Published in final edited form as:

Cancer Immunol Immunother. 2015 November ; 64(11): 1437–1447. doi:10.1007/s00262-015-1745-x.

Gene Expression Patterns in CD4+ Peripheral Blood Cells in Healthy Subjects and Stage IV Melanoma Patients

Sara J. Felts^{†,*}, Virginia P. Van Keulen^{†,*}, Adam D. Scheid[†], Kathleen S. Allen[†], Renee K. Bradshaw[¶], Jin Jen[†], Tobias Peikert^{†,‡}, Sumit Middha^{§,1}, Yuji Zhang^{§,2}, Matthew S. Block^{†,¶}, Svetomir N. Markovic^{†,¶,||}, and Larry R. Pease^{†,#}

[†]Department of Immunology, Mayo Clinic College of Medicine, Rochester, MN, USA

[‡]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Mayo Clinic College of Medicine, Rochester, MN, USA

[¶]Division of Biomedical Statistics and Informatics, Department of Health Sciences, Mayo Clinic College of Medicine, Rochester, MN, USA

[§]Department of Oncology, Mayo Clinic College of Medicine, Rochester, MN, USA

^{||}Division of Hematology, Department of Internal Medicine, Mayo Clinic College of Medicine, Rochester, MN, USA

Abstract

Melanoma patients exhibit changes in immune responsiveness in the local tumor environment, draining lymph nodes, and peripheral blood. Immune-targeting therapies are revolutionizing melanoma patient care as increasingly, studies show patients derive clinical benefit from these newer agents. Nonetheless, predicting which patients will benefit from these costly therapies remains a challenge. In an effort to capture individual differences in immune responsiveness, we are analyzing patterns of gene expression in human peripheral blood cells using RNAseq. Focusing on CD4+ peripheral blood cells, we describe multiple categories of immune regulating genes which are expressed in highly ordered patterns shared by cohorts of healthy subjects and stage IV melanoma patients. Despite displaying conservation in overall transcriptome structure, CD4+ peripheral blood cells from melanoma patients differ quantitatively from healthy subjects in

[#]Corresponding author: Department of Immunology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN, 55905 USA; Telephone: 507-284-8177; pease.larry@mayo.edu.

^{*}Equal contribution to the study

¹Present address: Memorial Sloan Kettering Cancer Center, New York, NY, USA

²Present address: Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA

AUTHOR CONTRIBUTIONS

SJ Felts and LR Pease designed the study, analyzed data, and wrote the paper. VP Van Keulen performed experiments, analyzed data, and wrote the paper. AD Scheid, and T Peikert analyzed data and wrote the paper. KS Allen and J Jen performed experiments. KS Allen, R Bradshaw, MS Block and SN Markovic recruited, consented, and clinically characterized the study subjects. S Middha and Y Zhang analyzed the data.

CONFLICT OF INTEREST

No author has a financial and/or commercial conflict of interest.

COMPLIANCE WITH ETHICAL STANDARDS

Written informed consent was obtained from all individual participants included in this study. All procedures involving human participants were in accordance with the ethical standards and approval of the Institutional Review Board (protocols #12-002580 and #13-002293) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

the expression of more than 2000 genes. Moreover, 1300 differentially expressed genes are found in transcript response patterns following activation of CD4+ cells ex vivo, suggesting widespread functional discrepancies differentiate the immune systems of healthy subjects and melanoma patients. While our analysis reveals that the transcriptome architecture characteristic of healthy subjects is maintained in cancer patients, the genes expressed differentially among individuals and across cohorts provide opportunities for understanding variable immune states as well as response potentials, thus establishing a foundation for predicting individual responses to stimuli such as immunotherapeutic agents.

PRÉCIS

The transcriptomes of freshly-isolated and stimulated CD4+ blood cells from normal subjects and cancer patients reveal common gene expression architecture and notable differences in immune status and immune response potential.

Keywords

Gene expression; Immune response; RNAseq; Transcriptome

INTRODUCTION

Attempts to mobilize the immune response to destroy disseminated cancer cells have met with variable success and are challenged by local [1, 2] and systemic immune suppressive effects [3, 4] of advancing cancers, degradation of the immune system by standard cancer therapies, and variability introduced by genetic and epigenetic determinants of host responsiveness [5–7]. In order to delineate how these variables impact the status of the immune system and immune response potential, we have begun assessing the underlying gene expression profiles of cells comprising the immune system. As these cells circulate throughout the body and survey the cellular and molecular environments of most tissues and draining lymph nodes, we began our studies by focusing on cells from peripheral blood. This material is accessible from most patient populations, and we assume that the reactive “state” of circulating cells is an indication of how those cells are poised to respond to stimuli encountered within an inflamed tissue or tumor.

