bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# *Phenotyping the human tumor-immune microenvironment (TiME) in vivo by cellular-level optical imaging reveals unique combination phenotypes with variable inflammation and endothelial anergy*

Aditi Sahu[1]*, Teguru Tembo[1,2], Kivanc Kose[1], Anthony Santella[1], Anabel Alfonso[1,3], Madison Li[1], Miguel Cordova[1], Melissa Gill[2,4], Christi Fox[5], Salvador Gonzalez[6], Amber Weiching Wang[7], Nicholas Kurtansky[1], Pratik Chandrani[8], Piyush Kumar[9], , Shen Yin[1], Haaris Jilani[3], Paras Mehta[1], Cristian Navarrete-Dechent[1,10], Gary Peterson[1], Kimeil King[1], Stephen Dusza[1], Ning Yang[1], Shuaitong Li[1], William Phillips[1], Anthony Rossi[1], Allan Halpern[1], Liang Deng[1], Melissa Pulitzer[1], Ashfaq Marghoob[1], Chih-Shan Jason Chen[1], Milind Rajadhyaksha[1]*

1. Memorial Sloan Kettering Cancer Center, New York, NY, United States. 2. SUNY Downstate Medical Center, Brooklyn, NY, United States. 3. Georgia Tech, Atlanta, GA, United States. 4. SkinMedical Research and Diagnostics PLLC, Dobby Ferry, NY United States, 5. Caliber Imaging and diagnostics, Rochester, NY, United States, 6. Alcala University, Madrid, Spain, 7. Genentech Inc, San Francisco, CA, United States, 8. Tata Memorial Hospital, Mumbai, India 9. Icahn School of Medicine at Mt. Sinai, New York, NY, United States 10. Pontificia Universidad Católica de Chile, Santiago, Chile.

**Corresponding authors:**

Aditi Sahu, Milind Rajadhyaksha
Dermatology Service
Memorial Sloan Kettering Cancer Center
530E, 74th street,
New York-10021.
Email: sahua@mskcc.org, aditisahu@gmail.com, rajadhym@mskcc.org

Word Count: 3045
Figures: 4
Tables: 1
Supplementary Figures: 8
Supplementary Table: 1
Supplementary Files: 3

**Conflicts of Interest:**

Melissa Gill: consulting investigator for DBV technologies; research consultant: Dermatology Service, MSKCC. Christi Alessi-Fox: employee of and owns equity in Caliber I.D., manufacturer of the VivaScope RCM. Dr. Rossi: Mavig (travel accommodation), Merz, DynaMed, Canfield Scientific, Evolus, Biofrontera, QuantiaMD, Lam Therapeutics, Cutera (consultant); Allergan (advisory board). Allan Halpern: consultant to Canfield Scientific and an advisory board member of Scibase. L.D. is a cofounder and holds equity in IMVAQ Therapeutics. She has patents on applications related to work on

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

oncolytic viral therapy. Ashfaq A. Marghoob: honorarium for dermoscopy lectures (3GEN), royalties for books/book chapters, dermoscopy equipment for testing, payment for organizing and lecturing (American Dermoscopy Meeting). Chih-Shan Jason Chen: research funding from Apollo Medical Optics, Inc. Milind Rajadhyaksha: was employee of and owns equity in Caliber I.D. VivaScope is the commercial version of a laboratory prototype he developed at Massachusetts General Hospital, Harvard Medical School.

**Acknowledgement:**

Dr. Anjali Rajadhyaksha for assistance with experimental planning and scientific discussion. We would like to acknowledge MSKCC Cores: Integrated Genomics Operation Core, funded by the NCI Cancer Center Support Grant (CCSG, P30 CA08748), Cycle for Survival, and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Molecular Cytology Core Facility, Pathology Core, Flow Cytometry Core; funding sources NIH/NCI Cancer Center Support Grant P30 CA008748 and NCI/NIBIB R01EB020029, Melanoma Research Alliance (Aditi Sahu) and the Chan-Zuckerberg Initiative (Anthony Santella). Ms. Cassidy Cobbs, Mr. Eric Chan for assistance with experimental planning and data analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Abstract:**

Immunotherapies, especially immune checkpoint blockade therapy have shown unprecedented clinical benefits in several malignancies, however, responses are variable emphasizing the need for effective biomarkers for patient stratification[1]. Phenotyping of tumors into hot, altered, or cold[2] based on T-lymphocyte infiltration in tumor biopsies fails to explain and/or predict response to immunotherapy seen in a subset of patients[3,4]. One of the primary reasons for this suboptimal prediction by a single immune marker could be attributed to the fact that additional mechanisms within the tumor microenvironment modulate anti-tumor immunity and outcomes, including dynamic events such as tumor-angiogenesis and leukocyte trafficking[2,5,6]. We report novel tumor phenotypes through non-invasive spatially-resolved cellular-level analysis of the tumor immune microenvironment (TiME) and major determinants of anti-tumor immunity. Using skin cancers as a model and optical imaging using reflectance confocal microscopy (RCM)[7], we determined four major phenotypes based on unsupervised clustering for relative prevalence of vasculature (Vasc) and inflammation (Inf) features: $Vasc^{high}Inf^{high}$, $Vasc^{high}Inf^{low}$, $Vasc^{low}Inf(intratumoral)^{high}$ and $Vasc^{mod}Inf^{low}$. The $Vasc^{high}Inf^{high}$ phenotype correlate with high immune and vascular signatures while $Vasc^{high}Inf^{low}$ with endothelial anergy. Automated quantification of TiME features demonstrates moderate accuracy and high correlation with corresponding gene expression. Prospective testing of TiME features prior to topical immunotherapy response shows highest response in the $Vasc^{low}Inf(IT)^{high}$ phenotype, and revealing the added value of vascular features in predicting treatment response. This novel *in vivo* phenotyping combining dynamic immune and vascular features has the potential to advance fundamental understanding of the highly dynamic TiME, identify novel druggable pathways and develop robust predictors for immunotherapy outcomes.

**Main:**

Immunotherapy, especially immune checkpoint blockade therapy, has revolutionized cancer management by providing near durable responses in several cancers. However, only a subset of patients derives clinical benefit, highlighting a clinical need to develop effective biomarkers for patient stratification[1,8]. Phenotyping of tumors into hot, cold or altered based on quantifying T-lymphocyte infiltration at the tumor center and margin, along with PD-L1 expression on immunohistochemistry (IHC) and tumor mutation burden are important determinants for immunotherapy in solid cancers[9,10]. Although hot versus cold tumor phenotyping has shown some association with response, clinical response is not assured in the inflamed phenotypes, suggesting immune-cell infiltration is necessary but insufficient for inducing anti-tumor immunity[4,11]. Thus, tumors likely use additional mechanisms for evading immune response while establishing an immune-suppressive microenvironment, further complicating patient stratification strategies through dynamic tumor/host immune interactions and baseline tumor biology[9,12,13].

Tumor vasculature (both blood and lymphatic vessels) serve important immunomodulatory roles, and contribute to the immune evasion of tumors[14]. Angiogenesis promotes immune evasion through induction of a highly immunosuppressive TiME by inhibiting dendritic cell (DC) maturation, inhibiting T-cell development and function, and very importantly, limiting access of effector immune cells[15] to tumors by modulating leukocyte trafficking. In addition, tumor vasculature can display decreased expression of adhesion molecules, and non-responsiveness to inflammatory cytokines leading to

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

endothelial anergy. By downregulating trafficking of effector immune cells, vascular endothelial anergy contributes to ineffective anti-tumor immune responses and immune evasion[16-18].

Towards addressing the dynamic, complex and highly interdependent vascular-inflammation axis inside the TiME, *in vivo* phenotyping based on a combination of dynamic vascular and immune features, rather than *ex vivo* phenotyping based on static pathological evaluation of tumor infiltrating cells may facilitate a deeper understanding and achieve higher predictive power for patient stratification for immunotherapies. High-resolution non-invasive *in vivo* imaging is fundamental to this combination phenotyping, since static *ex vivo* analyses on patient tissue are limited in recapitulating dynamic vascular and immune attributes[19]. We report novel combination phenotypes detected *in vivo* using reflectance confocal microscopic (RCM) imaging. RCM is a high-speed (pixel times ~ 0.10 µsec, frame rates 10-30 per second) cellular-level label-free imaging approach based on backscattered light and endogenous tissue contrast[7,20]. Large image mosaics (64 $mm^2$ in 50 seconds) imaged to a depth of ~0.25 mm enables spatial resolution of TiME features. RCM is routinely used for real-time skin cancer diagnosis and management at the bedside. Few studies have demonstrated RCM imaging of vessels and leukocyte trafficking in humans[21,22]. Using skin cancer as a model, we define novel combination TiME phenotypes using six features: vessel diameter, vessel density, leukocyte trafficking, intratumoral inflammation, peritumoral inflammation and perivascular inflammation (Fig S1) and report their subsequent molecular correlation with inflammatory, angiogenic and trafficking signatures. We also and, in a prospective pilot study, examine their investigated relative importance of TiME features in predicting response to topical immunotherapy.

**Results:**

Towards *in vivo* combination tumor phenotyping, agreement between RCM feature evaluation (Fig S1) by two independent readers was investigated, and correlated with same features on well-validated and gold-standard histopathology by a board-certified dermatopathologist (Table 1). Substantial to almost perfect agreement (k=0.62-1.0) was observed for most RCM features. Good to very good agreement ($AC_1$: 0.74-1.0) was found between histopathology and average RCM evaluation, confirming the visualization of validated histopathological TiME features on RCM. Unsupervised clustering on these TiME features using Principal Component Analysis (PCA) on 27 skin cancer patients revealed four major phenotypes (Fig 1A): $Vasc^{high}Inf^{high}$, $Vasc^{high}Inf^{low}$, $Vasc^{low}Inf(IT)^{high}$ and $Vasc^{mod}Inf^{low}$ based on the PC loadings and vectors indicated in green arrows (Fig S2A). The largest loadings (PC1: dilated vessels, trafficking and PC2: peritumoral, perivascular inflammation) were used to establish the phenotypes on 27 patients. Patients between the PC1 and PC2 loading vectors were denoted as $Vasc^{high}Inf^{high}$ (red), patients proximal to the PC1 loadings were denoted $Vasc^{high}Inf^{low}$ (blue) while distal to PC1 loadings were $Vasc^{mod}Inf^{low}$ (light blue), and patients distal to main PC2 loadings but proximal to intratumoral (IT) inflammation loading vector were denoted as $Vasc^{low}Inf(IT)^{high}$. For phenotyping, the 6 RCM features used for unsupervised clustering were grouped into Vasculature (dilated vessels, number of vessels and trafficking) and Inflammation (peritumoral, perivascular and intratumoral inflammation). The 'high' and 'low' indicate the prevalence of the feature in the phenotype, high corresponds to manual score of 2 or 3, low to 0 or 1 (on a scale of 0-3, Fig S1D) while 'mod' denotes moderate feature prevalence (between 1-2). The *in vivo* phenotypes were correlated with the total number of $CD3^+$ T cells (to assess with respect to T-cell based pathological phenotyping[2]) and density of tertiary lymphoid structures (Fig S3), a recent hallmark

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of an inflamed micro-environment and positive cancer outcomes[23] on 27 patients (Fig 1A). The largest number of T-cells were seen in the $Vasc^{low}Inf(IT)^{high}$ phenotype (mean: 783 cells/mm$^2$, range: 148-1489) as compared to $Vasc^{high}Inf^{high}$ (407 cells/mm$^2$, range: 69-704), $Vasc^{mod}Inf^{low}$ (429 cells/mm$^2$, range: 236-623) and $Vasc^{high}Inf^{low}$ (371 cells/mm$^2$, range: 81-803). Higher TLS density was found in $Vasc^{low}Inf(IT)^{high}$ phenotype (mean: 0.037 mm$^2$, range:0-0.155) followed by $Vasc^{mod}Inf^{low}$ (0.037 mm$^2$, range: 0.01-0.006) while comparable densities were found in $Vasc^{high}Inf^{high}$ (0.01 mm$^2$, range:0-0.04), and $Vasc^{high}Inf^{low}$ (0.01 mm$^2$, range:0-0.06).

Transcriptomic analysis on 14 patient samples harboring sufficient RNA quality and quantity for bulk RNA-sequencing (RNA-seq QC in Fig S4A) was performed to enable phenotype correlation with gene expression profiling. The RCM tumor phenotyping on this subset of patients reveals similar phenotypes, with clustering driven mainly by trafficking (PC2) and intratumoral trafficking (PC2) (Fig 1B).The phenotyping was correlated with gene expression of CD3E (T-cells) to mimic the comparison with $CD3^+$ T cells on immunohistochemistry. Highest CD3E expression (7.7, 2.5-11 transcripts/million) was seen in $Vasc^{high}Inf^{high}$ while lowest CD3E expression (0.9, 0.6-1.1 transcripts/million) was seen in $Vasc^{high}Inf^{low}$ phenotype (Fig 1B). Unsupervised clustering to explore similar phenotypes in gene expression data was also explored using PCA for inflammation, angiogenesis[24] and trafficking gene signatures. Similar tendency of classification (Fig 1C), especially between the red/pink and blue/light-blue phenotypes was observed in inflammation (driven mainly by TNFAIP2 expression), angiogenesis (mainly SPARCL1) and trafficking (mainly CXCL12). Other major loadings or determinants driving classification in individual PCA included IL10RA, CD68, CD2, IFNGR1, JAML (inflammation), RGS5, CLEC3B, EDNRB, A2M and PDGFD (angiogenesis), Cav1, CD99, CXCL14, CCL 21 (trafficking) (Fig S2D-F).

