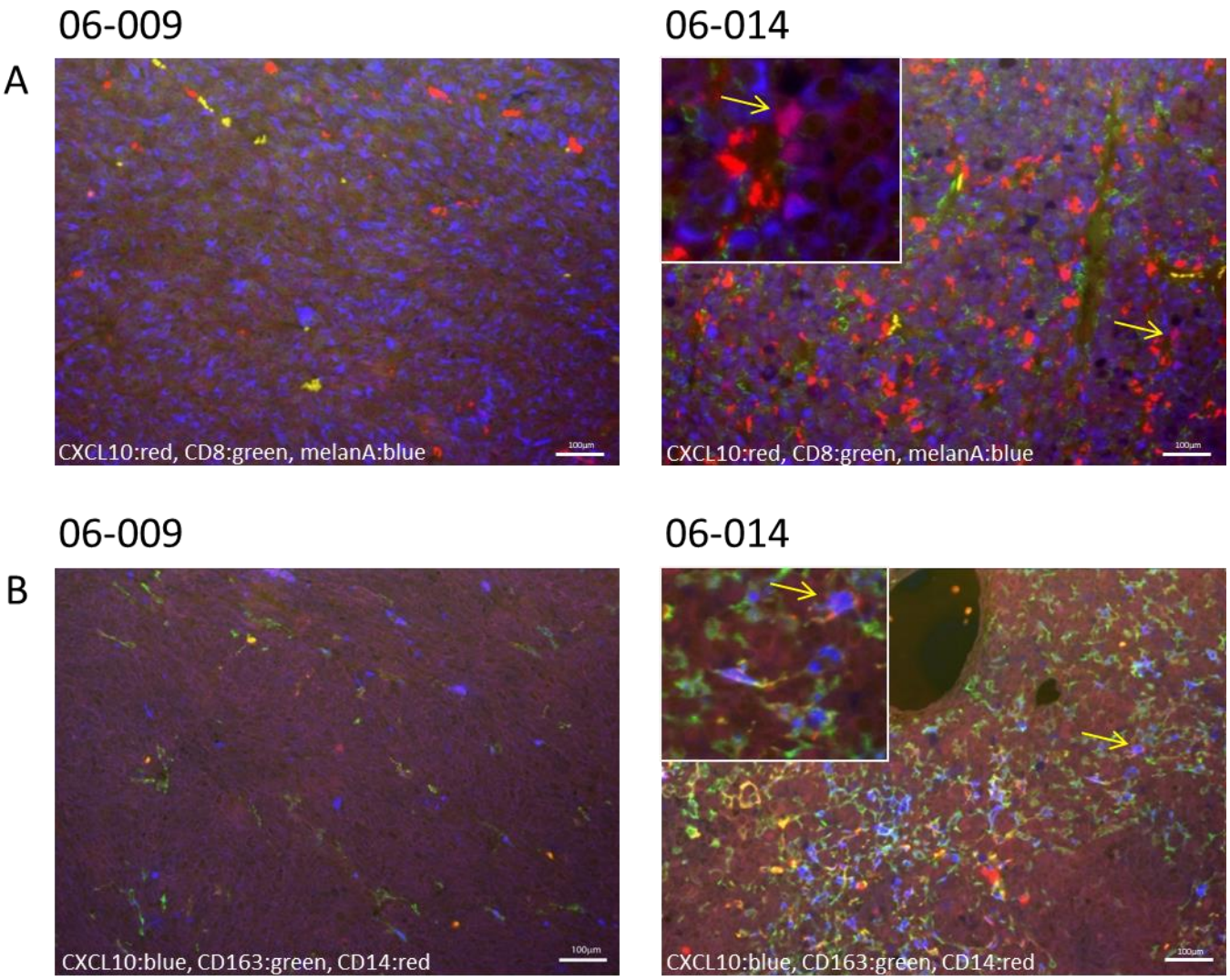
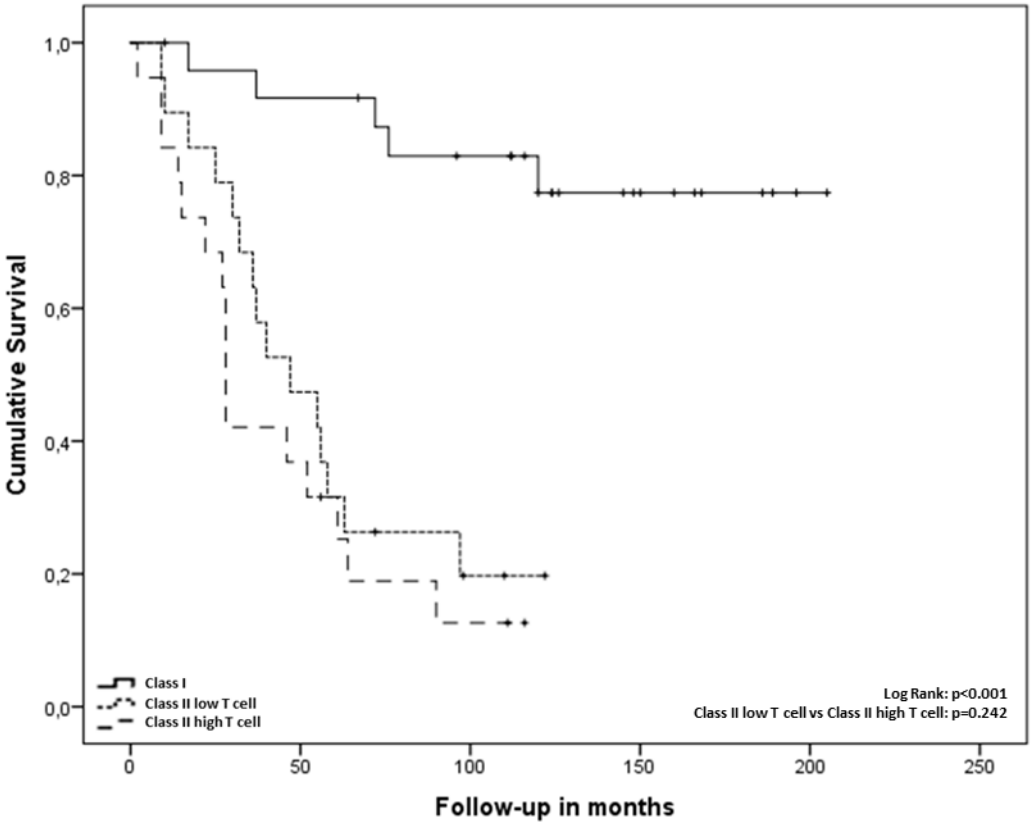