Cutting edge technologies reveal the inner workings of immune cells and define molecular determinants of the human immune response [8]. Ability to measure these molecular determinants is fundamental to evaluating individual inflammatory and immune responses to environmental cues, yet traditional approaches focusing on antibody and cytokine production or the expression of select cell-associated molecules are not sufficient to provide informative answers. More comprehensive approaches such as characterizing the transcriptome of peripheral blood mononuclear cells have been stymied by the complexity of the cell types present in these mixed populations [9, 10]. Efforts to investigate factors differentiating closely related cell types are beginning to define pathways governing cellular differentiation [11–13] and define parameters of individual variability [14–16]. Technologies have advanced to the point where the transcriptome of individual cells can be analyzed systematically [17].

We have adopted an intermediate strategy by analyzing subsets of cells that share major lineage markers used to define cellular elements of the immune system. This approach has the advantage of substantially reducing complexity of the cell populations studied, while retaining some elements of complexity that might be important for determining immune response status or defining individual-specific characteristics of immunity. For example, finding that one individual had a different composition of CD4⁺ cellular subsets may be a determining trait of that individual. By quantitating the transcriptomes of 10⁶ cells, the response patterns observed are averages, which remove the cell-to-cell temporal variation in gene expression observed when analyzing single cells. Whereas recent studies investigating the transcriptomes of immune cells have used semi-quantitative gene array profiling [14–16], we used RNAseq [18, 19] to glean a more complete and complex look at the mRNA content of cells at the individual gene level. Recent comparisons of technical platforms for assessing global gene expression indicate that RNAseq is superior to hybridization-based approaches for measuring mRNA levels of genes with transcript levels below the 50th percentile [19]. We reason that many genes of importance may be expressed at these middling-to-lower levels, even in lineage-enriched cell populations, and thus chose to sequence transcriptomes of individuals at the depth of 37 million reads to capture accurately RNA present at levels as low as 1 read per million (RPM).

In this study we delineate structural features of the human CD4⁺ transcriptome. CD4⁺ peripheral blood lymphocytes (PBL) of T cell lineages (e.g. Th1, Th2, Th17, Treg, [20, 21]) and other less frequent cell types such as CD4⁺ macrophages [22] should be captured in this assay. We chose to focus initially on CD4⁺ cells (primarily T cells) because these cells not only provide helper and regulatory functions for CD8 T cells, but they also interact with and modulate the functions of many other immune cells [23]. Our thesis is that individual variation in CD4⁺ subpopulations is a trait detectable by the relative abundance of specific RNAs recovered from the lineage positive PBL, a view consistent with recent studies suggesting gene expression, as opposed to levels of regulation is the major factor determining phenotype [24]. At any given time, the variation in relative gene expression by lineage positive populations distinguishing individuals is expected to be the product of differences in genetic hard wiring and epigenetic changes stemming from biological rhythms and environmental stimuli. Here, we demonstrate RNAseq as a quantitative tool for human studies and document the remarkable homogeneity in mRNA expression patterns among both healthy individuals as well as patients with stage IV melanoma for genes specifying many shared functions of CD4⁺ T cells. Our data show that other populations of mRNA are more variably expressed and suggest that these more complex expression patterns can be used as a source for elucidating mechanisms of systemic immune dysfunction in melanoma and how immune therapies can be appropriately harnessed to improve patient outcomes.

MATERIALS AND METHODS

Study Subjects

Participant recruitment and blood sample collection were conducted using protocols approval by the Mayo Clinic Institutional Review Board (IRB#12-002580 and IRB#13-002293). Normal healthy subjects (7 females and 3 males; median age 34 years)

were from a Southeastern Minnesota community cohort of adults self-identifying as having no autoimmunity, allergy, immunodeficiency, immunosuppressive treatments, or cancer. The second cohort (5 females and 5 males; median age 60) was comprised of patients with a diagnosis of stage IV melanoma, but having different degrees of disease burden and a variety of treatment histories, planning to receive immune therapy (see Supplementary Table 1).

CD4+ T-Cell Isolation and Stimulation

Blood was drawn into Sodium Heparin tubes (APP Pharmaceuticals, NDC 63323-540-11); CD4+ cells were isolated from whole blood using anti-CD4 Microbeads (Miltenyi Biotec) and MS column purification as per manufacturer's instructions. A small portion of the cell isolate was set aside for flow cytometry using anti CD3 APC-H7 and anti CD4 PE-Cy7 (BD Pharmingen) to assess purity. Isolated cells were re-suspended at 1 million/ml in RPMI containing 10% FBS and either immediately lysed in 0.7 ml QIAzol lysis reagent (Qiagen) or incubated for 4 hours at 37°C with 25 µl/ml anti-CD3/anti-CD28 Human T-Activator Dynabeads (Invitrogen) prior to lysis. Lysates were stored at -80°C until processing for RNAseq.

RNA and cDNA Library Preparation

Samples with RNA integrity values of ≥ 7.0 by Agilent Bioanalyzer (Santa Clara, CA) were processed by the Mayo Clinic Medical Genome Facility Gene Expression Core into TruSeq libraries according to the manufacturer's instructions (RNA Prep Kit v2, Illumina, San Diego, CA) and published elsewhere [25]. Paired-end DNA adaptors (Illumina) with a single "T" base overhang at the 3' end were immediately ligated to the 'A tailed' cDNA population. Unique indexes, included in the standard TruSeq Kits (12-Set A and 12-Set B), were incorporated at the adaptor ligation step for multiplex sample loading on the flow cells. Libraries (8–10 pM) were loaded onto paired-end flow cells to generate cluster densities of 700,000/mm² following Illumina's standard protocol for cBot and cBot Paired end cluster kit version 3. The flow cells were sequenced as 51 X 2 paired-end reads on an Illumina HiSeq 2000 using TruSeq SBS sequencing kit version 3 and HCS v2.0.12 data collection software. Base-calling was performed using RTA version 1.17.21.3.