Next, we performed extensive group-level transcriptomic analysis using CIBERSORT[25], DGEA and pathway analysis, gene set enrichment analysis (GSEA) and validated DGEA results using hierarchical cluster analysis (HCA). Owing to the small sample size, we merged the four phenotypes (Fig 1B) into two groups: $Inf^{high}$ (red) and $Vasc^{mod/high}$ (blue) (Fig 2A) to characterize molecular attributes specific to the prominent inflammation or vasculature groups. Subsequently, we performed additional unsupervised k-means clustering on CIBERSORT output (Fig 2B, Fig S4B) that two distinct $Inf^{high}$ and $Vasc^{mod/high}$ similar to RCM (Fig 2A), with one misclassified patient in each cluster (P21 and P22). DGEA on the CIBERSORT output on the $Inf^{high}$ and $Vasc^{mod/high}$ phenotypes found 11 significant differentially expressed genes (DEGs) with upregulated JAK-STAT signaling, NK-cell medicated cytotoxicity and chemokine signaling in the red cluster. Relative immune cell proportions in the red and blue clusters indicated higher CD4 memory resting, CD4 memory active and M1, M2 macrophages in the red cluster (Fig S4C, D). DGEA on the entire transcriptome guided by the RCM TiME phenotypes ($Inf^{high}$ and $Vasc^{mod/high}$) found 114 differentially expressed genes that separated RCM phenotypes into the same 2 groups with hierarchical cluster analysis (HCA) (Fig 4F), 85 genes were overexpressed in the $Inf^{high}$ while 29 genes in the $Vasc^{mod/high}$ cluster (Fig 2C). Pathway analysis using available gene sets (MSigDB_Hallmark_2020, NCI_human, WikiPathways_2019_Human, Bioplanet_2019) and GO biological processes demonstrate enrichment of mainly pro-inflammatory and anti-tumor genes in $Inf^{high}$ while enrichment of leukocyte-endothelial interactions, ephrin B2 pathway, vasodilation and neovascularization in the $Vasc^{mod/high}$ based on selected pathways (Fig 2D, Fig S6). Gene set enrichment

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

analysis (GSEA) for immune and angiogenesis signatures suggest higher scores in the $Inf^{high}$ cluster suggesting a higher immune activation and vascular enrichment (Fig 2E). HCA on Nanostring Pancancer Immune Panel Genes[26] (Fig S5E-F) indicates presence of four phenotypes of which P7, P14 and P17 (RCM $Vasc^{mod/high}$) were found to have comparatively lower expression of immune genes as compared to P5, P6, P21, P23, P36 (RCM $Inf^{high}$). These results suggested $Inf^{high}$ may have more inflamed phenotypic attributes while $Vasc^{mod/high}$ phenotype were associated with lower immune densities and immune activation in presence of high vessels and/or trafficking, suggestive of endothelial anergy[18]. After successfully verifying tumor phenotypes on RCM and gene expression datasets, we also investigated tumor phenotyping on 3 GEO datasets[27-29] using PCA and differential gene expression analysis (DGEA) towards analyzing similar phenotypic trends in additional diverse cohorts (Fig S5).

To characterize patients within each *in vivo* phenotype in terms of specific inflammation, vascular (angiogenesis, trafficking, endothelial anergy) and tumor intrinsic pathways driving the phenotypic classification, we investigated genes from critical pathways within each of these TiME components. All major inflammation genes were systematically upregulated in the $Vasc^{high}Inf^{high}$ while downregulated in the $Vasc^{high}Inf^{low}$ phenotype (Fig 3A). However, the $Vasc^{high}Inf^{low}$ phenotype was characterized by the highest M2/M1 macrophage ratio (from CIBERSORT output). While the $Vasc^{mod}Inf^{low}$ showed comparable immune expression for most genes with the $Vasc^{low}Inf(IT)^{high}$ phenotype, we found lower CD86 and lower CD8, GZMA/B expression. Both $Vasc^{high}Inf^{high}$ and $Vasc^{low}Inf(IT)^{high}$ showed similar degree of immune-inhibition (FOXP3, IDO1, TIGIT, CD274), higher than the remaining two phenotypes. In terms of vascular features, $Vasc^{high}Inf^{high}$ as compared to $Vasc^{high}Inf^{low}$ exhibited higher or similar expression of angiogenic genes such as CAV1, CAV2, VEGF-A and ALCAM (Fig 3B). We found highest VEGF-A expression in 1 patient (P17) in the $Vasc^{high}Inf^{low}$; most patients in this phenotype demonstrated higher VEGF-D and endothelin-2 expression than other phenotypes. $Vasc^{high}Inf^{high}$ and $Vasc^{low}Inf(IT)^{high}$ showed the highest VCAM, ICAM1, ICAM2, SEL-L, CXCL12 and CXCL9 levels. Relatively higher (or comparable to $Vasc^{high}Inf^{high}$) expression of ITGA3, FUT4 was observed in the $Vasc^{high}Inf^{low}$ phenotype (Fig 3C). We also investigated tumor intrinsic pathway differences and found relatively higher β-catenin, PTEN, TP53 and COX11 and lower STAT1, NF-κB1/2 and TLR7 expression in the $Vasc^{high}Inf^{low}$ phenotype (Fig 3D). Immunohistochemical correlation of phenotypes with $CD3^{+}$ T-cells, $CD20^{+}$ B-cells and TLS number/density found least immune cells and absence of tertiary lymphoid structures in $Vasc^{high}Inf^{low}$ phenotype (Fig 3E), additionally confirming the gene expression distribution pattern of immune cells and anti-tumor mechanisms in these phenotypes.

Of thirteen patients who underwent prospective evaluation of TiME features for correlation of novel phenotypes with response to topical immunotherapy, 7 patients responded to the immunotherapy treatment while 6 were non-responders. Within this pilot study cohort, most responders (5 out of 7) belonged to the $Vasc^{low}Inf(IT)^{high}$ phenotype (Fig 3F). Additional RCM features were investigated for modeling TiME features and tumor phenotypes with response to treatment. Higher frequency of leukocyte trafficking, greater number of stromal vessels and stromal macrophages were present in 50%, 100% and 86% of the non-responders. Linear regression models for predicting responders and non-responders demonstrate low predictive power of inflammation as a variable, either as "tumor-infiltrating lymphocytes" or "intratumoral inflammation"

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with accuracy of 46% and 61%, respectively (Fig 3F). The same regression models were used for identifying predictor variables from the TiME features using specificity or AIC coefficient as outcome measures. The best performance among the AIC prioritizing models was 85% sensitivity, 66% specificity and AIC=-15.06 with 8 variables while the best performance among the specificity prioritizing model was 71% sensitivity, 83% specificity and AIC= -15.06 with 13 variables (Fig S6). Three of eight features (stromal vessels, peritumoral vessels and stromal macrophages) overlapped between the two models. Linear separability plots also confirmed stromal vessels and stromal macrophages variables led to high separation between responders and non-responders. Addition of stromal vessels to intratumoral inflammation or tumor-infiltrating lymphocytes (Fig 3F) as features in the linear regression model resulted in best model performance (71% sensitivity, 83% specificity and 76% accuracy).

Finally, to enable objective quantitative comparisons between RCM Time features and phenotypes, immunohistochemical markers and gene expression, we investigated automated quantification of RCM TiME features— immune cells, leukocyte trafficking and vascular features using machine learning and image processing. Quantification of immune cell density was explored using a machine learning segmentation model, U-Net. Representative images and segmentations are shown in Fig S7. Image-processing was used for quantification of vascular features (vessel area, diameter and number) and leukocyte trafficking; vessel segmentation was performed using statistical filtering and component analysis-based algorithms while trafficking quantification was explored on a custom pipeline involving TrackMate (FIJI)[30]. Each analysis was validated on a subset of data (described in Methods), the results are summarized in Fig 4A. Correlation between manual evaluation and automated quantification demonstrated strong correlation of automated leukocyte area (green, Fig S7A) with inflammation (r=0.64), moderate correlation of automated total inflammation with inflammation (r=0.49) and weak correlation for trafficking (r=0.39) and vessel diameter with dilated vessels (0.29). Subsequently, correlation of RCM TiME features with corresponding gene expression show strong correlations between RCM total inflammation area with CSF1R (macrophage, r=0.73), CD1E (dendritic cells, r=0.64) and CD3E (lymphocytes, r=0.51), between total leukocyte area and CD8B (r=0.6) and GZMA (r=0.53). Vessel diameter correlated well with VEGF-D expression (r=0.45), PDGFD (r=0.538) and negatively corelated with VEGF-A (r=-.477) while leukocyte trafficking correlated positively with CCL18 (r=0.56), CAV-1 (r=0.468) and negatively with CCL25 (r=-0.771) (Fig 4D).

**Discussion and Conclusion:**

Predicting response/resistance immunotherapy as well as understanding how tumors escape immunity can facilitate effective treatment strategies. Currently known predictive mechanisms for immunotherapy including the IFN-G gene panel and the "Tumor Inflammation Signature" have shown some promise in predicting response to immune checkpoint blockade[31,32]. A more comprehensive and quantitative analysis of major determinants of anti-tumor immunity within the TiME will improve patient stratification, which will ultimately improve current cancer management paradigms and introduce a more personalized immunotherapy platform, similar to the IMPACT™[33]. Dynamic, *in vivo* noninvasive imaging is crucial for studying these interactions within TiME, since active vascular processes such as leukocyte trafficking are optimally studied as a live dynamic process, and *ex vivo* tissue studies on vasculature have shown inconsistent vessel measurements[19]. We

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

demonstrate the proof-of-concept for unperturbed characterization and tumor phenotyping inside patients. Furthermore, feasibility of automated quantification to generate stronger quantitative correlations was also demonstrated.

Our preliminary studies on a skin cancer cohort demonstrate presence of four unique *in vivo* tumor phenotypes based on their cumulative vascular and inflammatory attributes: $Vasc^{High}Inf^{Low}$, $Vasc^{high}Inf^{low}$, $Vasc^{low}Inf(IT)^{high}$ and $Vasc^{mod}Inf^{low}$ where $Vasc^{high}Inf^{high}$ demonstrates high immune and angiogenic signatures, including features characteristic of inflamed phenotypes such as CXCL9 (T-cell recruitment inside tumors)[34], induction of suppressor pathways PD-L, FOXP3, IDO-1 and exhaustion markers TIGIT[35] and higher memory cells (associated with improved overall survival[36]). This phenotype likely had normalized vasculature, as seen in high expression of VCAM1, ICAM1, L-selectin, CCL2 and higher intratumoral inflammation[14,16]. Conversely, $Vasc^{high}Inf^{low}$ show features characteristic of endothelial anergy and immunosuppression (poor immune infiltration in tumors) , including downregulation of major adhesion molecules (VCAM, ICAM1, ICAM2, SELL), higher VEGFD, relatively higher EDN2 and ITGA3 and higher tumor-intrinsic factors known to induce immunosuppression (CTNNB1, PTEN, COX11)[37,38]. VEGFD is implicated in blood and lymph vessel dilation and shows association with dilated vessels[39], genes such as ITGA3, CCL-28, CAV-1 and EDN2 may likely be compensating for decreased adhesion molecule expression in patients with higher trafficking[40]. Higher CCL28 correlation was seen in the $Vasc^{high}$ phenotype with high trafficking, suggesting the possible role of CCL-28 in trafficking of T-regulatory cells. Additionally, the upregulation of WNT/β-catenin pathway is one of the major factors associated with immune exclusion[37] and could be responsible for the endothelial anergy and immune suppression in the $Vasc^{high}Inf^{low}$ phenotype. Alternatively, $Vasc^{low}Inf(IT)^{high}$ corresponded to a highly inflamed phenotype with lower immunosuppressive vascular features, also reflected in the prospective imiquimod study where highest proportion of responders were found to belong to the $Vasc^{low}Inf(IT)^{high}$ phenotype. Relatively higher expression of TLR7 (agonist of Imiquimod) in this phenotype could also explain higher Imiquimod response. Additionally, immunophenotyping to correlate RCM phenotypes in a pilot study on 3 BCC tumors indicated higher activated $CD8^{+}GZMb^{+}$ and $CD8^{+}Ki\text{-}67^{+}$ cells in a patient with features of $Inf^{high}$ as compared to a patient with $Vasc^{high}$ also suggesting $Inf^{high}$ phenotype are inflamed (Fig S8).