06-009
06-014
A
CXCL10:red, CD8:green, melanA:blue
CXCL10:red, CD8:green, melanA:blue
06-009
06-014
B
CXCL10:blue, CD163:green, CD14:red
CXCL10:blue, CD163:green, CD14:red

Figure 5


Cumulative Survival
1,0
0,8
0,6
0,4
0,2
0,0
0
50
100
150
200
250
Follow-up in months
Class I
Class II low T cell
Class II high T cell
Log Rank: p<0.001
Class II low T cell vs Class II high T cell: p=0.242

Molecular T cell quantification assisted deconvolution of the microenvironment in uveal melanoma

Supplemental methods

Table of Contents

Initialize libraries

Dependencies for this project are `RColorBrewer`, `grid`, a modified version of `pheatmap` and `dendsort`:

```
# Load pheatmap, with support for text rotation in row and col names (stevepe
d/pheatmap commit 5dcb7f00e4ec64c0f61e500b40abab96567201c1)
source("rpheatmap.R")

# Load other libraries
library("RColorBrewer")
library("grid")
library("dendsort")

# Default heatmap settings
hm_palette = colorRampPalette(c("#BF0080", "#CE6EAE", "#dddddd", "#6EAE6E", "
#008000", "white"))(n = 6)
hm_breaks = c(-6, -3.6, -1.2, 1.2, 3.6, 6,9)
hm_legend_breaks = c(-5.7,-3.8,-1.9,0,1.9,3.8,5.7,7.5,8.5)
hm_legend_labels = c("     -6",
                     "     -4",
                     "     -2",
                     "     0",
                     "     2",
                     "     4",
                     "     6",
                     "expression",
                     "Relative")

# Apply dendsort callback for pheatmap
dendsort_callback = function(hc, ...) { dendsort(hc, type = "min") }
```

Loading data

Loading data (gene expression data of 1538 most significant class I/II classifier probes and T-cell fractions).

```
load("../data/1538.RData")
data[1,1:8]
```

```
##        probeid  gene locus entrezid   01-042  01-074  02-158  02-167
## 1 ILMN_2193980 ABCB6  2q36    10058 8.043216 7.60644 8.19975 7.34011
```

Correlations with T-cell fraction

T-cell fraction is available in 63/64 samples.

```
# Which tumors should be correlated? All class II, without sample 07-005 (T-c
ell fraction NA)
selection = c(26:36,38:64)
tumordata = data[,5:68]
tcf_selection = tcf[selection]

# Multiple testing correction: divide alpha by number of tested probes (1538)
alpha = 0.05/1538

# Initialize overall results table
results = NULL

# For every selected probe
for (i in 1:nrow(data)) {
  # Get probe information
  expression_selection = as.numeric(tumordata[i,selection])
  probeid = as.character(data[i,1])
  gene = as.character(data[i,2])
  entrezid = as.character(data[i,4])

  # Calculate pearson R
  pearson_r = cor(x = tcf_selection,
                  y = expression_selection,
                  method = "pearson")

  # Calculate pearson p-value
  pearson_p = cor.test(x = tcf_selection,
                       y = expression_selection,
                       method = "pearson",
                       alternative="two.sided")$p.value

  # Check whether pearson p-value is below alpha (is correlation significant?
)
  pearson_significant = pearson_p < alpha
```

```
  # Bind probe results to overall results table
  results = rbind(results, c(probeid,
                             gene,
                             entrezid,
                             pearson_r,
                             pearson_p,
                             pearson_significant))
}

# Add col names to overall results table
colnames(results) = c("probeid",
                      "gene",
                      "entrezid",
                      "pearson_r",
                      "pearson_p",
                      "pearson_significant")
```

Results are saved in correlation.csv:

```
write.csv(results,"../results/correlation.csv")
```

Selection of significant probes

We select unique genes from probes with a significant correlation:

```
# Select probes with significant correlation
significant_selection = results[which(results[,6] == TRUE),]

# Find unique genes represented by these probes
significant_genes = unique(significant_selection[,2])
significant_ids = unique(significant_selection[,3])
```