Bioinformatics Analysis

The Illumina standard pipeline version 1.4 was used to process raw images to make base calls and to generate sequence reads. The generated FASTQ sequence reads (51 bp long) were aligned to human genome assembly hg19 using Tophat aligner [26] with the default parameters. The raw read counts of each annotated gene in GENCODE v12 [27] were calculated by the HTSeq program (<http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>) for further downstream analysis. The initial reference genome selected is applied consistently throughout the analysis, even though updated references periodically become available. However, inferences about differences in gene expression for weakly annotated or previously unidentified loci from this presentation should be tempered until a deeper analysis of the distribution of the reads mapped to the genomic reference to verify that the correct match been the putative locus and mapped transcripts has been identified. For well-defined coding sequences, this is not an issue. For most analyses, the raw read counts of each gene were normalized by the total reads of each individual sample and then

standardized to reads per million (RPM, gene counts/total counts of each sample $\times 1$ million). Genes without any reads across all samples were eliminated from downstream analysis. As there were genes with zero count in some of samples, 1 was added to all gene counts in all samples after converting to RPM. Log₂-transformed RPM values were used for additional analyses. Non-transformed data were used to calculate fold differences from arithmetic means.

Statistical Analyses

The average (mean) and variation in mRNA expression across 10 individual subjects were determined and characterized using the coefficient of variation ($CV = STDEV/mean$) using Microsoft Excel software. Additional data analyses were conducted using Prism (GraphPad). The distributions of the CV measurements were assessed using the F test. The variances of the distributions of CV were found in some comparisons to be highly significantly different. Consequently, all subsequent comparisons were conducted using nonparametric rank analyses (Mann-Whitney or Kruskal–Wallis) as indicated where appropriate. A probability threshold of less than 0.05 was used to determine statistical significance. A gene-by-gene comparison of means and CV for the described gene sets was conducted by regression analysis in Prism. Two-way ANOVA to derive differentially expressed gene sets, stable between duplicate blood draws, were conducted using Partek [28].

RESULTS

Peripheral blood CD4⁺ cells were isolated directly from whole blood of healthy donors (HD, $n = 10$; Supplementary Table 1) and patients with stage IV melanoma (Mel, $n = 10$) using magnetic beads (Supplementary Figure 1). Typical enrichment efficiency (Supplementary Figure 2) for CD4⁺ cells was approximately 1000-fold. In order to identify similarities and differences in gene expression as a reflection of both immune state and response potential, RNA was extracted from freshly isolated (i.e., unstimulated, T0) cells, as well as cells placed into short-term culture (4 h in the presence of anti-CD3/anti-CD28 beads, T4) and analyzed by RNAseq. The number of individual mRNAs is reported as log₂ of the mapped reads per million reads plus 1 to account for zero reads in some samples ($\log_2 [RPM+1]$). Approximately 12,000 different mRNAs were present at greater than a value of 2 which is ~ 1 copy per cell, assuming a typical resting lymphocyte contains approximately 3.5×10^5 mRNA molecules [29, 30, 17]. Although sequences in this range could be present in only a fraction of the cells, they may represent subset-specific transcripts as CD4 T cells are known to differentiate into several functionally distinct lineages (e.g. Th1, Th2, Th17, Treg). We have also seen that previously unappreciated cell types, present at low frequencies ($< 1\%$), can have profound functional significance [31]. Consequently, the ability to measure mRNA species at low frequency is important for our approach.

We first sought to determine the frequency of mRNAs encoding gene products key to CD4⁺ T cell functions: (1) 58 genes encoding elements of the T cell receptor (CD3 + CD28) signaling pathway, (2) 41 genes encoding transcription factors regulating immune functions and polarization of CD4 T cells, (3) 17 genes encoding cytokines up-regulated more than 10-fold following CD3/CD28 ligation, (4) genes expressed at the highest levels (>400 RPM,

non-log2 transformed) in freshly isolated cells (~140 genes), and (5) genes expressed at the highest levels and up-regulated more than 10-fold following stimulation (~60 genes).