Thus, these results suggest the feasibility of identifying not just the TiME phenotypes, but the mechanism of immunosuppression that can be exploited for treatment of cold, non-inflamed and non-responsive tumors. Anti-angiogenic agents can overcome endothelial cell anergy and reinduce EAM expression, resulting in increased leukocyte infiltration into tumours. For example, phenotypes with the $Vasc^{high}Inf^{low}$ phenotype could benefit from additional synchronous treatment for vessel normalization or pharmacological targeting of the WNT/β-catenin pathway or using anti-angiogenic topical treatments (COX-2, bFGF inhibitors) towards treatment optimization[41]. Additionally, immune-cell infiltration into tumors does not necessarily warrant response, thus turning excluded anergized tumors into inflamed phenotypes may fail as a treatment strategy. As an alternative, longitudinal and non-invasive monitoring of the treatment-induced alterations in the TiME phenotype for such newer TiME targeting therapies can help assess response and resistance mechanisms and further treatment optimization.

While this study demonstrates a combination of high-resolution spatially resolved and dynamic imaging, limitations included grayscale-limited specificity tissue contrast and imaging depth to 0.2-0.25 mm. The label-free approach enables

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

visualization of all TiME features, but is limited in specificity for functional phenotyping. This potential limitation may be overcome by accounting for the longitudinal spatio-temporal distributions in patients to monitor immune changes. The limited depth of imaging fails to capture deeper TiME features. With the current state of RCM devices and technology, this approach is currently restricted to accessible diseases and cancers, namely skin cancer, head-neck cancer, cervical cancers, cutaneous lymphomas, cutaneous metastasis. In the future, extensive validation with targeted molecular correlations on precision biopsies[42] will enable better correlations. Exhaustive molecular validation using flow cytometry and single-cell RNA-seq and spatial transcriptomics[43,44] on subsequent models will facilitate improved understanding of RCM phenotyping. Complementary multimodal approaches[45] such as dynamic optical coherence tomography (OCT, for deeper imaging of blood vessels[46]), multiphoton microscopy (MPM, for better contrast and collagen delineation[47]), photoacoustic microscopy (PAM, for deeper vessels with oxygen saturation/desaturation) and fluorescence lifetime imaging (FLIM, for immune cell specificity and activation states[48]) may be necessary to further enhance *in vivo* TiME visualization and enhance current TiME phenotyping. Through robust prospective studies, fundamental basis of phenotyping and their correlation with variable treatment responses in cancer immunotherapy systems will be explored for better patient stratification. These initial findings will enable hypothesis-driven research for developing novel druggable targets and gaining mechanistic insights regarding host anti-tumor immune response in various bedside cancer settings in human patients.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Methods:**

**Methods:**

***Patient recruitment and imaging***

Patients referred for physician consultation, Mohs surgery or wide local excision at Memorial Sloan Kettering Cancer Center (MSKCC), NY were prospectively enrolled for this study under MSKCC-IRB approved protocols. Patients (aged 18 or older) with either a previously biopsied or clinically suspected keratinocytic [basal cell carcinoma (BCC), squamous cell carcinoma (SCC), actinic keratosis(AK)] , melanocytic (melanoma) lesion or drug rash were accrued consecutively at Memorial Sloan Kettering Cancer Center (MSKCC) after written informed consent. Patients with suspected BCC selected for topical immunotherapy Imiquimod (n=9) were also enrolled.

***In vivo imaging***

In vivo RCM imaging was performed prospectively on 97 lesions using an RCM (either VivaScope 1500 or a handheld VivaScope 3000, Caliber I.D., Rochester, NY) and/or an integrated handheld RCM-OCT prototype. Images were acquired and interpreted in real-time at the bedside to select representative areas with tumor, immune cells and blood vessels across the lesion by 2 investigators (M.C. and A.S.) having more than 4 years of RCM reading experience. Mosaics (large area sampling), stacks (depth sampling), scanning and single field-of-view (FOV) videos were acquired and saved in an online database (Vivanet, Caliber ID) or on a local drive. Individual images (0.75 x 0.75 mm) from stacks and temporal single FOV frames with were used for automated quantification of immune cells, and vascular features, respectively.

***Patient tissue:***

Biopsies (targeted or non-targeted) taken as standard-of-care or for research use were used for histopathological, immunohistochemical, RNA-sequencing and flow cytometry correlations. Formalin-fixed paraffin embedded (FFPE) specimens from 34 lesions were used for histopathological and immunohistochemical correlations. Eight tissue sections were provided by a collaborator for tertiary lymphoid structure studies. RNA-extraction was on 25 lesions and RNA-seq was performed on 14 lesions. Imaging-guided targeted biopsy was performed on 7 lesions; frozen sections followed by pathology/IHC on 3 lesions and flow cytometry was performed on 4 lesions.

***RCM Data:***

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Stacks and videos from 97 lesions were analyzed in this study. Machine learning-based immune cell quantification was explored on 1026 frames from 93 cancer lesions. Each lesion contributed 5-27 independent images. The algorithm was tested on 652 independent images from 33 lesions at ~20 representative images/lesion. For vascular feature quantification, 438 single FOV videos (39, 813 frames) from 48 cancer lesions were selected. Each lesion contributed 1-31 videos. Quantified values from 270 videos in 31 patients in the analyzed set were used for subsequent correlations with manual evaluation and gene expression.

***Quantification:***

***Machine learning for immune cell:*** A pixel-wise segmentation model was trained for 4 different morphological patterns (dendritic cells, macrophages, leukocytic round-ellipsoid cells and miscellaneous immune cells) imperative for TiME analysis. We binned them into 2 classes as class 1: Dendritic cells and Melanophages (macrophages), Class 2: Leukocytes and Miscellaneous immune cells. As a third class we also labelled areas that did not contain any of these patterns as background. 1026 RCM images from 93 lesions were labelled pixelwise for these 3 classes in a non-exhaustive manner, where we only labelled examples of these patterns (Fig S7A). A total of 12% of the pixels were labelled (6% Class 1, 3% Class 2 and 91% Class 3). We trained a 3 class UNet[49]segmentation model using the MONAI framework[50]. We used 926 images for training and 100 independent images for testing the model. Based on our former studies[51,52], we downsampled the RCM images to 256 by 256 pixels (corresponding to 2 µm resolution) for the sake of computational efficiency. We use a learning rate of 5e-2, batch size of 64, and SGD optimizer with Nesterov momentum. We also used image augmentation such as random rotation, flipping, elastic-affine deformation, intensity scaling, to increase the training dataset size. The model is trained for 90 epoch using DICE loss. After 90 epochs we did not see any improvement in the loss. Dice similarity coefficient of 0.72 was found for these 3 classes (Fig 4A).

***Vascular features***

For all video frames, a two-step image stabilization procedure was used to remove the significant motion found in each movie segment. Firstly, a linear pre-alignment is performed to minimize large scale motion in Fiji[53] using the SIFT feature plugin Plugins->Registration-> Linear Stack Alignment with SIFT and default parameters. Stabilized images are then automatically cropped in Matlab (mathworks.com) to remove black background and include only areas within the field of view during the entirety of the movie segment. The crop rectangle is computed automatically by iteratively removing the row or column of pixels which contains the most blank pixels in a temporal min image until all outer edge rows and columns that contain more than three quarters blank pixels are removed. A second custom nonlinear stabilization was then performed in Matlab to remove large scale tissue deformations over time. Frame t+1 first has its histogram equalized to match frame t and then is aligned to frame t using the imregdemon procedure with four pyramid levels and steeply decreasing iterations of alignment at

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

successively finer scales (iterations, [100,50,10,1]). Frame t+2 is then aligned with the transformed frame t+1 and so on. Cropping of all regions not in view throughout the movie is again performed via the same procedure.

***-Trafficking***

**Background Subtraction** A background image is estimated for each frame as the median per pixel over a temporal window of 6 s centered on the current frame. Where movie temporal resolution differs, the window in frames is adjusted accordingly. This background estimate is subtracted out of the current frame, largely isolating moving cells on a dark background. We experimented with sparse linear methods for background subtraction, but found increased distortion in extracted foreground cells were a persistent problem across methods (data not shown). Mean and min background estimates were also tried, as well as dividing through by, rather than subtracting, background estimates, which desirably enhances dim cells. This advantage was offset by noise enhancement in non-vessel voids in the tissue (data not shown).

**Tracking** Background subtracted images are exported from Matlab as 32bit OME tiffs and imported into Fiji. Tracking is then performed in Trackmate[30] using DoG spot detection (subpixel=true; radius=7.5pixels (7.5/1.33= 5.63 micron) ; threshold=1.6) and the LAP tracker with no splitting, merging or gap closing, and a max match distance of 20 pixels (20/1.33 = 15.03 micron). The tracklets found are then filtered in Matlab to remove spurious tracklets corresponding to imperfectly removed background elements (this occurs particularly during changes in z during imaging) or tracks strung together from different fast moving circulating blood cells while preserving the desired target population. Features used to measure tracklet desirability are detailed below. Thresholds were set quantitatively and automatically to maximize correspondence between automated results and manual counts on an initial training set of 40 movies (approx. 10% of overall data). Three different temporal windows ranging from 0.6s, 0.8 s and 1s were investigated for total quantification of rolling, crawling and adherent cells. Constrained optimization within a restricted range was adopted, although fully independent threshold optimization was also investigated (Fig S7C-D). Moderate-high correlation (0.79-0.82) was observed during first optimization following which trafficking was quantified on remaining videos. Final validation using manual counts on a subset of videos (~2.5% of total data) by two readers with high inter-reader concordance (Fig S7E) found high correlation (0.74-.0.89) for different temporal windows (Fig 4A, Fig S7F). Temporal window 3 was selected for subsequent analysis to ensure inclusion of especially faster trafficking processes (rolling cells) in shorter blood vessels. The correlation was worse for videos with remnant motion after two-step motion minimization strategy, suggesting need for minimizing in axial and lateral motion during data acquisition, and use of more efficient motion removal algorithms in future.

The Tracklet Parameters used are as follows:

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Displacement=[15.41,16.92,16.92]um([20.5,22.5,22.5]px)

Consistency=[58,58,58] degrees

Quality=[1.6,1.65,1.75] arbitrary units;

Length=[0.6 s, 0.8 s or 1 s])

Where,

Displacement- total displacement between tracklet start and end point, in pixels (tracklets with lower displacement are discarded)

Motion Consistency – average angle between the motion vector of the track at successive timepoints in degrees (tracklets with higher angular difference are discarded)

Quality- average quality of detections making up the tracklet as measured by Trackmate (lower average quality tracklets are discarded)

Length – duration in s of tracklet, in all cases this was set to the thresholds used in manual counts (shorter tracklets are discarded)

***-Blood vessel segmentation:***

Manual segmentation of blood vessels was performed using an open-source segmentation platform called 3D Splicer (https://www.slicer.org/)[54] on 25 randomly selected videos. Two videos were discarded from analysis due to extreme Z-motion. The remaining 23 videos were processed to display only every 10$^{th}$ frame to mimic the automated segmentation approach; each frame in the resulting file was segmented. The entire video segmentation was exported as a Nifti (.nii) file format and imported into Matlab as a 3D image array, where consecutive images in the array correspond to consecutive frames in the RCM video. Ensuring that the consecutive frames are registered, our assumption for detecting the vessels was that the areas of high variation between consecutive frames correspond to vessels. In order to suppress the variation due to speckle noise in the RCM images, we first applied a gaussian smoothing filter (sigma = 1px). Then we applied a finite impulse response high pass filter (F = [0.5,-1,0.5]) and smooth out the extracted pixel-wise variation in time using a 7-by-7 median filter. We then subtract the mean variation of each frame to eliminate the slowly varying areas, and obtain a variation map for the whole video by accumulating the variation over the entire video sequence. We finally apply otsu thresholding the final variation map to find the areas of vessels in the videos. To smooth the border of the vessels and clean out the noise in the segmentation, we applied morphological closing operation on the binary segmentation map and clean segmented areas smaller than 0.1% and larger 10% of the entire frame. Dice similarity coefficients were calculated for comparing manual and automated vessel segmentation (Fig S7B, 4A).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

***RCM manual evaluation***

RCM features were manually evaluated (Fig S1A-D) by 4 readers with at least 4 years' experience (AS) or >20 year experience in interpreting RCM images for the main study on 33 cancers (MC, SG) or treatment response study on 14 cancers (CMAF). The major features evaluated on manual reading included number of vessels, dilated vessels, trafficking, intratumoral inflammation, peritumor inflammation and perivascular inflammation. These features were graded on a scale of 0-3 after exhaustive review of data from each patient. For imiquimod response study, type and spatial distribution of vessels and three immune morphologies were accounted for assessing time features that associate with response.

***Histopathological evaluation***

Same TiME features evaluated on RCM were also graded on digitized histopathological slides of 33 patients by a board-certified dermatopathologist (MG).