... which results in the following selection of 60 genes from 68 probes:

```
significant_genes
```

```
##  [1] "AIF1"     "APOL3"    "ARHGDIB"  "BTN3A3"   "C1QA"     "C1QB"
##  [7] "C1QC"     "CCL5"     "CD2"      "CD209"    "CD247"    "CD3D"
## [13] "CD74"     "CD86"     "CD8A"     "CECR1"    "CXCL10"   "TYMP"
## [19] "EOMES"    "FCER1G"   "FGD2"     "FGL2"     "GBP1"     "GBP2"
## [25] "GBP4"     "GIMAP4"   "GIMAP7"   "GZMA"     "GZMK"     "HAVCR2"
## [31] "HCLS1"    "HLA-DMA"  "HLA-DMB"  "HLA-DOA"  "HLA-DPA1" "HLA-DQA1"
## [37] "HLA-DRA"  "HLA-DRB6" "IRF1"     "IRF8"     "LAG3"     "LAPTM5"
## [43] "LCP1"     "LGALS3"   "HLA-DQB1" "MS4A6A"   "MS4A7"    "NKG7"
## [49] "PTPN6"    "RARRES3"  "RFTN1"    "RNASE6"   "SLC7A7"   "SLCO2B1"
## [55] "STAT1"    "TMSB4X"   "TNFRSF1B" "TYROBP"   "WARS"     "WAS"
```

Load BioGPS data

We observe that some genes are very well correlating with T-cell infiltrate: the more T cells, the more of expression of these genes. From which cell types does expression of these genes come from? In order to answer this question, we use data from the Primary Cell Atlas, obtained via the BioGPS website. This service provide cell-type specific gene expression profiles, which we want to integrate with our list of correlating genes.

```
# Initialize BioGPS results table
biogps_results = NULL
a_is = NULL
na_is = NULL

# For every significantly correlated, unique gene
for (i in 1:length(significant_ids)) {
  # Get data from URL based on gene_id
  biogps_file = tryCatch({read.csv(paste0("http://ds.biogps.org/dataset/csv/B
DS_00013/gene/",
                                          as.numeric(significant_ids[i]),"/")
,
                                   header = TRUE,
                                   sep = ",",
                                   quote = "\"",
                                   dec = ".",
                                   fill = TRUE)},
                         warning = function(w) {"NA"},
                         error = function(e) {"NA"})

  # If BioGPS data is available for this gene
  if (length(biogps_file) > 1) {

    # If data from >1 probe for a gene is available
    if (ncol(biogps_file) - 1 > 1) {
      n_select = 2
      n_select_mean = 0
      # Check which probe has the highest range of expression
      for (n in 2:ncol(biogps_file)) {
        if (n_select_mean < max(biogps_file[,n]) - min(biogps_file[,n])) {
          n_select_mean = max(biogps_file[,n]) - min(biogps_file[,n])
          n_select = n
        }
      }
      # And use data from that probe
      biogps_data = biogps_file[,n]
    }

    # If only data from 1 probe for a gene is available
    else {
      # Use data from that probe
```

```
      biogps_data = biogps_file[,2]
    }

    # Normalize obtained data
    biogps_data_normalized = log2(biogps_data / mean(biogps_data))

    # Attach normalized data to `biogps_results`
    biogps_results = rbind(biogps_results, biogps_data_normalized)

    # Mark this gene as `BioGPS data available`
    a_is = c(a_is,i)
  }

  # If NO BioGPS data is available for this gene
  else {
    # Mark this gene as `BioGPS data NOT available`
    na_is = c(na_is,i)
  }
}

# Match column names with genes
rownames(biogps_results) = significant_genes[a_is]
colnames(biogps_results) = read.table("../data/groteset_template.csv",
                                      sep = ",",
                                      comment.char="#",
                                      header=T,
                                      dec=".")[,1]
```

For 60/60 (=100%) genes BioGPS data is available.

Merge BioGPS data of replicates

Data from 745 cell types is available in the Primary Cell Atlas. Merge data (based on mean) of cell type replicates:

```
biogps_results_merged = NULL

# Determine which cell types are available and which are unique
all_celltypes = colnames(biogps_results)
unique_celltypes = unique(colnames(biogps_results))

# For every unique cell type
for (i in 1:length(unique_celltypes)) {
  to_merge = which(all_celltypes == unique_celltypes[i])

  # If >1 cell type `replicates` are available
  if (length(to_merge) > 1) {
    # Take the mean of those `replicates`
```

```
    merged = rowMeans(biogps_results[,to_merge])
  }

  # If only 1 cell type `replicate` is available
  else {
    # Take data from that `replicate`
    merged = biogps_results[,to_merge]
  }

  # Merge results from all cell types
  biogps_results_merged = cbind(biogps_results_merged,merged)
}

# Cell types as col names
colnames(biogps_results_merged) = unique_celltypes
```

... resulting in available BioGPS data for 188 cell types.