The T cell receptor and the CD28 costimulatory receptor signaling pathways comprise central mechanisms of intracellular communication used by CD4 T cells [20, 21]. Transcripts encoding these elements, common housekeeping genes, and approximately 5000 other annotated loci were present in CD4+ peripheral blood resting cells at levels ranging from 5.5–9.5 (Supplementary Table 2 and data not shown). A comparison of the average expression of each member of this gene set from cohorts of healthy donors (HD) and melanoma patients (Mel) revealed shared, highly ordered patterns of gene expression both before and after stimulation (Fig. 1a, d). Parallel analyses focusing on transcription factors (Supplementary Table 3) and the most highly expressed transcripts (present at greater than 400 reads per million ($\log_2 \sim 8.6$; ~150 transcripts, see Supplementary Table 4) revealed similar patterns (Fig. 1b, c, e). Among these most highly expressed mRNA are those encoding cell surface markers known to be expressed by CD4+ cells, including CD4, CD3e, and IL-10Ra; 12 of the 58 molecules associated with the TcR /CD28 signaling pathways were also in this category. Other transcripts in this list encode functions associated with the mechanics of translation, the cytoskeleton or energy metabolism and some have no known function.

To assess immune responsiveness, CD4+ PBL were stimulated for four hours ex vivo in medium containing beads coated with antibodies specific for CD3 epsilon and the co-stimulatory receptor CD28. The levels of cytokine-encoding mRNAs in stimulated cells were increased greater than 10-fold over resting levels (Fig. 2a, Supplementary Table 5), yet the stimulated levels of these cytokines was highly similar in cells from melanoma patients and healthy donors (Fig. 2b). Genes expressed at high levels after ex vivo stimulation and whose expression levels were increased at least 10-fold were also highly comparable between these cohorts (Fig. 2c; Supplementary Table 6).

We next considered the intra-cohort variability in gene expression, calculating coefficients of variation (CV) for the gene sets described above. In unstimulated CD4+PBL, variations in the genes of the TcR pathway from healthy donors and melanoma patients were relatively small, with median CVs of 0.019 and 0.015, respectively (Fig. 3a, b). The median CV of expression for transcription factors in freshly isolated cells from healthy subjects (T0, HD) was significantly higher than the median CV for the signaling molecules (Fig. 3a, $P = 0.0202$, Mann Whitney U). Much of this difference can be traced to variation in a subset of the transcription factors with CV values over 0.07, as the top quartile of the transcription factors has a median CV of 0.161 in comparison to the equivalent quartile of the TcR/CD28 signaling pathway with a median CV of 0.039 (Supplementary Tables 3 and 2, respectively). Included in this group are genes TBX21 and RORC encoding master regulators of the Th1 and Th17 response patterns, EOMES, a factor associated with terminally differentiated cells [32, 33], as well as heterodimeric regulators such as CREB5 and FOSB [34]. This disparity in CVs for genes associated with TcR signaling pathway (shared by all CD4+ T cells) and transcription factors in the CD4+ PBL also was observed in cells from the cohort of 10 melanoma patients (Fig. 3b, $P = 0.0048$). These data suggest that despite disease course and

treatment regimens endured by these individual cancer patients, the underlying architecture of CD4+ PBL gene expression pattern is preserved.

When data from stimulated cells were analyzed, cytokine gene expression showed a substantial variation among the healthy individuals in comparison to TcR pathway genes (Fig. 3c, median CV = 0.101 vs. 0.024, $P < 0.0001$, Mann-Whitney U). A parallel analysis of the transcripts in melanoma patient CD4+ PBL revealed a similar disparity between the CVs of TcR pathway elements and cytokine genes in the stimulated cells (Fig. 3d, $P < 0.0001$). These findings indicate that some key modulating factor(s) distinguishing individual responses impacts the variability in cytokine gene expression downstream of signaling molecules which are expressed less variably. Relatively low variability was observed for a set of 58 genes expressed at greater than 400 copies per million and up-regulated greater than 10-fold following activation (Fig. 3c, Supplementary Table 6; comparing median CVs of 0.101 vs 0.0250, $P < 0.05$, Kruskal-Wallis, Dunn's multiple comparison test). A very similar pattern was observed in our analysis of CD4+ PBL from melanoma patients (Fig. 3d, $P < 0.05$). These findings suggest that some genes that are strongly upregulated post activation are more variably up-regulated across individuals within cohorts than are others and that this variation of gene expression is not necessarily higher among highly expressed genes.

Next, we examined the transcriptomes from our cohort of healthy subjects for additional RNA transcripts expressed by CD4+ cells with higher variability, a subset of genes which might provide a basis for discriminating individual differences in immune function. Using a high threshold of CV > 0.1 , 1818 transcripts were expressed at greater than 2 copies per cell (RPM > 3.3) in fresh CD4+ PBL. Of those genes, ~40 have known roles in immune responses, yet were present in similar quantities and with similar CVs in HD and melanoma patient CD4+ cells (Supplementary Table 7). Similarly, ~400 transcripts present in HD CD4+ PBL following ex vivo stimulation were identified, 37 with known roles in immune responses (Supplementary Table 8). Again, overall levels and variability of expression in melanoma patients paralleled those in healthy donors.

To further investigate whether the patterns of conservation observed in healthy subjects and melanoma patients were interrelated and ordered, we performed a regression analysis comparing the CV for each gene in the sets encoding the CD3/CD28 signaling molecules, the transcription factors subset, and the induced cytokines before and after stimulation. By ordering the measures from healthy subjects (X-axis) and cancer patients (Y-axis) according to average CV (Supplementary Figure 3), a strongly correlated pattern for each comparison was revealed ($P < 0.0001$). The strongest correlations were for the TCR/CD28 signaling pathways and for transcription factor genes expressed after stimulation. We conclude that the transcriptomes of both healthy subjects and melanoma patients are highly ordered with a strong pattern specifying relative rankings in gene expression levels and patterns of variation.