***Agreement and correlation studies***

Agreements between two readers' manual evaluations for binary RCM feature presence were quantified using Cohen's kappa coefficients. For the evaluations between RCM and histology, agreement regarding the extent of each feature presence was quantified using linearly weighted Gwet's AC1 for each of the two RCM readers to the single histology reader. The simple average of the two Gwet's AC1 scores were reported for each feature. Binary feature presence on RCM versus histology was derived the same way after recoding the manual evaluations.

Correlations between automated and manual features were computed using Spearman's correlation (1-tailed, confidence interval-95%). Spearman correlation between RCM quantified features and immune-related (nanostring), trafficking-related (Gene ontology reference) and angiogenesis score (ref).

***Statistical clustering for TiME phenotyping***

Unsupervised statistical clustering on TiME features was performed to explore classification trends or phenotypes. Principal component analysis (PCA) was used for clustering for manual evaluation and quantified RCM features. Centered method (data scaled such that mean =0, Sd unchanged) or standardized method (data scaled such that mean=0, SD=1) for manual and automated PCA, respectively. PCs were selected such that largest eigenvalues together accounted for 95% of the total variance. Loadings and biplot are presented along with scatter plot. Six RCM features were selected for (intratumoral, peritumoral and perivascular inflammation, vessel area/number, dilated vessels, trafficking). Based on loadings in the manual PCA, patients in the area

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

under the inflammation , trafficking and vessel loading vectors were termed as red phenotype. Patients in the area under the inflammation vector were pink phenotype, patients under the vessel and trafficking vectors were blue phenotype. Patients close to vessel vectors were categorized as light blue phenotype. PCA for individual gene groups (inflammation[26], angiogenesis[24] and trafficking(GO Pathways[55])) were also analyzed to derive correlations with RCM phenotypes. Another PCA on response to imiquimod was done using similar parameters. PCA was performed in Graphpad 9.0.

***Immunohistochemistry***

CD3, CD68 and CD20 IHC were performed on BOND RX(Leica) while MPO IHC was performed on Ultra Discovery platform(Roche). The protocol for the Bond Rx platform included ER2 (High pH buffer) -30 minutes for Heat retrieval followed by 30 minutes incubation time for Primary Abs (Santa Cruz Biotech, US). Polymer Detection was through DAB Kit (catalog DS9800). For the dual CD3/CD20 sequential stain, ER2 -30 minutes for heat retrieval, 30 minute incubation time for Primary Ab followed by Polymer Detection kit. This was followed by ER2-20 minute, 30 minute -incubation with second Ab, Polymer refine Red detection KIT , (catalog # DS9390). The protocol for Discovery Ultra involved CC1-32 minutes for Heat Retrieval, 32 minutes -incubation of Primary Ab. OptiView DAB IHC Detection Kit (catalog # 760-700).

***IHC evaluation and quantification***

CD3, CD20 and MPO stains were evaluated on 33 cases by a board-certified dermatopathologist (MG) for presence/absence of ulceration/erosion and CD3+ T-cells , CD20+ B-cells , total lymphocytes (CD3+ CD20+) and neutrophils. In addition, for each immune marker, features were evaluated on a scale of 0-3 where 0 is absent and 3 is highest. These features included predominant distribution, TILs, trafficking and distribution at tumor periphery. For MPO, an additional category called intravascular cells was evaluated. Tertiary lymphoid structures (TLS) labeled by dual CD3/CD20 staining on 40 cases were also analyzed for total TLS numbers, TLS dimensions (maximum dimensions in X and Y) and tissue size (maximum dimensions in X and Y) to compute TLS numbers/mm2 and TLS area coverage/$mm^2$. Within defined TLS and non-TLS areas (used as control), tumor killing as defined on histopathology was noted, and TILs in TLS-adjacent tumor nests were counted. For TLS positive patients, both TLS and non-TLS areas were evaluated for TILs and tumor killing, in TLS negative patients, TILs and tumor killing were specified in defined areas (Fig S3).
Confidence intervals for median proportion of TLS area coverage was derived from percentile bootstrapping approach. Mann-Whitney U tests were used to quantify the statistical significance in differences between median proportion of TLS coverage across binary clinical factors such as ulceration presence and NMSC classification. Generalized estimating equations (GEE) were used to estimate association between local TiME

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TLS presence with both TILs presence and tumor killing by clustering on histologic specimen using an exchangeable correlation structure. This approach was applied to the binary classification of local TLS area versus local control area to the continuous outcome of local TILs cell count density. GEE was applied exclusively to TLS regions to model the local TLS area coverage to both the continuous outcome of TILs cell count using Gaussian link function and to the binary outcome of tumor killing using the logit link function.

***RNA analysis on GEO datasets***

Two previously-published RNA-seq data sets (GSE125285, GSE128795) and one microarray data set (GSE53462) for Basal Cell Carcinoma samples were downloaded from the Gene Expression Omnibus. GSE125285 and GSE128795 contained pre-processed RNA-seq data, however, the microarray data was indexed to Illumina Probes[56,57] . First, those probes with high detection p-values ($p > 0.05$ for 13 out of the 26 samples) were filtered out, leaving 23, 176 probes remaining from an initial value of 47, 323. ProbeIDs were matched to common gene identifiers using illuminaID2nuID[12]. Of the remaining probes, 5, 778 did not have a unique gene associated with them. We took the value with the highest expression to have each gene represented only once, leaving 17398 genes. We created phenotype groupings a priori via unbiased clustering through immune-related genes provided by Nanostring[58]©. For the whole transcriptome analysis, we used the built-in R heatmap function (stats 4.0.2) to create phenotype clusters. The heatmap revealed 2 groups on which the DGEA analysis was performed. Functional enrichment analysis was performed using GO enrichment analysis (https://go.princeton.edu/tmp/5497206//query_results.html), and each enriched ontology hierarchy (false discovery rate (FDR) $< 0.05$) was reported with two terms in the hierarchy: (1) the term with the highest significance value and (2) the term with the highest specificity

***RNA extraction***

FFPE sections were deparaffinized using the mineral oil method. Briefly, 800µL mineral oil was mixed with the sections and the sample was incubated at 65°C for 10 minutes. Phases were separated by centrifugation in 360µL Buffer PKD and Proteinase K was added for digestion. After a three-step incubation (65°C for 45', 80°C for 15', 65°C for 30') and additional centrifugation, the aqueous phase containing RNA was removed and DNase treated. The RNA was then extracted using the RNeasy FFPE Kit (QIAGEN catalog # 73504) on the QIAcube Connect (Qiagen) according to the manufacturer's protocol with 285µL input. Samples were eluted in 13µL RNase-free water.

***Transcriptome sequencing***

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 356-500ng of total RNA with DV200% varying from 88-93 underwent ribosomal depletion and library preparation using the TruSeq Stranded Total RNA LT Kit (Illumina catalog # RS-122-1202) according to instructions provided by the manufacturer with 8 cycles of PCR. Samples were barcoded and run on a HiSeq 4000 in a PE100 run, using the HiSeq 3000/4000 SBS Kit (Illumina). On average, 78 million paired reads were generated per sample and 20% of the data mapped to mRNA.

***CIBERSORT analysis***

CIBERSORT was used for the immune cell analysis to delineate immune subsets using 584 genes for 22 immune cell types[25]. Transcript per million values were used as input. CIBERSORT chooses the record with the highest mean expression across the mixtures during analysis. The gene expression file with 14 cases was uploaded to CIBERSORT as a mixture file, and CIBERSORT was run with the following options: relative and absolute modes together, LM22 signature gene file, 100 permutations, and quantile normalization disabled. Sample distance matrix resulting from immune cell distribution, k-means clustering and differential gene expression analysis (DGEA) were used to interpret CIBERSORT output.

***Differential gene expression analysis (DGEA)***

DESeq2 (ver 1.28.1) was used to perform differential gene expression analysis comparing RCM groups 1 vs 2. Genes with an absolute log2 fold change of >= 0.5 an adjusted p-value of < =0.1 were considered significantly changed. Log transformation was then performed on the full gene expression matrix with the *rlog* function and the transformed read counts of the 114 significantly changed genes were extracted for unsupervised hierarchical clustering analysis with pheatmap (ver 1.0.12, clustering_method = "complete")

***Gene set enrichment analysis***

471 angiogenesis genes were identified to generate the angiogenesis core gene set and 547 immune genes were extracted from CIBERSORT analysis as mentioned above. Top 10% genes differentially expressed between the RCM groups 1 vs 2 ranked by absolute fold change were identified and ranked from the highest to the lowest fold change. Gene set enrichment analysis was then performed on the 2,529 genes to calculate the enrichment score for the angiogenesis and immune gene sets with the R package fgsea (ver 1.14.0) and the *fgseaMultilevel* function.

***Response to Immunotherapy analysis***

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Correlation of TiME features and phenotypes with response to topical immunotherapy Imiquimod were analyzed on 13 lesions (MSK). The patients were imaged at baseline (T0) and the TiME features were further analyzed with respect to response to treatment. Linear regression modeling was undertaken to quantitatively identify the predictor variables for response to imiquimod and compared against the known "standard" which is tumor-infiltrating lymphocytes and intratumoral inflammation. In order to measure the predictive power of each feature, we train predictive models in a leave-one-out cross-validation fashion and measure the model performance by inferring on the left-out test sample (out-of-bag estimates). This procedure is followed in an iterative manner, where we select a single feature that gives the highest performance and add a new feature that provides the highest performance in each iteration. Model performance was measured calculating specificity (higher the better) on the out-of-bag estimates and Akaike Information Criterion (lower the better) value of the model. In this way, the features are prioritized according to their predictive power. Moreover, we also examined the linear separability of (i) individual features by looking at the histogram of feature values for each sample, and (ii) each pairwise feature combination by examining kernel density estimation plots.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**References:**


1 Darvin, P., Toor, S. M., Nair, V. S. & Elkord, E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp. Mol. Med.* **50**, 1-11 (2018).

2 Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960-1964 (2006).

3 Topalian, S. L., Taube, J. M., Anders, R. A. & Pardoll, D. M. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nature Reviews Cancer* **16**, 275-287 (2016).

4 De Guillebon, E. *et al.* Beyond the concept of cold and hot tumors for the development of novel predictive biomarkers and the rational design of immunotherapy combination. *Int. J. Cancer* **147**, 1509-1518 (2020).

5 Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541 (2018).

6 Gnjatic, S. *et al.* Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy. *Journal for immunotherapy of cancer* **5**, 1-18 (2017).

7 Rajadhyaksha, M., Anderson, R. R. & Webb, R. H. Video-rate confocal scanning laser microscope for imaging human tissues in vivo. *Appl. Opt.* **38**, 2105-2115 (1999).

8 Friedman, C. F. & Postow, M. A. Emerging tissue and blood-based biomarkers that may predict response to immune checkpoint inhibition. *Curr. Oncol. Rep.* **18**, 21 (2016).

9 Filipovic, A., Miller, G. & Bolen, J. Progress Toward Identifying Exact Proxies for Predicting Response to Immunotherapies. *Frontiers in Cell and Developmental Biology* **8**, 155 (2020).

10 Fridman, W. H., Zitvogel, L., Sautès–Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nature reviews Clinical oncology* **14**, 717 (2017).

11 Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature* **541**, 321-330 (2017).

12 Paijens, S. T., Vledder, A., de Bruyn, M. & Nijman, H. W. Tumor-infiltrating lymphocytes in the immunotherapy era. *Cell. Mol. Immunol.*, 1-18 (2020).

13 Tang, T. *et al.* Advantages of targeting the tumor immune microenvironment over blocking immune checkpoint in cancer immunotherapy. *Signal Transduction and Targeted Therapy* **6**, 1-13 (2021).

14 Schaaf, M. B., Garg, A. D. & Agostinis, P. Defining the role of the tumor vasculature in antitumor immunity and immunotherapy. *Cell Death Dis.* **9**, 115 (2018).

15 Hendry, S. A. *et al.* The role of the tumor vasculature in the host immune response: implications for therapeutic strategies targeting the tumor microenvironment. *Front. Immunol.* **7**, 621 (2016).

16 Klein, D. The tumor vascular endothelium as decision maker in cancer therapy. *Front. Oncol.* **8**, 367 (2018).

17 Motz, G. T. & Coukos, G. The parallel lives of angiogenesis and immunosuppression: cancer and other tales. *Nature reviews Immunology* **11**, 702-711 (2011).

18 Huinen, Z. R., Huijbers, E. J. M., van Beijnum, J. R., Nowak-Sliwinska, P. & Griffioen, A. W. Anti-angiogenic agents — overcoming tumour endothelial cell anergy and improving immunotherapy outcomes. *Nature Reviews Clinical Oncology*, doi:10.1038/s41571-021-00496-y (2021).

19 Fisher, D. T. *et al.* Intraoperative intravital microscopy permits the study of human tumour vessels. *Nature communications* **7**, 1-9 (2016).

20 Rajadhyaksha, M., Grossman, M., Esterowitz, D., Webb, R. H. & Anderson, R. R. In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. *J. Invest. Dermatol.* **104**, 946-952 (1995).