Among these cell types less relevant ones, e.g. `Gametocytes:spermatocyte`, are also present, which should be excluded. Immune cell types are positively selected:

```
# Make a selection of immune cell types
selection_immunological = c(112:123,
                            126,
                            132:133,
                            136,
                            138:140,
                            #146,
                            155:157,
                            162,
                            166:169,
                            171:174,
                            175,
                            178:180)

biogps_selection_immunological = biogps_results_merged[,selection_immunologic
al]
```

...resulting in the selection of the following 35 cell types:

```
colnames(biogps_results_merged[,selection_immunological])
```

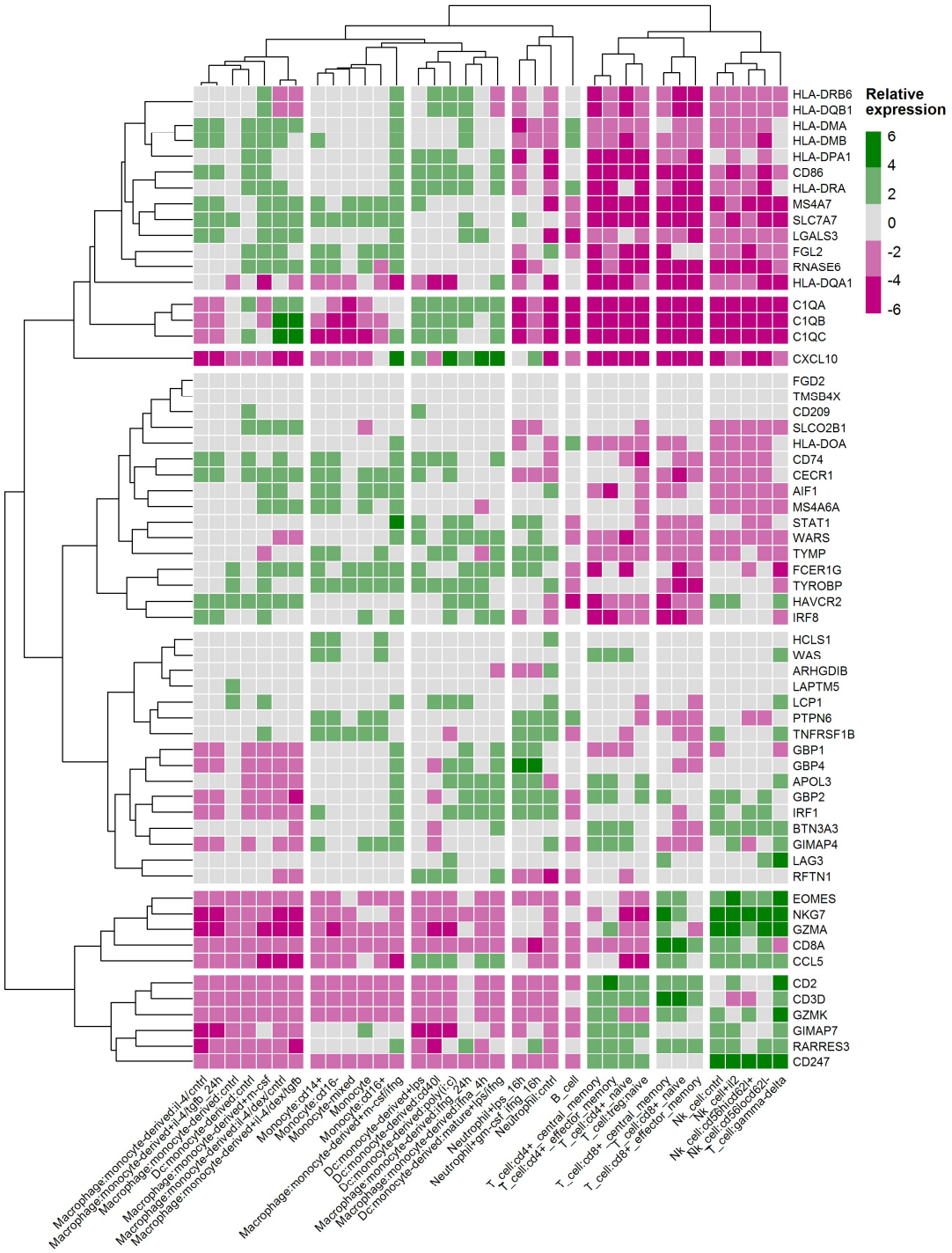
```
##  [1] "Monocyte-mixed"
##  [2] "Monocyte:cd14+"
##  [3] "Monocyte:cd16-"
##  [4] "Monocyte:cd16+"
##  [5] "Monocyte"
##  [6] "Macrophage:monocyte-derived:il-4/cntrl"
##  [7] "Macrophage:monocyte-derived:il-4/dex/cntrl"
##  [8] "Macrophage:monocyte-derived+il-4/dex/tgfb"
##  [9] "Macrophage:monocyte-derived+il-4/tgfb_24h"
```

```
## [10] "Macrophage:monocyte-derived:cntrl"
## [11] "Macrophage:monocyte-derived+m-csf"
## [12] "Macrophage:monocyte-derived+m-csf/ifng"
## [13] "Macrophage:monocyte-derived:ifng_24h"
## [14] "Macrophage:monocyte-derived:ifna_4h"
## [15] "Dc:monocyte-derived:cntrl"
## [16] "Dc:monocyte-derived:mature+lps/ifng"
## [17] "Dc:monocyte-derived+lps"
## [18] "Dc:monocyte-derived:poly(i:c)"
## [19] "Dc:monocyte-derived:cd40l"
## [20] "T_cell:cd4+_naive"
## [21] "T_cell:cd4+_central_memory"
## [22] "T_cell:cd4+_effector_memory"
## [23] "T_cell:cd8+_central_memory"
## [24] "T_cell:cd8+_effector_memory"
## [25] "T_cell:cd8+_naive"
## [26] "T_cell:gamma-delta"
## [27] "T_cell:treg:naive"
## [28] "Nk_cell:cntrl"
## [29] "Nk_cell+il2"
## [30] "Nk_cell:cd56hicd62l+"
## [31] "Nk_cell:cd56locd62l-"
## [32] "Neutrophil:cntrl"
## [33] "Neutrophil+lps_16h"
## [34] "Neutrophil+gm-csf_ifng_16h"
## [35] "B_cell"
```