Having established shared underlying architecture of the transcriptomes of the healthy and melanoma patient cohorts, we evaluated the data further for differences in gene expression patterns in a head-to-head comparison. To control for the variability in expression of some

genes irrespective of subject cohort, we analyzed RNAseq data for which we had duplicate, independent blood samples (10 HD and 7 Mel) using a two-way ANOVA. In fresh CD4+ PBL transcriptomes, approximately 6500 loci were found present at consistent levels for these 17 individuals (using a threshold of $P = 0.5$). In stimulated CD4+ PBL transcriptomes, that number was closer to 6000 loci. There were 2153 genes that differed between the cohorts at $P < 0.01$ (Table 1); 1413 genes (66%) were expressed at higher levels ($P < 0.01$) in cells from melanoma patients. For cells stimulated 4 h, this relationship reversed with 66% of the 1328 stable genes differentiating healthy and patient cohorts expressed at higher levels in cells from the healthy cohort (Table 1). Regression analyses of the mean expression values confirmed the previous conclusions that the transcriptomes of both fresh and stimulated CD4+ PBL from healthy subjects and metastatic melanoma patients were highly correlative (Fig. 4a, b). Nonetheless, more than 100 genes within the sets differentiating healthy subjects and melanoma patients at $P < 0.01$ exceeded a threshold of 2-fold difference (Fig. 4c, d). These patterns of differential gene expression were also observed using data from two additional healthy subjects and two additional melanoma patients where only single blood draw data were available (not shown).

The top 50 genes whose expression differed by at least two-fold between HD and Mel were ranked by P value and used to construct a heat map for both T0 and T4 data (Fig. 5a, b). Of note, the genes comprising this list derived from T0 data were over-expressed in the Mel CD4+PBL compared to HD cells whereas the genes derived from T4 data were under-expressed in the Mel cells compared to HD samples. As the ages of the subjects in our cohorts were not the same, additional analyses of an age-matched sample subset were performed, showing that the contribution of age to the gene expression patterns was not substantial (Fig. 5c, Supplementary Tables 9 and 10, and data not shown).

Finally, the cumulative data analyzed in this report were evaluated to determine whether the cohort of melanoma patients displayed greater variability in gene expression relative to the healthy subjects. The distributions of the evaluated CVs for the expression of the 236 genes in freshly isolated CD4+ PBL (Supplementary Tables 2, 3, and 4) and the 75 genes in stimulated PBL (Supplementary Tables 5 and 6) were found significantly more variable in the melanoma patients ($P = 0.015$ and $P = 0.002$, respectively by Mann-Whitney U analysis), perhaps reflecting differences in disease burden and treatment history.

DISCUSSION

Stage IV melanoma patients have been a targeted population for immunotherapy clinical trials. To investigate the status of the immune system of these advanced cancer patients, the structure and patterns of variability within the transcriptome of CD4+ PBL were compared to the transcriptomes of CD4+ PBL from healthy subjects. We used RNAseq to develop global assessments of immune cells and found this approach measures gene expression with precision, as indicated by the low coefficients of variations (<0.02) for 50% of the genes evaluated. This initial assessment of CD4+ PBL from very advanced cancer patients reveals that, despite having had different treatment histories and tumor burdens, the foundation of the CD4+ cell transcriptome from these patients is well conserved in terms of both transcript abundance and variability between members of the cohort. Moreover, the transcripts

dictating CD4 T cell functions, including those encoding the T cell receptor and associated signaling molecules are expressed very similarly in healthy individuals and patients with stage IV melanoma. Nonetheless, there are approximately 2000 genes whose expression levels are statistically different levels in the melanoma patients and healthy subject cohorts, even after controlling for variability between repeated blood draws. Comparison of age-matched subsets from these cohorts revealed similar findings. The variability in the expression levels of specific genes in cells from melanoma patients is greater in response to ex vivo stimulation compared to the variability of the same genes in cells from healthy subjects. We conclude that systemic changes in the immune status of melanoma patients are clearly evident and that these changes are reflected in altered immune responsiveness to strong stimuli.