21 González, S., Sackstein, R., Anderson, R. R. & Rajadhyaksha, M. Real-time evidence of in vivo leukocyte trafficking in human skin by reflectance confocal microscopy. *The Journal of investigative dermatology* **117**, 384 (2001).

22 Ahlgrimm-Siess, V. *et al.* The vasculature of nonmelanocytic skin tumors in reflectance confocal microscopy: vascular features of basal cell carcinoma. *Arch. Dermatol.* **146**, 353 (2010).

23 Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nature Reviews Cancer* **19**, 307-325 (2019).

24 Masiero, M. *et al.* A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell* **24**, 229-241 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.


25 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453-457 (2015).

26 Cesano, A. nCounter<sup>®</sup> PanCancer Immune Profiling Panel (NanoString Technologies, Inc., Seattle, WA). *Journal for ImmunoTherapy of Cancer* **3**, 42, doi:10.1186/s40425-015-0088-7 (2015).

27 Sand, M. *et al.* Dicer Sequencing, Whole Genome Methylation Profiling, mRNA and smallRNA Sequencing Analysis in Basal Cell Carcinoma. *Cell. Physiol. Biochem.* **53**, 760-773 (2019).

28 Jee, B. A. *et al.* Molecular classification of basal cell carcinoma of skin by gene expression profiling. *Mol. Carcinog.* **54**, 1605-1612 (2015).

29 Wan, J. *et al.* Distinct transcriptomic landscapes of cutaneous basal cell carcinomas and squamous cell carcinomas. *Genes & Diseases* (2019).

30 Tinevez, J.-Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80-90 (2017).

31 Gide, T. N. *et al.* Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. *Cancer Cell* **35**, 238-255. e236 (2019).

32 Cristescu, R. *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy. *Science* **362** (2018).

33 Cheng, D. T. *et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *The Journal of molecular diagnostics* **17**, 251-264 (2015).

34 Harlin, H. *et al.* Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. *Cancer Res.* **69**, 3077-3085 (2009).

35 Gajewski, T. F. *et al.* Cancer immunotherapy targets based on understanding the T cell-inflamed versus non-T cell-inflamed tumor microenvironment. *Tumor immune microenvironment in cancer progression and cancer therapy*, 19-31 (2017).

36 Bindea, G. *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782-795 (2013).

37 Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231-235 (2015).

38 Yang, L., Li, A., Lei, Q. & Zhang, Y. Tumor-intrinsic signaling pathways: key roles in the regulation of the immunosuppressive tumor microenvironment. *J. Hematol. Oncol.* **12**, 1-14 (2019).

39 Stacker, S. A. & Achen, M. G. Emerging roles for VEGF-D in human disease. *Biomolecules* **8**, 1 (2018).

40 Paluskievicz, C. M. *et al.* T regulatory cells and priming the suppressive tumor microenvironment. *Front. Immunol.* **10**, 2453 (2019).

41 Lucena, S. R. *et al.* Combined treatments with photodynamic therapy for non-melanoma skin cancer. *Int. J. Mol. Sci.* **16**, 25912-25933 (2015).

42 Navarrete-Dechent, C. *et al.* Optical imaging guided-'precision' biopsy of skin tumors: a novel approach for targeted sampling and histopathologic correlation. *Archives of dermatological research*, 1-13 (2020).

43 Mantri, M. *et al.* Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nature communications* **12**, 1-13 (2021).

44 Hunter, M. V., Moncada, R., Weiss, J. M., Yanai, I. & White, R. M. Spatially resolved transcriptomics reveals the architecture of the tumor/microenvironment interface. *bioRxiv*, 2020.2011.2005.368753, doi:10.1101/2020.11.05.368753 (2021).

45 Sahu, A. *et al.* Evaluation of a Combined Reflectance Confocal Microscopy–Optical Coherence Tomography Device for Detection and Depth Assessment of Basal Cell Carcinoma. *JAMA dermatology* **154**, 1175-1183 (2018).

46 Schuh, S. *et al.* Imaging blood vessel morphology in skin: dynamic optical coherence tomography as a novel potential diagnostic tool in dermatology. *Dermatology and therapy* **7**, 187-202 (2017).

47 Lin, R. *et al.* Longitudinal label-free optical-resolution photoacoustic microscopy of tumor angiogenesis in vivo. *Quantitative imaging in medicine and surgery* **5**, 23 (2015).

48 Walsh, A. J. *et al.* Label-free Method for Classification of T cell Activation. *bioRxiv*, 536813 (2019).

49 Ronneberger, O., Fischer, P. & Brox, T. 234-241 (Springer International Publishing).

50 1.The MONAI Consortium. Project MONAI. (2020). doi:10.5281/zenodo.4323059.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

51 Kose, K. *et al.* Segmentation of cellular patterns in confocal images of melanocytic lesions in vivo via a multiscale encoder-decoder network (med-net). *Med. Image Anal.* **67**, 101841 (2021).

52 D'Alonzo, M. *et al.* Semantic segmentation of reflectance confocal microscopy mosaics of pigmented lesions using weak labels. *Sci. Rep.* **11**, 1-13 (2021).

53 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682 (2012).

54 Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **30**, 1323-1341 (2012).

55 Smith, C. L. & Eppig, J. T. The mammalian phenotype ontology: enabling robust annotation and comparative analysis. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **1**, 390-399 (2009).

56 Shi, W., Oshlack, A. & Smyth, G. K. Optimizing the noise versus bias trade-off for Illumina whole genome expression BeadChips. *Nucleic Acids Res.* **38**, e204-e204 (2010).

57 Du, P., Kibbe, W. A. & Lin, S. M. nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. *Biol. Direct* **2**, 1-7 (2007).

58 Cesano, A. nCounter® PanCancer immune profiling panel (NanoString technologies, Inc., Seattle, WA). *Journal for immunotherapy of cancer* **3**, 1-3 (2015).

59 Zhu, J. & Paul, W. E. CD4 T cells: fates, functions, and faults. *Blood, The Journal of the American Society of Hematology* **112**, 1557-1569 (2008).

60 Wang, Z. & Zavislan, J. in *Frontiers in Optics.* FWD5 (Optical Society of America).

61 Actor, J. K. Elsevier's Integrated Review Immunology and Microbiology E-Book: with STUDENT CONSULT Online Access. (2011).

62 Manley, H. R., Keightley, M. C. & Lieschke, G. J. The neutrophil nucleus: an important influence on neutrophil migration and function. *Front. Immunol.* **9**, 2867 (2018).

63 Willetts, L., Ochkur, S. I., Jacobsen, E. A., Lee, J. J. & Lacy, P. in *Eosinophils* 111-128 (Springer, 2014).

64 McWhorter, F. Y., Wang, T., Nguyen, P., Chung, T. & Liu, W. F. Modulation of macrophage phenotype by cell shape. *Proceedings of the National Academy of Sciences* **110**, 17253-17258 (2013).

65 Romani, N., Clausen, B. E. & Stoitzner, P. Langerhans cells and more: langerin‐expressing dendritic cell subsets in the skin. *Immunol. Rev.* **234**, 120-141 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Table:**

| TiME Feature | RCM 2-Reader Agreement (Cohen's kappa) | Histology vs. Average 2-reader (Binary) |
|---|---|---|
| # of vessels | 0.62 | 0.84 |
| Dilated vessel presence | 0.72 | 0.89 |
| Trafficking | 0.76 | 0.72 |
| Regression | 0.74 | -0.13 |
| Intratumor inflammation | 1.00 | 1.00 |
| Peritumor inflammation | NA | 0.97 |
| Mucin | 0.93 | 0.74 |
| Perivascular inflammation | NA | 1.00 |

**Table 1. High agreement confirms the reproducibility and presence of TiME features on histopathology**

Substantial to almost perfect agreement (k=0.62-1.0) was observed for RCM features between 2-readers, k for peritumor inflammation and perivascular inflammation agreement could not be computed due to 100% prevalence.
Good to very good agreement ($AC_1$: 0.74-1.0) was found in the binary analysis between average RCM evaluation and dermatopathologist grading of features.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Figures.**

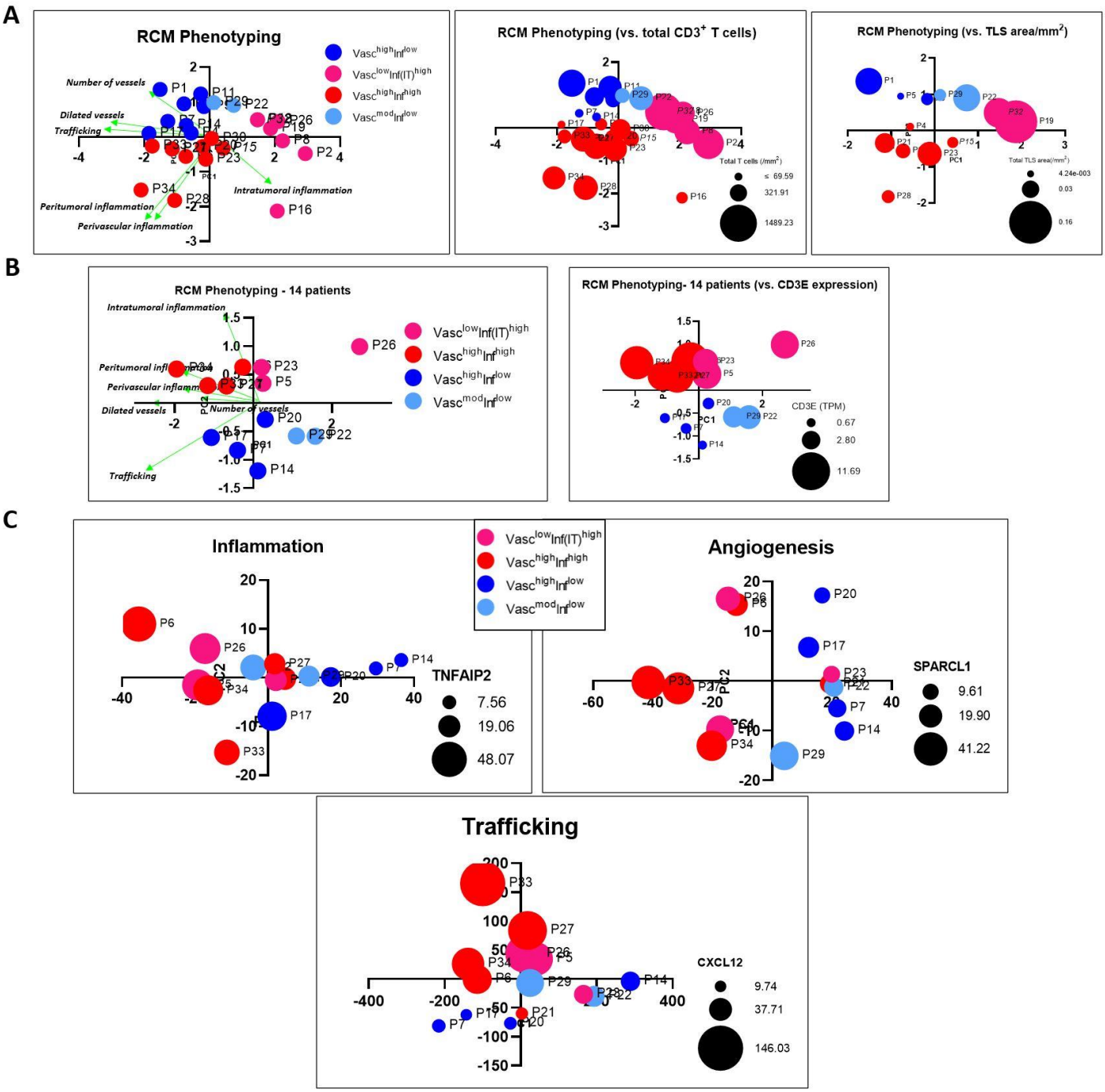



**Figure 1. Novel TiME phenotypes correlate with underlying immune and vascular features**

**A.** Unsupervised principal component analysis (PCA) on 6 RCM features (intratumor, peritumor and perivascular inflammation, vessel number, diameter and trafficking) demonstrate four phenotypes: $Vasc^{high}Inf^{high}$, $Vasc^{high}Inf^{low}$, $Vasc^{low}Inf(IT)^{high}$ and $Vasc^{mod}Inf^{low}$ on 27 patients with PC1 loadings (dilated vessels, trafficking) and PC2 (peritumoral and perivascular inflammation) contributing high variance in classification. Patients that correspond to both dilated vessels, trafficking and peritumor/perivascular inflammation were termed as $Vasc^{high}Inf^{high}$ phenotype, patients located proximal or