Heatmap on BioGPS data

Create heatmap to visualize BioGPS data of selected immune cell types:

```
pheatmap(
  biogps_selection_immunological,
  clustering_callback = dendsort_callback,
  treeheight_row = 125,
  cutree_rows = 7,
  cutree_cols = 8,
  breaks = hm_breaks,
  color = hm_palette,
  legend_breaks = hm_legend_breaks,
  legend_labels = hm_legend_labels,)
```



... which is also stored in ../results/heatmap.tiff in 600 dpi.

```
tiff("../results/heatmap.tiff",
     width=9,
     height=12,
     units="in",
```

```
    res=600,
    compression = "lzw")

pheatmap(
  biogps_selection_immunological,
  clustering_callback = dendsort_callback,
  treeheight_row = 125,
  cutree_rows = 7,
  cutree_cols = 8,
  breaks = hm_breaks,
  color = hm_palette,
  legend_breaks = hm_legend_breaks,
  legend_labels = hm_legend_labels)

dev.off()

## png
##   2
```

Principal component analysis

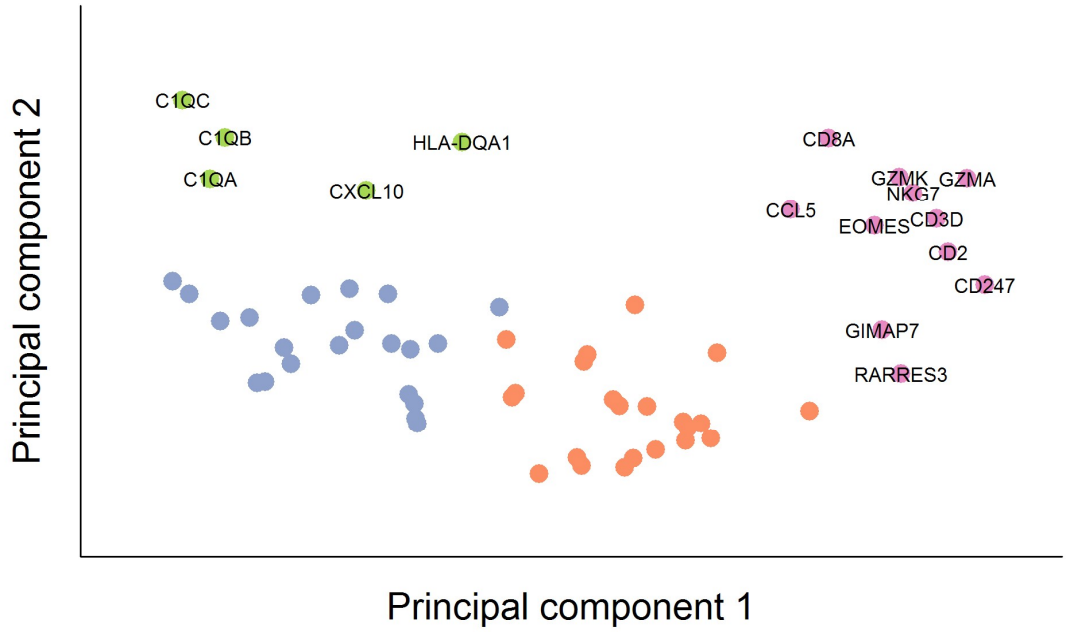
Using PCA to visualize to BioGPS results with our gene selection.

```
pca = prcomp((biogps_selection_immunological))
```

First two PCA components are plotted in a 2D scatterplot. Every dot represents one gene, genes clustered together show similar expression profile across different cell types.

```
set.seed(1)
kmeans = kmeans(biogps_selection_immunological, 4,nstart=1)
col = c("#FC8D62","#8DA0CB","#E78AC3","#A6D854")

par(mar=c(3,3,1,1))
plot(pca$x[, 1], pca$x[, 2], xlim = c(-16,18), ylim = c(-10,15), xlab="", yla
b="", bty='l',xaxt='n',yaxt="n",pch=16,cex=1.5, col=col[kmeans$cluster])
mtext("Principal component 1",side=1,line=1,font=1,cex=1.3)
mtext("Principal component 2",side=2,line=1,font=1,cex=1.3)
labels = rownames(pca$x)
labels[which(kmeans$cluster<3)] = ""
text(pca$x[, 1], pca$x[, 2], labels=labels, cex=0.7)
```



... which is also stored in ../results/pca.tiff in 600 dpi.

```
tiff("../results/pca.tiff",
     width=6.5,
     height=4,
     units="in",
     res=600,
     compression = "lzw")


par(mar=c(3,3,1,1))
plot(pca$x[, 1], pca$x[, 2], xlim = c(-16,18), ylim = c(-10,15), xlab="", yla
b="", bty='l',xaxt='n',yaxt="n",pch=16,cex=1.5, col=col[kmeans$cluster])
mtext("Principal component 1",side=1,line=1,font=1,cex=1.3)
mtext("Principal component 2",side=2,line=1,font=1,cex=1.3)
labels = rownames(pca$x)
labels[which(kmeans$cluster<3)] = ""
text(pca$x[, 1], pca$x[, 2], labels=labels, cex=0.7)


dev.off()

## png
##   2
```