Transcriptome analysis allows for simultaneous evaluation of cell type-specific molecules for their roles in dictating cellular functions and also as a way to seek out mechanisms contributing to individual variations in cellular (and here, immune) responses. Our data show that the most abundantly expressed mRNA molecules (more than 400 copies per 10^6 transcripts) identified from freshly isolated human CD4⁺ cells include CD4 itself, elements of the T cell receptor (TcR constant region and CD3 components), other cell surface leukocyte markers, kinases, adaptor molecules and transcription factors and that these basic immune response elements are largely intact in advanced melanoma patients. Moreover, the individual subject variability in the frequency of differentiated CD4 T cell subtypes (e.g. Th1, Th2, Th17, or Treg cells), based on mRNA levels and their variability for transcription factors that differentiate these cell types (e.g., T-bet (TBX21), GATA3, ROR γ t (RORC), FOXP3; [35]) were similar between PBL of healthy subjects and melanoma patients. However, transcripts encoding the AP1 components Fos, Jun, and JunB were repressed in freshly isolated CD4 PBL from melanoma patients relative to our healthy donors. Functional annotation clustering using DAVID Bioinformatics Resources [36, 37] revealed that many genes differentially expressed by cells from melanoma patients were enriched for inflammatory response genes and encoded membrane proteins involved in cell-cell communication, cell adhesion and migration, and molecule transport. How specific changes in gene expression relate to altered states of immune responsiveness will require additional study.

Our data also underscore how little is understood about the connections between genes expressed in specific types of cells. For example, stimulating CD4 T cells ex vivo for 4 hours resulted in ~2000 transcripts increased or decreased by at least a factor of 4. How the signals are elaborated from the engaged cell surface receptors to such a large network of genes in just a few hours is not understood. Furthermore, our data show that most of the central players identified in CD4 cells are maintained in late stage cancer patients and that those cells can respond to stimuli. How this information can be used to deduce critical points of dysregulation central to restoring homeostasis and focusing immunity against the cancer remains a challenge. Our RNAseq analysis reveals that many of the genes differentiating CD4 cells from healthy subjects and melanoma patients are expressed at relatively low levels which may be best addressed using sensitive quantitative assays such as RNAseq. We hypothesize that a gene expression signature from immune cells that has clinical consequence will be one that takes into account a summation of the contributions of several

hundred genes, each expressed at measurably different levels by subjects in healthy and cancer-burdened cohorts. In addition, the contribution of the transcriptomes from other PBL subtypes will likely be important.

During the course of this work, a study by Benoist and colleagues identified sources of individual variation of gene expression linking RNA expression to GWAS analyses and assessed their relationships to cellular activation and autoimmunity [16]. Our approach is different, focused on a summation of variation across many genes to create patterns that may identify sources of individuality in immune responsiveness. It will be of interest to determine to what extent the differentially-expressed or the variably-expressed gene subsets identified in our study correlate with other measures of differences in immune responses among individuals in various cohorts. Future studies will aim to determine which of these gene expression patterns can be resolved into predictive measures of how cancer patients respond to therapeutic interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Bruce Eckloff and staff of the Medical Genome Facility Sequencing Core at Mayo Clinic for the RNAseq services that are at the foundation of this study. This work was supported by the Mayo Clinic Center of Individualized Medicine and a Richard M Schulze Family Foundation Award in Cancer Research. Additional support came from the National Institutes of Health (NIH) T32 AI07425.

Abbreviations used

CV	coefficient of variation
HD	healthy donors
HGNC	HUGO Gene Nomenclature Committee
Mel	melanoma patients
PBL	peripheral blood leukocyte
RPM	reads per million
TcR	T-cell receptor

REFERENCES

1. Baitsch L, Baumgaertner P, Devereux E, Raghav SK, Legat A, Barba L, Wieckowski S, Bouzourene H, Deplancke B, Romero P, Rufer N, Speiser DE. Exhaustion of tumor-specific CD8(+) T cells in metastases from melanoma patients. *J Clin Invest*. 2011; 121(6):2350–2360. [PubMed: 21555851]
2. Siska PJ, Rathmell JC. T cell metabolic fitness in antitumor immunity. *Trends Immunol*. 2015; 36(4):257–263. [PubMed: 25773310]
3. Leontovich AA, Dronca RS, Suman VJ, Ashdown ML, Nevala WK, Thompson MA, Robinson A, Kottschade LA, Kaur JS, McWilliams RR, Ivanov LV, Croghan GA, Markovic SN. Fluctuation of

systemic immunity in melanoma and implications for timing of therapy. *Front Biosci (Elite edition)*. 2012; 4:958–975.