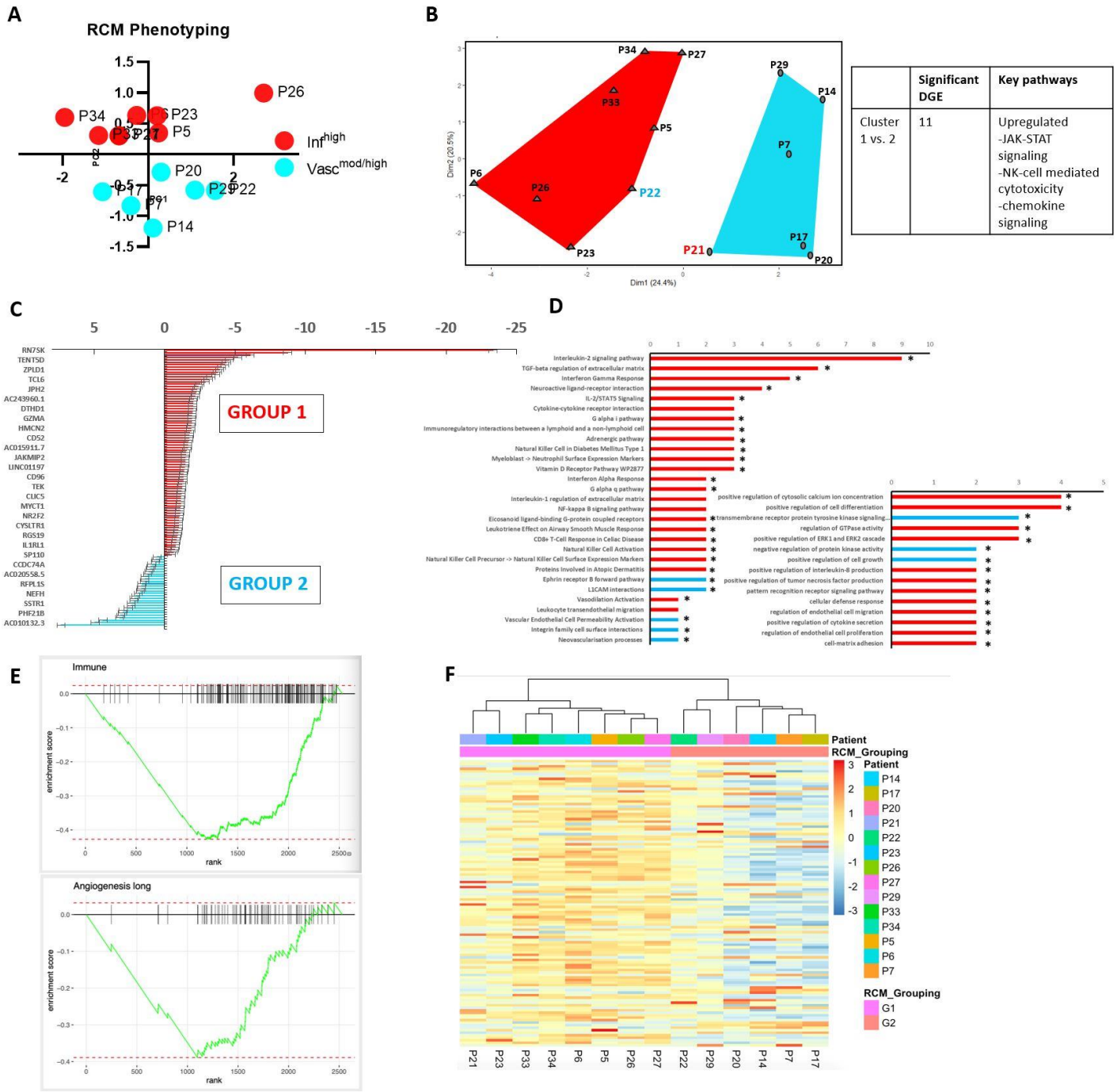
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

distal to PC1 loadings (vessels) were termed as $Vasc^{high}Inf^{low}$ and $Vasc^{mod}Inf^{low}$ , respectively. Patients distal to peritumoral and perivascular inflammation and close to intratumor inflammation loading vector were termed as $Vasc^{low}Inf(IT)^{high}$. Largest number of T-cells (denoted by size of the bubble) seen in $Vasc^{low}Inf(IT)^{high}$ along with highest density of TLS.

**B.** Unsupervised principal component analysis (PCA) on 6 RCM feature on a subset of patients shows similar phenotypes, classification driven by intratumoral inflammation and trafficking. Correlation with gene expression shows highest CD3E expression in $Vasc^{high}Inf^{high}$ phenotype.

**C.** Unsupervised clustering for phenotypes in inflammation, angiogenesis[24] and trafficking gene signatures demonstrate similar clustering as RCM phenotyping (same color code for each patient as in 1B) driven mainly by TNFAIP2, SPARCL1 and CXCL12, respectively

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




| | Significant DGE | Key pathways |
|---|---|---|
| Cluster 1 vs. 2 | 11 | Upregulated<br>-JAK-STAT signaling<br>-NK-cell mediated cytotoxicity<br>-chemokine signaling |

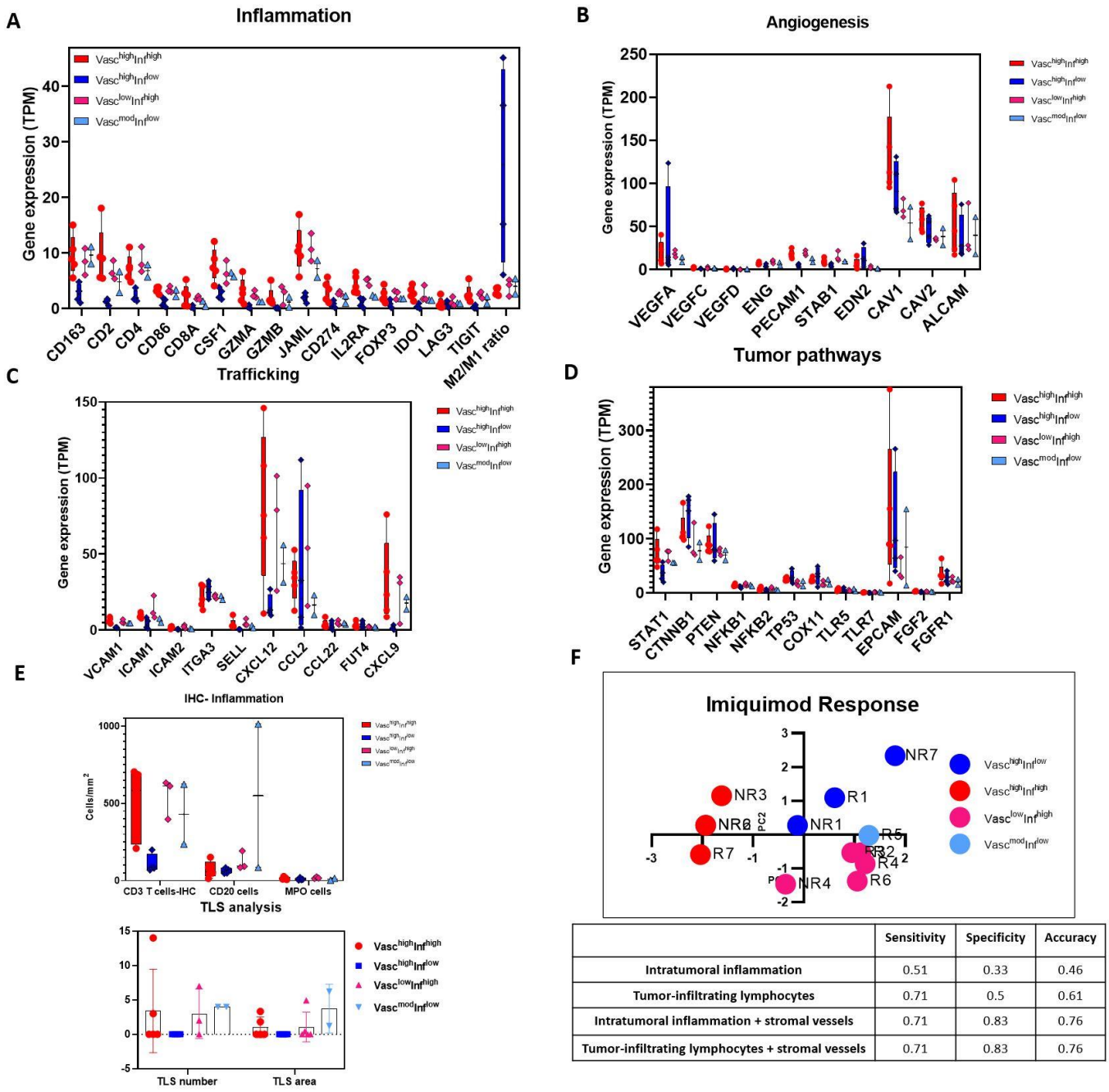
**Figure 2. Gene expression explains the biological basis of RCM phenotyping**

**A.** PCA plot of RCM data grouped in Inf$^{high}$ and Vasc$^{mod/high}$ for gene expression analysis. **B.** CIBERSORT analysis performed to select immune fractions (574 genes from 22 immune cell types) analyzed through k-means clustering also yields two groups similar to RCM phenotypes. Differential gene expression analysis (DGEA) on CIBERSORT output indicate upregulation of JAK-STAT, chemokine signaling and NK-mediated cell cytotoxicity in the red cluster. **C.** DGEA based on RCM clustering of Inf$^{high}$ and Vasc$^{mod/high}$ groups on all genes found 114 differentially expressed genes (p adj-0.1). Of these 85 genes were upregulated in Inf$^{high}$ and 29 genes in Vasc$^{mod/high}$. **D.** Gene ontology and pathway analysis was performed on the 114 genes. Pro-inflammatory pathways were overexpressed in Inf$^{high}$ group, while few genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

corresponding to leukocyte interactions and vascularity were upregulated in $Vasc^{mod/high}$ group (* denotes statistically significant results). GO pathway analysis revealed higher biological processes, cell differentiation, cytosolic calcium signaling, cytokine secretion, cellular defense in red group while negative tyrosine kinase and positive cell proliferation in blue group. **E.** Gene set enrichment analysis on immune and angiogenesis signatures demonstrates higher scores for $Inf^{high}$ group. **F.** Unsupervised HCA on the 114 genes found same clustering as DGEA.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




| | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| Intratumoral inflammation | 0.51 | 0.33 | 0.46 |
| Tumor-infiltrating lymphocytes | 0.71 | 0.5 | 0.61 |
| Intratumoral inflammation + stromal vessels | 0.71 | 0.83 | 0.76 |
| Tumor-infiltrating lymphocytes + stromal vessels | 0.71 | 0.83 | 0.76 |

**Figure 3. Molecular analysis on individual phenotypes indicate unique inflammation, angiogenic, trafficking and tumor-intrinsic signatures, reveals inflamed and anergized phenotypes and correlates inflamed phenotype with response to topical immunotherapy**

**A.** Highest immune cell gene signatures in Vasc$^{high}$Inf$^{high}$ and the Vasc$^{low}$Inf(IT)$^{high}$ phenotypes, except the M2/M2 ratio which was highest in Vasc$^{high}$Inf$^{low}$ phenotype. **B.** High END2, VEGFD in Vasc$^{high}$Inf$^{low}$ phenotype while other angiogenic genes higher in the Vasc$^{high}$Inf$^{high}$ phenptype, **C.** Lower expression of major adhesion molecules (VCAM, ICAM1, ICAM2 and SEL-L while higher ITGA3, CCL2 range in the Vasc$^{high}$Inf$^{low}$ phenotype. CXCL12 expression delineates Vasc$^{high}$Inf$^{low}$ from all other phenotypes. **D.** Higher CTNNB1, PTEN, COX11 and TP53 in the Vasc$^{high}$Inf$^{low}$ phenotype while higher STAT1 and NF-kB seen in Vasc$^{high}$Inf$^{high}$ and Vasc$^{low}$Inf(IT)$^{high}$ phenotypes. **E.** Higher T-cell, B-cell counts

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

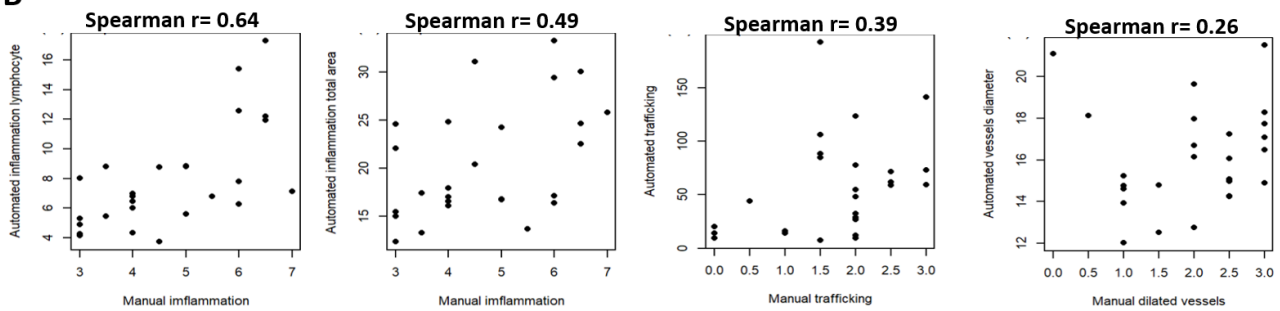
and TLS density found in all phenotypes except $Vasc^{high}Inf^{low}$ **F.** PCA on TiME features of 13 prospective patients receiving imiquimod indicates most responders (5 of 7) belonged to the $Vasc^{low}Inf(IT)^{high}$ phenotypes; phenotypic association with relatively high TLR expression noted (Fig 3D). Incorporating a vascular feature in the regression model for predicting response delivers higher accuracy than inflammation alone.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**A**

| Quantification of TiME Feature | Validation Data | Validation Result |
|---|---|---|
| Inflammation ( Machine learning UNEt model) | 9.7 % | Dice coefficient: 0.72 |
| Vessel segmentation | 5.2% | Dice coefficient: 0.29-0.78 |
| Trafficking – optimization | 10.9% | Spearman r: 0.73-0.84 |
| Trafficking – validation | 2.5% | Spearman r: 0.74-0.89 |