Heatmap on BioGPS data with PCA annotations

Create heatmap with PCA annotations:

```r
ann = data.frame(kmeans$cluster)
rownames(ann) = rownames(pca$x)
colnames(ann) = "cluster"

pheatmap(
  biogps_selection_immunological,
  clustering_callback = dendsort_callback,
  treeheight_row = 125,
  cutree_rows = 8,
  cutree_cols = 8,
  breaks = hm_breaks,
  color = hm_palette,
  legend_breaks = hm_legend_breaks,
  legend_labels = hm_legend_labels,
  annotation_row = ann,
  annotation_legend = FALSE,
  annotation_colors = list(cluster = col),
  clustering_method = "complete",
  annotation_names_row = FALSE)
```

... which is also stored in ../results/heatmap_pca_annotated.tiff in 600 dpi.

```
tiff("../results/heatmap_pca_annotated.tiff",
     width=9,
     height=12,
     units="in",
```

```
     res=600,
     compression = "lzw")


pheatmap(
  biogps_selection_immunological,
  clustering_callback = dendsort_callback,
  treeheight_row = 125,
  cutree_rows = 8,
  cutree_cols = 8,
  breaks = hm_breaks,
  color = hm_palette,
  legend_breaks = hm_legend_breaks,
  legend_labels = hm_legend_labels,
  annotation_row = ann,
  annotation_legend = FALSE,
  annotation_colors = list(cluster = col),
  clustering_method = "complete",
  annotation_names_row = FALSE)

dev.off()
```

```
## png
##   2
```

Supplementary table 1

probeid	gene	entrezid	pearson_r	pearson_p	pearson_significant
ILMN_1792473	AIF1	199	0,69	2,10E+08	TRUE
ILMN_1756862	APOL3	80833	0,72	3,95E+07	TRUE
ILMN_1678143	ARHGDIB	397	0,72	4,46E+07	TRUE
ILMN_2373831	BTN3A3	10384	0,62	3,19E+09	TRUE
ILMN_1737918	C1QA	712	0,71	5,14E+07	TRUE
ILMN_1796409	C1QB	713	0,64	1,31E+09	TRUE
ILMN_1785902	C1QC	714	0,67	4,22E+08	TRUE
ILMN_1773352	CCL5	6352	0,73	1,48E+07	TRUE
ILMN_2098126	CCL5	6352	0,73	1,73E+07	TRUE
ILMN_1695025	CD2	914	0,78	1,02E+06	TRUE
ILMN_1676372	CD209	30835	0,64	1,24E+09	TRUE
ILMN_2377669	CD247	919	0,79	3,81E+05	TRUE
ILMN_2261416	CD3D	915	0,74	1,29E+07	TRUE
ILMN_2325837	CD3D	915	0,74	9,02E+06	TRUE
ILMN_1736567	CD74	972	0,63	2,23E+09	TRUE
ILMN_1761464	CD74	972	0,69	1,79E+08	TRUE
ILMN_2379644	CD74	972	0,63	2,27E+09	TRUE
ILMN_1714602	CD86	942	0,69	1,40E+08	TRUE
ILMN_1768482	CD8A	925	0,74	9,33E+06	TRUE
ILMN_2353732	CD8A	925	0,74	9,12E+06	TRUE
ILMN_1751851	CECR1	51816	0,63	2,69E+07	TRUE
ILMN_1791759	CXCL10	3627	0,62	2,96E+09	TRUE
ILMN_2109708	TYMP	1890	0,67	5,21E+08	TRUE
ILMN_1760509	EOMES	8320	0,76	4,13E+06	TRUE
ILMN_2123743	FCER1G	2207	0,63	2,46E+09	TRUE
ILMN_2115005	FGD2	221472	0,75	5,68E+06	TRUE
ILMN_1693009	FGL2	10875	0,68	2,26E+08	TRUE
ILMN_2148785	GBP1	2633	0,64	1,82E+09	TRUE
ILMN_1701114	GBP1	2633	0,68	2,51E+08	TRUE
ILMN_1774077	GBP2	2634	0,68	2,90E+08	TRUE
ILMN_1771385	GBP4	115361	0,71	5,99E+07	TRUE
ILMN_1748473	GIMAP4	55303	0,73	2,56E+07	TRUE
ILMN_1776678	GIMAP7	168537	0,77	1,84E+06	TRUE
ILMN_1779324	GZMA	3001	0,76	3,89E+06	TRUE
ILMN_1710734	GZMK	3003	0,75	6,02E+06	TRUE
ILMN_1693826	HAVCR2	84868	0,66	7,14E+08	TRUE
ILMN_1727402	HCLS1	3059	0,78	1,01E+06	TRUE
ILMN_1695311	HLA-DMA	3108	0,67	4,35E+08	TRUE
ILMN_1761733	HLA-DMB	3109	0,64	1,28E+09	TRUE
ILMN_1659075	HLA-DOA	3111	0,68	2,40E+08	TRUE
ILMN_1772218	HLA-DPA1	3113	0,65	8,55E+08	TRUE
ILMN_1808405	HLA-DQA1	3117	0,71	7,31E+07	TRUE
ILMN_1689655	HLA-DRA	3122	0,68	3,11E+08	TRUE
ILMN_2157441	HLA-DRA	3122	0,67	3,44E+08	TRUE
ILMN_2066066	HLA-DRB6	3128	0,64	1,68E+08	TRUE
ILMN_1708375	IRF1	3659	0,62	2,97E+09	TRUE
ILMN_1666594	IRF8	3394	0,76	3,45E+06	TRUE