4. Wesa AK, Mandic M, Taylor JL, Moschos S, Kirkwood JM, Kwok WW, Finke JH, Storkus WJ. Circulating type-1 anti-tumor CD4(+) T cells are preferentially pro-apoptotic in cancer patients. *Front Oncol*. 2014; 4:266. [PubMed: 25325015]
5. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shores N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015; 518(7539):337–343. [PubMed: 25363779]
6. Huang E, Wells CA. The ground state of innate immune responsiveness is determined at the interface of genetic, epigenetic, and environmental influences. *J Immunol*. 2014; 193(1):13–19. [PubMed: 24951823]
7. Lim PS, Li J, Holloway AF, Rao S. Epigenetic regulation of inducible gene expression in the immune system. *Immunology*. 2013; 139(3):285–293. [PubMed: 23521628]
8. Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, Stichweh D, Blankenship D, Li L, Munagala I, Bennett L, Allantaz F, Mejias A, Ardura M, Kaizer E, Monnet L, Allman W, Randall H, Johnson D, Lanier A, Punaro M, Wittkowski KM, White P, Fay J, Klintmalm G, Ramilo O, Palucka AK, Banchereau J, Pascual V. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity*. 2008; 29(1):150–164. [PubMed: 18631455]
9. Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, Clark HF. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One*. 2009; 4(7):e6098. [PubMed: 19568420]
10. Lahdesmaki H, Shmulevich L, Dunmire V, Yli-Harja O, Zhang W. In silico microdissection of microarray data from heterogeneous cell populations. *BMC Bioinformatics*. 2005; 6:54. [PubMed: 15766384]
11. Hashimoto S, Nagai S, Sese J, Suzuki T, Obata A, Sato T, Toyoda N, Dong HY, Kurachi M, Nagahata T, Shizuno K, Morishita S, Matsushima K. Gene expression profile in human leukocytes. *Blood*. 2003; 101(9):3509–3513. [PubMed: 12522010]
12. Lyons PA, Koukoulaki M, Hatton A, Doggett K, Woffendin HB, Chaudhry AN, Smith KG. Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. *BMC Genomics*. 2007; 8:64. [PubMed: 17338817]
13. Palmer C, Diehn M, Alizadeh AA, Brown PO. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics*. 2006; 7:115. [PubMed: 16704732]
14. Gruden K, Hren M, Herman A, Blejec A, Albrecht T, Selbig J, Bauer C, Schuchardt J, Or-Guil M, Zupancic K, Svajger U, Stabuc B, Ihan A, Kopitar AN, Ravnkar M, Knezevic M, Rozman P, Jeras M. A "crossomics" study analysing variability of different components in peripheral blood of healthy caucasoid individuals. *PLoS One*. 2012; 7(1):e28761. [PubMed: 22253695]
15. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, Wang E, Olnes MJ, Narayanan M, Golding H, Moir S, Dickler HB, Perl S, Cheung F. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell*. 2014; 157(2):499–513. [PubMed: 24725414]
16. Ye CJ, Feng T, Kwon HK, Raj T, Wilson MT, Asinovski N, McCabe C, Lee MH, Frohlich I, Paik HI, Zaitlen N, Hacohen N, Stranger B, De Jager P, Mathis D, Regev A, Benoist C. Intersection of population variation and autoimmunity genetics in human T cell activation. *Science*. 2014; 345(6202):1254665. [PubMed: 25214635]
17. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014; 42(14):8845–8860. [PubMed: 25053837]
18. Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol*. 2011; 9:34. [PubMed: 21627854]
19. Wang C, Gong B, Bushel PR, Thierry-Mieg J, Thierry-Mieg D, Xu J, Fang H, Hong H, Shen J, Su Z, Meehan J, Li X, Yang L, Li H, Labaj PP, Kreil DP, Megherbi D, Gaj S, Caiment F, van Delft J, Kleinjans J, Scherer A, Devanarayan V, Wang J, Yang Y, Qian HR, Lancashire LJ, Bessarabova M, Nikolsky Y, Furlanello C, Chierici M, Albanese D, Jurman G, Riccadonna S, Filosi M, Visintainer

- R, Zhang KK, Li J, Hsieh JH, Svoboda DL, Fuscoe JC, Deng Y, Shi L, Paules RS, Auerbach SS, Tong W. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nat Biotechnol.* 2014; 32(9):926–932. [PubMed: 25150839]
20. Acuto O, Michel F. CD28-mediated co-stimulation: a quantitative support for TCR signalling. *Nat Rev Immunol.* 2003; 3(12):939–951. [PubMed: 14647476]
21. Huse M. The T-cell-receptor signaling network. *J Cell Sci.* 2009; 122(Pt 9):1269–1273. [PubMed: 19386893]
22. Szabo G, Miller CL, Kodys K. Antigen presentation by the CD4 positive monocyte subset. *J Leukoc Biol.* 1990; 47(2):111–120. [PubMed: 2303746]
23. Zhu J, Paul WE. CD4 T cells: fates, functions, and faults. *Blood.* 2008; 112(5):1557–1569. [PubMed: 18725574]
24. Schwanhaussier B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. *Nature.* 2011; 473(7347):337–342. [PubMed: 21593866]
25. Yamada D, Rizvi S, Razumilava N, Bronk SF, Davila JI, Champion MD, Borad MJ, Bezerra JA, Chen X, Gores GJ. IL-33 facilitates oncogene-induced cholangiocarcinoma in mice by an interleukin-6-sensitive mechanism. *Hepatology.* 2015; 61(5):1627–1642. [PubMed: 25580681]
26. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009; 25(9):1105–1111. [PubMed: 19289445]
27. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012; 22(9):1775–1789. [PubMed: 22955988]
28. Downey T. Analysis of a multifactor microarray study using Partek genomics solution. *Methods Enzymol.* 2006; 411:256–270. [PubMed: 16939794]
29. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, Linnarsson S. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 2011; 21(7):1160–1167. [PubMed: 21543516]
30. Jackson DA, Pombo A, Iborra F. The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. *Faseb J.* 2000; 14(2):242–254. [PubMed: 10657981]
31. Brickshawana A, Shapiro VS, Kita H, Pease LR. Lineage(-)Sca1+c-Kit(-)CD25+ cells are IL-33-responsive type 2 innate cells in the mouse bone marrow. *J Immunol.* 2011; 187(11):5795–5804. [PubMed: 22048767]
32. Knox JJ, Cosma GL, Betts MR, McLane LM. Characterization of T-bet and eomes in peripheral human immune cells. *Front Immunol.* 2014; 5:217. [PubMed: 24860576]
33. Ono C, Yu Z, Kasahara Y, Kikuchi Y, Ishii N, Tomita H. Fluorescently activated cell sorting followed by microarray profiling of helper T cell subtypes from human peripheral blood. *PLoS One.* 2014; 9(11):e111405. [PubMed: 25379667]
34. Hess J, Angel P, Schorpp-Kistner M. AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci.* 2004; 117(Pt 25):5965–5973. [PubMed: 15564374]
35. Evans CM, Jenner RG. Transcription factor interplay in T helper cell differentiation. *Brief Funct Genomics.* 2013; 12(6):499–511. [PubMed: 23878131]
36. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat Protoc.* 2009; 4(1):44–57. [PubMed: 19131956]
37. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009; 37(1):1–13. [PubMed: 19033363]