**B**




**C**

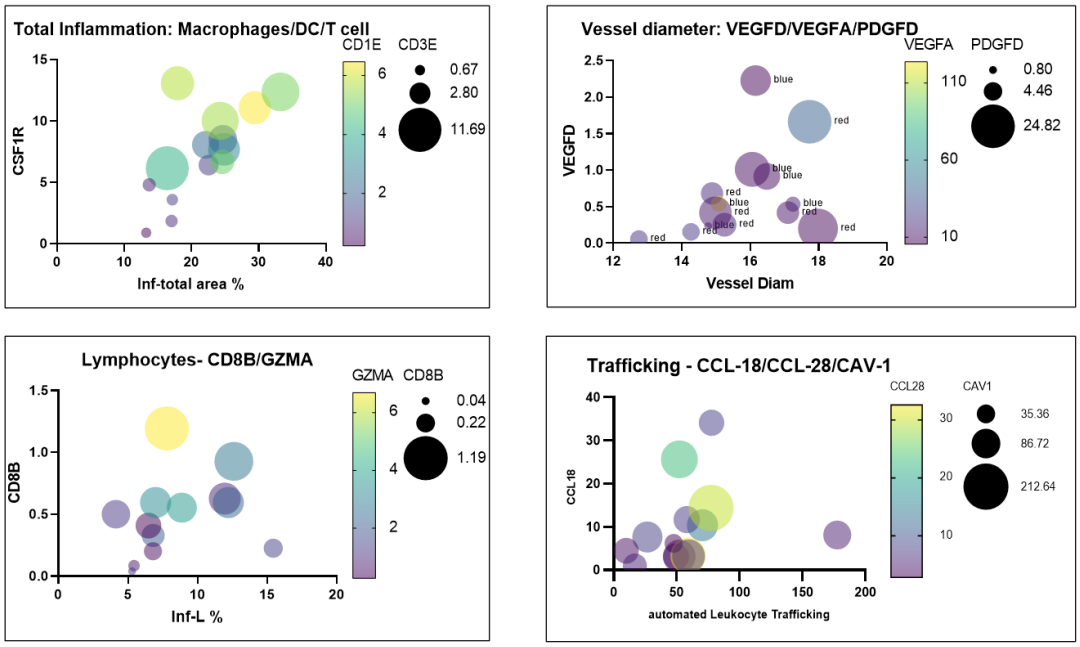



**Figure 4. Moderate-high correlation seen between automated feature quantification (immune cells and vasculature), manual evaluation, and corresponding gene signatures.**

**A.** Summary of results for validation performed after each automated quantification. **B.** Spearman correlation between manual grading and automated quantification: r=0.66 for leukocytic inflammation and manual inflammation, r=0.49 for total inflammation, r=0.34 for trafficking and r= 0.22 for vessel diameter **C.** Spearman correlation for TiME features and corresponding gene expression: Strong correlations were observed between RCM inflammation feature of total inflammation area with CSF1R (macrophage, r=0.73), CD1E (dendritic cells, r=0.64) and CD3E (T lymphocytes, r=0.51), between total lymphocyte area and CD8B (Cytotoxic T- lymphocyte, r=0.6) and GZMA (cytotoxic T-cell activation, r=0.53). In vascular features, vessel diameter correlated with VEGF-D (r=0.459) and VEGF-A (r=-0.471) and trafficking

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

positively correlated with CCL-18 (r=0.561), CAV-1 (r=0.46) and negatively with CCL25 (r=-0.7). p values were significant for all comparisons,

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Supplementary Table/Figures:**

| Cell | Cell Size (µ) | Shape | Nuclear Size (µ) | Nuclear Shape | Refractive Index | Scattering Coefficient | References |
|---|---|---|---|---|---|---|---|
| **T Cell** | 8-10 | Round | 6-7 | Irregular circular shape | 1.36 (1.34-1.37) | $0.54*10^4$ (STD=$0.076*10^4$) | [59, 60] |
| **B-Cell** | 8-10 | Round | 6-8 | Nucleus indented | 1.36 (1.34-1.37) | $0.58*10^4$ (STD=$0.075*10^4$) | [61, 60] |
| **Neutrophil** | 12-14 | Irregular-round | 3-4 | Multilobed nucleus | 1.4 (1.35-1.42) | $4*10^4$ (STD=$0.57*10^4$) | [61,62, 60] |
| **Eosinophil** | 12-15 | Round | 3-4 | Bilobed | 1.4 (1.35-1.42) | $4.6*10^4$ (STD=$1.04*10^4$) | [63, 60] |
| **Basophil** | 14-16 | Round, pleomorphic in tissue | Unknown | Bilobed | 1.4 (1.35-1.42) | $0.82*10^4$ (STD=$0.135*10^4$) | [61, 60] |
| **Macrophage** | 20-30 | Large, irregular, triangular | Unknown | Single-lobed, round centered | 1.384 ± 0.015 | Not Known | [64] |
| **Langerhans Cell** | 15-25 | Dendritic | Unknown | Indented "coffee bean" | Not Known | Not Known | [65] |

**Table S1. Distinct optical and cellular properties enable visualization and morphological distinction between major immune cell classes such as dendritic cells, macrophages and lymphocytes.**

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

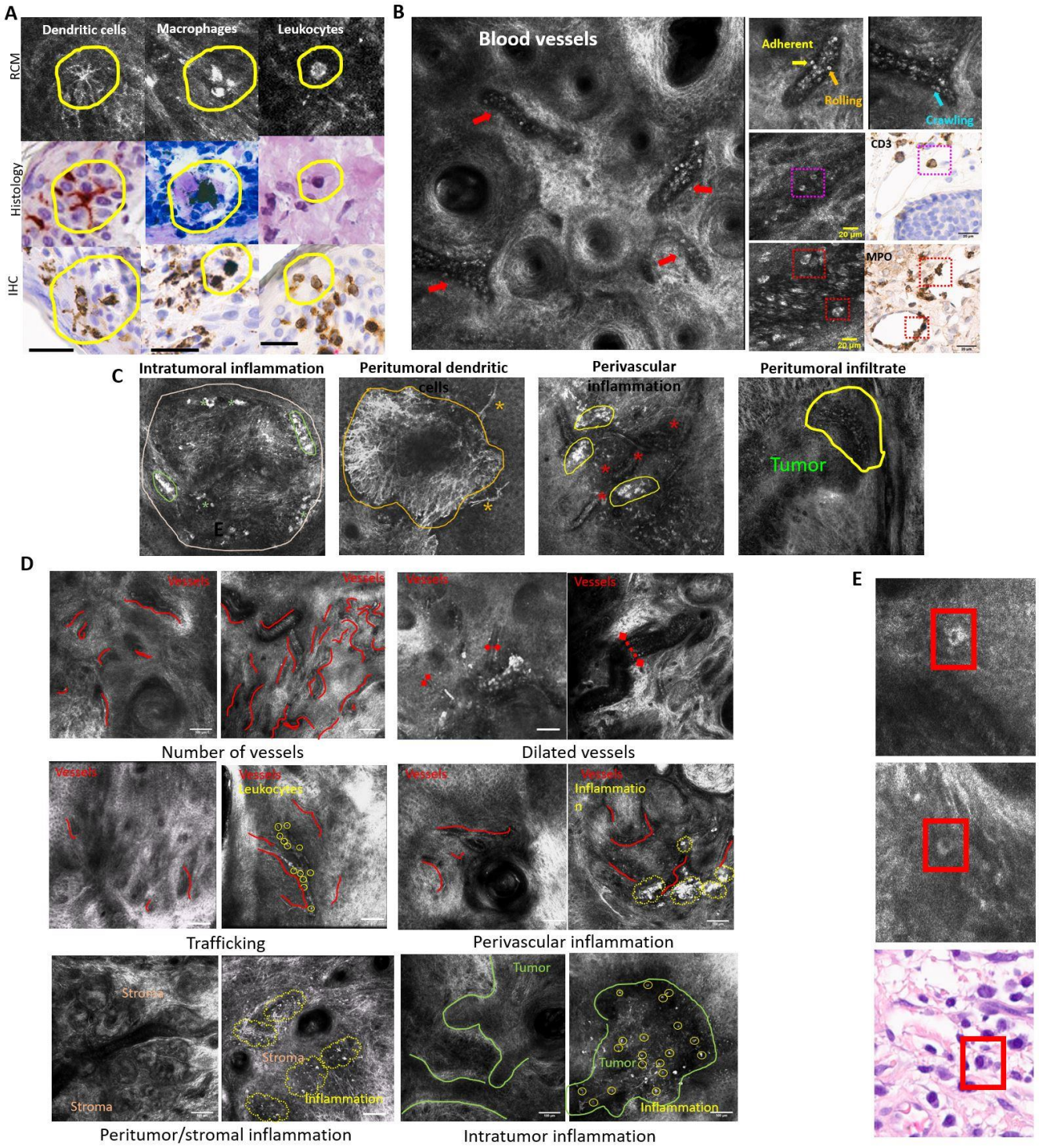



**Figure S1. Tumor-immune microenvironment and vasculature visualized and evaluated on RCM.**

**A.** Dendritic cells, macrophages and lymphocytes identified by RCM are confirmed using histology and immunohistochemistry (CD1, CD68 and CD3, respectively) on targeted biopsies. **B.** RCM image showing blood vessels (red arrows) and trafficking of leukocytes on the vessel edges and individual leukocyte trafficking steps: adhering, rolling and crawling. Trafficking of lymphocytes and neutrophils confirmed using CD3 and Myeloperoxidase IHC, respectively. **C.** Intratumoral, peritumoral and perivascular distribution of inflammatory cells as established on histopathology is seen on RCM. **D.** Representative example of features assessed during manual evaluation. For each feature, images with

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

no/minimal feature presence and scored either 0 or 1 are shown on the left, while images with high density of features scored as 2 or 3 shown on the right (Red curve= blood vessels, diameter of vessels indicated by length of lines between diamonds; trafficking leukocytes encircled in yellow; inflamed area adjacent to vessels or in stroma marked by yellow lines; intratumor immune cells encircled in yellow) **E.** Round to elliptical, large (15-20 micron) immune cells with an eccentric nucleus and bright cytoplasm in dermis, often adjacent to blood vessels (perivascularly) resembling plasma cells on histopathology (red boxes) with an eccentric nucleus, perinuclear hoff and cytoplasm, seen next to a blood vessel