ILMN_1813338	LAG3	3902	0,72	3,25E+06	TRUE
ILMN_1772359	LAPTM5	7805	0,70	8,83E+07	TRUE
ILMN_1662932	LCP1	3936	0,73	1,61E+07	TRUE
ILMN_1803788	LGALS3	3958	0,68	2,62E+08	TRUE
ILMN_3214389	HLA-DQB1	3119	0,67	4,59E+08	TRUE
ILMN_1721035	MS4A6A	64231	0,68	3,09E+08	TRUE
ILMN_1797731	MS4A6A	64231	0,73	2,25E+07	TRUE
ILMN_2331087	MS4A7	58475	0,73	1,68E+07	TRUE
ILMN_1682993	NKG7	4818	0,69	1,50E+08	TRUE
ILMN_1738675	PTPN6	5777	0,73	1,96E+07	TRUE
ILMN_1701613	RARRES3	5920	0,66	7,95E+08	TRUE
ILMN_1800787	RFTN1	23180	0,66	8,10E+08	TRUE
ILMN_1780533	RNASE6	6039	0,70	1,01E+08	TRUE
ILMN_1810275	SLC7A7	9056	0,74	1,14E+07	TRUE
ILMN_2087656	SLCO2B1	11309	0,70	8,40E+07	TRUE
ILMN_1690105	STAT1	6772	0,63	2,48E+09	TRUE
ILMN_3240316	TMSB4X	7114	0,73	1,82E+07	TRUE
ILMN_1764788	TNFRSF1B	7133	0,65	8,85E+08	TRUE
ILMN_1778977	TYROBP	7305	0,62	3,23E+09	TRUE
ILMN_1727271	WARS	7453	0,65	1,12E+09	TRUE
ILMN_1760027	WAS	7454	0,76	3,56E+06	TRUE

TUMOR SAMPLE	CLASS	[VOPP1]/[TTC5]	[VOPP1]/[TERT]	[TERT]/[TTC5]	[ΔB]/[TTC5]	[ΔB]/[VOPP1]	[ΔB]/[TERT]	T-CELL FRACTION	METHOD
01-042	I	0.979	1.040	0.941	0.969	0.989	1.029	2.1%	1
01-074	I	1.041	1.068	0.975	0.971	0.933	0.997	1.6%	2
01-091	IIb	1.019	1.040	0.979	0.935	0.918	0.955	7.3%	1
01-129	IIb	1.105	1.010	1.095	0.663	0.600	0.606	39.7%	3
01-131	IIa	1.003	1.076	0.932	0.922	0.919	0.989	7.9%	1
02-158	I	1.028	1.070	0.961	0.963	0.936	1.002	5.1%	1
02-167	I	1.027	1.029	0.998	0.987	0.961	0.988	1.2%	2
02-174	IIa	1.874	1.037	1.808	1.817	0.970	1.006	1.2%	3
02-199	I	0.976	1.125	0.867	0.943	0.967	1.088	4.5%	1
03-086	I	1.046	1.120	0.934	0.934	0.893	1.000	8.6%	1
03-087	I	1.033	0.757	1.364	1.000	0.968	0.733	1.6%	1
03-120	I	1.052	1.050	1.002	0.980	0.932	0.979	2.1%	2
03-129	I	0.986	1.015	0.972	0.901	0.913	0.927	9.3%	1
04-018	IIa	1.008	1.060	0.951	0.893	0.887	0.940	11.0%	1
04-035	IIb	1.000	0.970	1.031	0.750	0.750	0.728	25.0%	1
04-074	I	1.015	0.976	1.040	0.985	0.970	0.947	2.2%	1
04-075	I	1.057	1.011	1.045	0.956	0.905	0.915	9.0%	3
04-103	I	1.029	1.022	1.006	0.996	0.969	0.990	0.7%	2
04-112	IIb	1.062	0.982	1.081	0.902	0.850	0.835	15.8%	3
05-005	IIb	1.110	1.098	1.011	0.985	0.887	0.974	12.5%	4
05-020	I	1.066	1.050	1.015	0.917	0.860	0.903	9.0%	2
05-033	IIa	0.974	1.128	0.863	0.903	0.927	1.046	8.5%	1
05-034	IIa	1.042	1.106	0.943	0.942	0.903	0.999	7.8%	1
05-046	IIa	1.057	1.042	1.014	0.996	0.942	0.982	1.1%	2
05-058	I	0.960	0.960	1.000	0.930	0.969	0.930	7.0%	2
05-061	IIa	1.156	1.096	1.055	0.976	0.844	0.925	17.6%	4
06-004	IIb	1.125	1.042	1.080	0.948	0.843	0.878	14.0%	3
06-008	IIb	1.057	1.136	0.930	0.859	0.813	0.923	16.4%	1
06-009	IIa	1.027	1.030	0.997	1.022	0.995	1.025	-2.4%	2
06-010	I	1.020	0.996	1.024	0.986	0.967	0.963	3.5%	3
06-011	I	1.022	1.060	0.964	0.985	0.964	1.022	2.5%	1
06-014	IIb	1.072	1.127	0.951	0.566	0.528	0.595	41.9%	2
06-015	IIb	1.010	1.084	0.932	0.880	0.871	0.944	12.5%	1
06-023	IIb	1.084	1.038	1.044	0.828	0.764	0.793	22.2%	3
06-033	I	1.057	0.987	1.071	1.014	0.959	0.947	4.7%	3
06-036	I	1.527	1.528	0.999	1.460	0.957	1.462	6.6%	4
06-038	IIb	1.071	1.053	1.017	0.966	0.902	0.950	4.2%	2
06-041	IIb	1.045	1.095	0.955	0.982	0.939	1.028	4.0%	1
06-042	IIa	1.094	1.019	1.074	0.931	0.851	0.867	14.1%	3
06-045	IIb	1.077	1.129	0.954	0.802	0.745	0.841	28.2%	4
06-046	I	1.030	0.839	1.228	0.935	0.907	0.761	7.9%	1
06-047	IIb	1.099	1.127	0.975	0.952	0.866	0.976	14.9%	4
07-003	I	0.990	1.023	0.968	1.086	1.096	1.122	-9.1%	1
07-004	I	0.968	1.009	0.960	0.950	0.981	0.990	1.5%	3
07-005	IIa	1.109	4.060	0.273	0.680	0.613	2.488	n/a	5
07-007	I	1.052	1.000	1.052	0.972	0.924	0.924	7.6%	3
07-012	IIb	1.049	1.128	0.930	0.640	0.610	0.688	37.5%	1
07-030	IIa	1.036	1.160	0.893	0.933	0.901	1.045	8.3%	1