Fig. 1. Regression analysis of mean expression values (log₂[RPM+1]) for (a) genes encoding elements of the T cell receptor (TcR) signaling pathway, (b) genes encoding transcription factors typical of immune cells, and (c) highly expressed genes in freshly isolated CD4⁺ PBL of healthy donors (HD) and patients with stage IV melanoma (Mel). Data are available in Supplementary Tables 2-4. (d, e) Similar analysis of TcR pathway and transcription factor transcripts from stimulated CD4⁺ PBL. Data were calculated using Excel (Microsoft) and analyzed using Prism (GraphPad)

Fig. 2.

(a) Response to ex vivo stimulation assessed by quantitation of cytokine gene expression using RNA from unstimulated (T0) and stimulated (T4) CD4⁺ PBL from Healthy Donors (HD). Data are available in Supplementary Table 5. (b) Regression analysis of mean expression levels for cytokine genes and (c) other genes highly expressed and up-regulated 10-fold in stimulated CD4⁺ PBL from HD and Mel. Data are available in Supplementary Tables 5 and 6. Data were plotted and analyzed using Kruskal-Wallis and Dunn's Multiple Comparison tests using Prism

Fig. 3.

Coefficients of variation (CV) are shown for RNAseq quantitation of gene expression for known elements of the T cell receptor (TcR) signaling pathways and downstream transcription factors (TFs). RNA was from unstimulated (T0) CD4+ cells from (a) Healthy Donors (HD; n=10) and (b) patients with stage IV melanoma (Mel; n= 10; see Supplementary Tables 2 and 3). Data were plotted and analyzed using Mann-Whitney test in Prism (GraphPad). Each line indicates the median CV for each group. (c, d) Similar analyses for genes encoding TcR pathway and cytokines at T4 as well as highly expressed genes at

T4 induced 10-fold by simulation ex vivo. Data are available in Supplementary Tables 2, 5, and 6.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4.

Regression analysis of stable- and differentially-expressed genes in Mel compared to HD. (a) Comparison ($R^2 = 0.943$) of mean RPM values for ~2150 genes from fresh CD4+ PBL with stable expression values between duplicate blood draws ($P = 0.5$) and distinguishing HD and metastatic melanoma patients at $p < 0.01$. (b) Similar to (a) but with ~1300 genes from stimulated PBL ($R^2 = 0.938$). (c) 121 gene subset from (a) that differed between cohorts by at least 2-fold ($R^2 = 0.664$). (d) 113 gene subset from (b) that differed between cohorts by at least 2-fold ($R^2 = 0.601$).

Fig. 5.

(a) Heat map displaying 50 stably-expressed genes whose expression levels best differentiate (based on p-value) melanoma patients ($n = 7$) from healthy donors ($n = 10$). Expression levels of the genes are shown for each individual melanoma patient relative to healthy donor mean expression levels of those genes. Blue = gene expressed at lower RPM in Mel; Red = gene expressed at higher RPM in Mel. (b) Similar to (a) but with transcriptomes of stimulated (T4) CD4+ Mel ($n = 6$) relative to HD PBL ($n = 10$). (c) Similar to (a) and (b) but showing results of analysis using data only from aged-matched HD and Mel.

Table 1

Numbers of genes distinguishing the transcriptomes of CD4+ PBL from healthy donors and stage IV melanoma patients.

CD4+ PBL	Total stably-expressed genes distinguishing melanoma from healthy subjects over probability threshold (P<0.01)	Genes expressed at higher levels in melanoma patients (P<0.01)	Subset of Genes expressed > 2 fold higher in melanoma patients (P<0.01)	Subset of Genes expressed >2 fold higher in healthy subjects (P<0.01)
Fresh (T0)	2153	1413 (66%)	93 (4.3%)	28 (1.3%)
Stimulated (T4)	1328	449 (34%)	49 (3.7%)	64 (4.8%)