07-034	I	1.034	0.990	1.044	1.008	0.975	0.966	3.0%	3
07-047	IIa	1.045	1.073	0.974	0.907	0.868	0.931	8.1%	2
07-050	IIb	1.024	1.065	0.962	0.944	0.921	0.981	6.8%	1
08-004	IIb	1.093	1.106	0.988	0.878	0.803	0.888	21.7%	4
08-005	IIb	1.075	1.070	1.005	0.881	0.820	0.877	12.1%	2
08-008	IIb	1.028	1.048	0.981	0.932	0.907	0.950	5.9%	2
08-029	IIb	0.978	1.081	0.905	0.782	0.800	0.864	20.9%	1
20-005	IIa	1.031	1.084	0.951	0.962	0.934	1.012	5.2%	1
20-042	IIb	0.923	1.079	0.856	0.810	0.877	0.946	15.6%	1
20-125	I	0.984	1.092	0.901	1.003	1.019	1.112	-1.1%	1
20-128	IIb	1.112	1.130	0.984	0.908	0.816	0.922	20.6%	4
20-173	IIa	0.967	1.061	0.912	0.887	0.917	0.972	9.8%	1
20-178	IIb	1.137	1.059	1.073	1.079	0.949	1.005	2.3%	3
99-184	IIb	1.081	1.119	0.966	0.882	0.816	0.913	20.2%	4
99-187	I	0.985	1.117	0.881	0.995	1.011	1.129	-0.3%	1
99-239	IIb	1.105	1.003	1.102	1.049	0.949	0.952	5.0%	3

Legend

1: T-cell fraction calculation based on average $\frac{[\Delta B]}{[\text{REFERENCE}]}$ using reference genes VOPP1 and TTC5.

2: T-cell fraction calculation based on average $\frac{[\Delta B]}{[\text{REFERENCE}]}$ using reference genes TERT and TTC5.

3: T-cell fraction calculation based on average $\frac{[\Delta B]}{[\text{REFERENCE}]}$ using reference genes VOPP1 and TERT.

4: Adjusted T-cell fraction calculation based on average $\frac{[\Delta B]}{[\text{REFERENCE}]}$ using reference genes TERT and TTC5.

5: T-cell fraction not calculable.

Supl Table 3: amplicon sequence context

Assay	MiQE Context Sequence
TERT	CACCCCTTGGTGGCGGCTCACCTGTACGCCTGCAGCAGGAGGATCTTGTAGATGTTGGTGCACACCGTCTGGAGGCTGTTCACCTAGAGTCGCCAAGAAAGAGTGAGAAACGGTAGAAACCTC
TTC5	TGGTCGCGATGCCACTGTGGCAACAGCCTGGCTGCTGGATCCCTGAGGCT
	TCCCATTCACCACTAGCAGGAGGGGCGTCTCCACTCGAACACTGGAAAAG
	GAATAGTCCTAGAAAAGACAGAC
VOPP1	TATGGAGAGGGCCCGCACACAGCACCTGGAGCCACAGCAGTCCTCGTAGG
	AGCGGCATCTGTGGAGAGAGGCACAGGCTGGTCAGCACTGAATTGGAAGC
	AGCCACCGGACCAGCCATGCGGC
TRB	CCTGCCATCCTCTGCAGGCCATGCACTTTCCCTTTCGATGGACCCTCACAGAGGGAGCATCTGAATGGGGCATCCTTTGAAAAAGGAACCTAGGACCCTGTGGATGGACTCTGTCATTCTCCATG