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

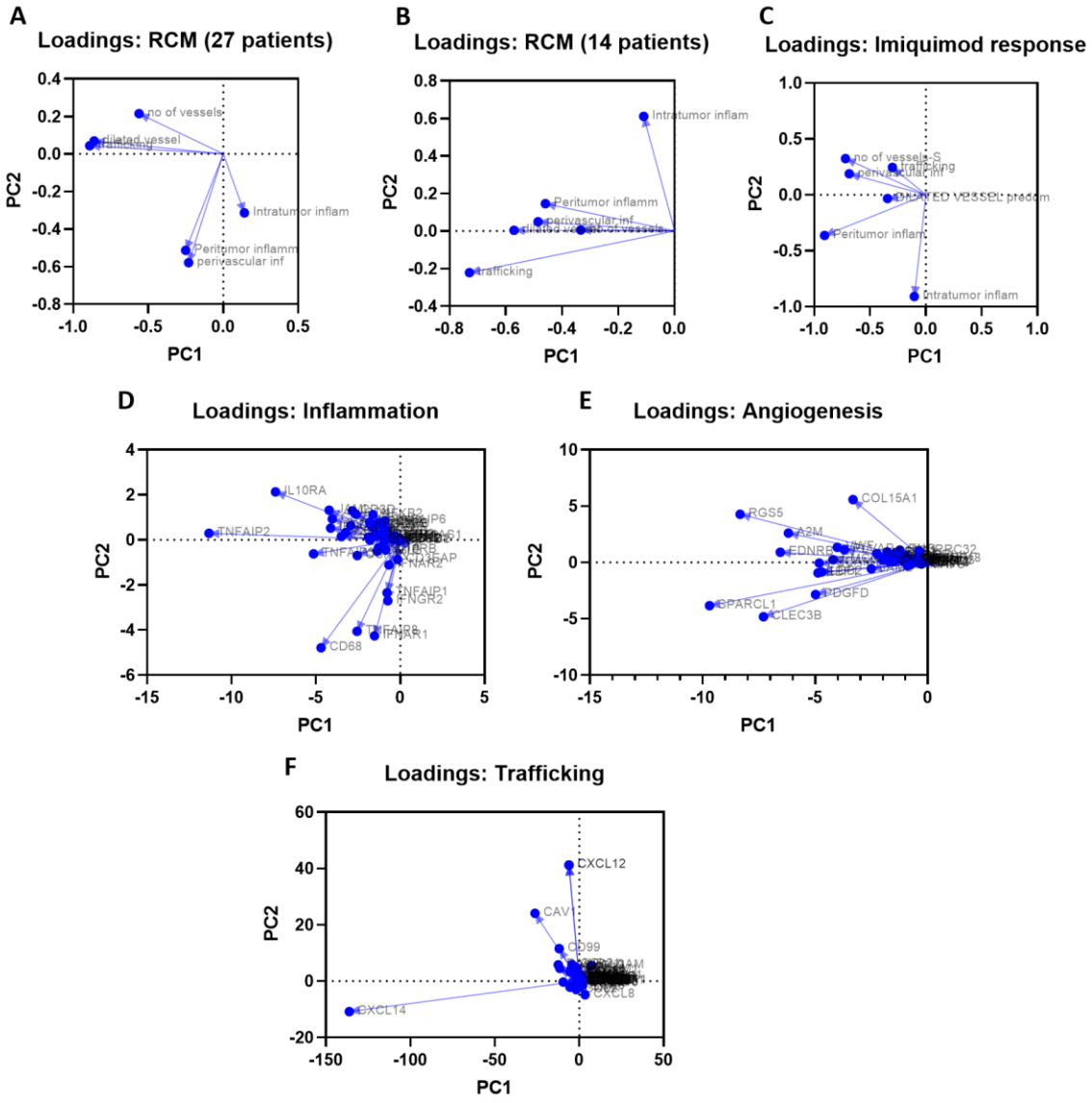



**Figure S2. PCA loadings for RCM and gene expression analysis.**

**A.** PC loadings for PCA on total BCC patients

**B.** PC loadings for PCA on BCC patients with gene expression

**C.** PC loadings for PCA on BCC patients undergoing imiquimod treatment

**D.** PC loadings for PCA on gene expression of major immune cells

**E.** PC loadings for PCA on angiogenic signatures

**F.** PC loadings for PCA on trafficking-associated genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

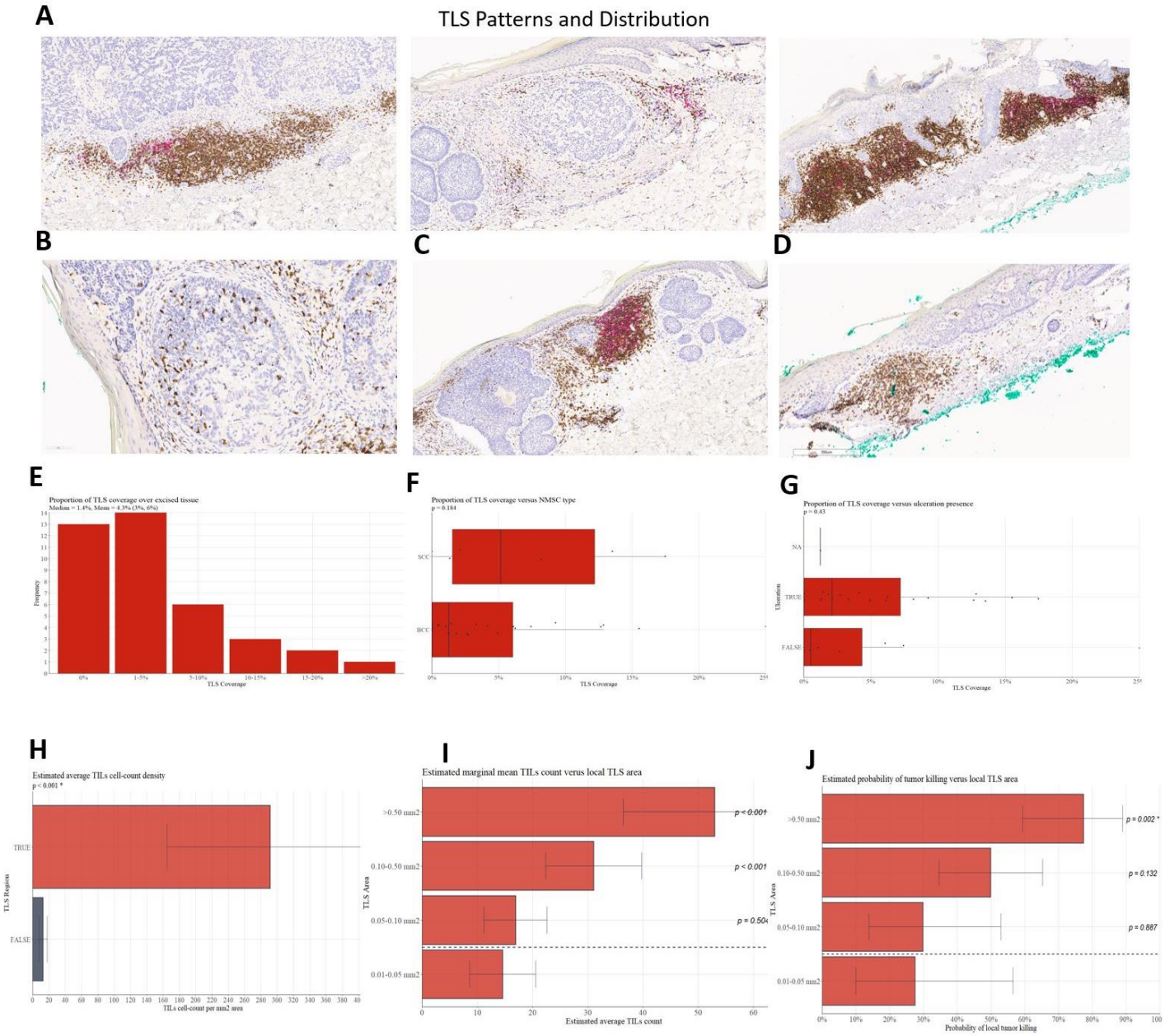



**Fig S3. Tertiary lymphoid structures abound in non-melanoma skin cancers and potentially represent anti-tumor immune response manifesting as higher TILs and tumor killing**

**A.** Representative images from CD3/CD20 staining of non-melanoma skin cancers (BCC and SCC) demonstrates differences in size, patterns and frequency of TLS in different patients. In addition to quantification of CD3, CD20 and total positive areas using software-based approaches, manual evaluation for number and size of TLS, local TILs and tumor killing were noted for a more spatial and context-dependent analysis; **B.** Representative image highlighting the high number of tumor-infiltrating CD3+ cells inside BCC tumor nest adjacent to TLS (bottom); **C.** Large TLS surrounding remains of a tumor as confirmed by H&E; **D.** Less TILs and no tumor killing was observed in non-TLS patients and regions; **E.** Relative distribution of TLS in 40 NMSC patients suggests a mean area coverage of 4.3% (CI: 3-6%); **F.** Higher TLS coverage in SCC as compared to BCC although the results did not achieve significance possibly because of

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

lower SCC cases in this cohort(p=0.18, ns); **G.** Although TLS presence corresponded with ulceration in most cases, the results were not significant (p=0.43, ns); **H.** TLS-positive areas demonstrated higher TILs as compared to TLS-negative areas (p<0.001) indicating an anti-tumor response; **I.** Higher TLS coverage (>0.1 mm2) correlated with higher TILs (p<0.001); **J.** Higher TLS area coverage was associated with a statistically higher probability of tumor killing.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

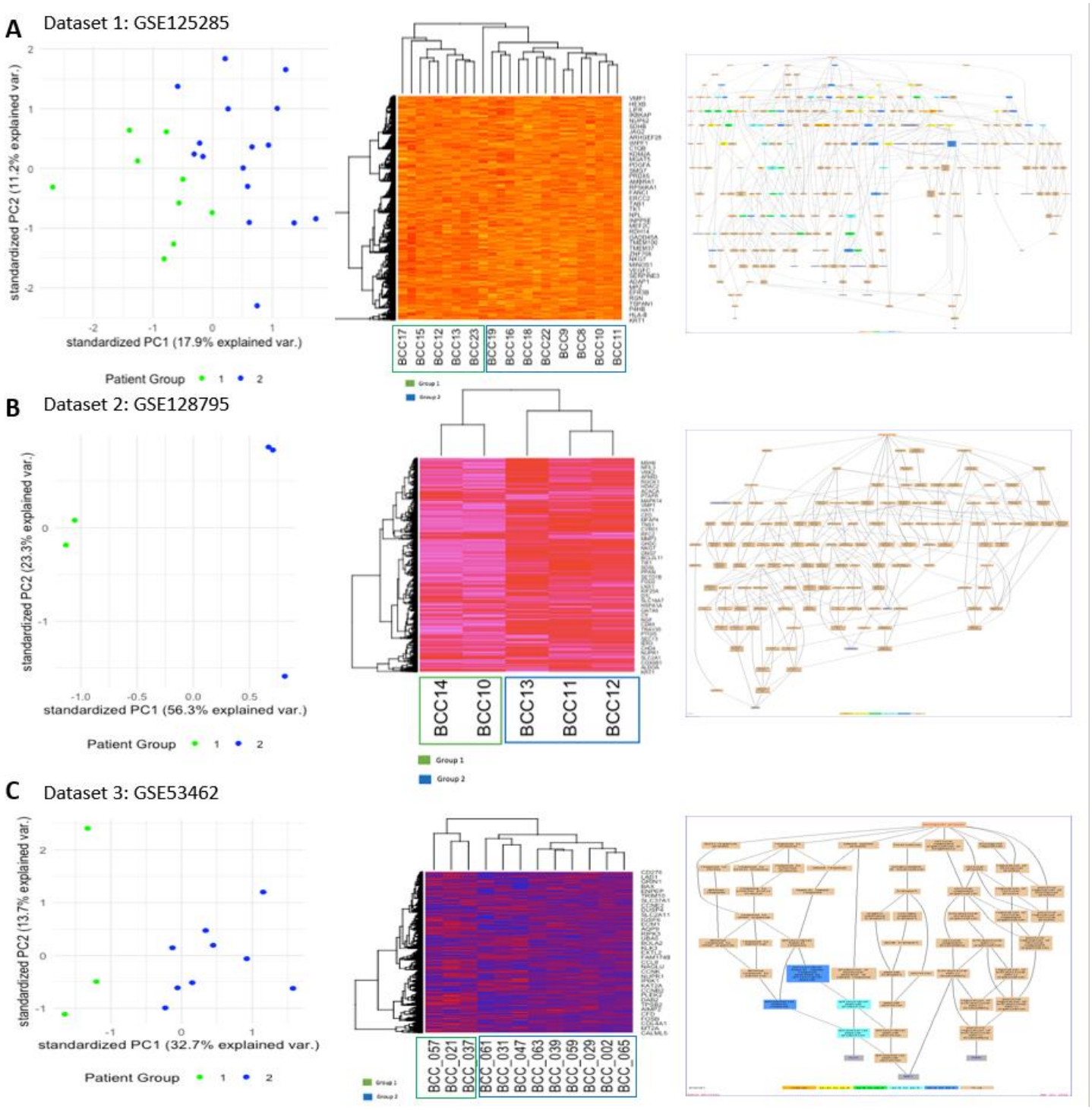



**Fig S4. Gene expression analysis on 3 distinct GEO datasets reveal phenotypes on Master Gene Panel (gene panel consisting of cancer-associated and immune-associated genes).**

**A.** PCA on GEO dataset GSE125285 Wan et al (n= 25 BCC) reveals presence of at least 2 phenotypes (green and blue). Unsupervised hierarchical cluster analysis 2 groups also results in two phenotypes. Gene ontology on the DGEA demonstrate the following enriched (FDR < 0.01) GO terms (between green and blue) in the top 10: immune system

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

process, cell surface receptor signaling pathway, lymphocyte activation, T cell activation, positive regulation of response to stimulus, leukocyte activation, positive regulation of immune system process, regulation of immune system process, and cell activation. **B.** PCA on GEO dataset GSE128795 Sand et al (n=5 BCC) also highlights distinct immune phenotypes. Top 4 GO Terms enriched (FDR <-0.05) in green vs blue group: leukocyte homeostasis, regulation of lymphocyte activation, cell adhesion, and biological adhesion.

**C.** PCA on GEO dataset GSE53462 (Jee et al) (n=12 BCC) also again reveals presence of at least 2 phenotypes (green and blue). Unsupervised hierarchical cluster analysis of 2 groups also results in two phenotypes. Gene ontology on the 5 DGEA ( |log2FC| > 0.3785 and FDR < 0.05) demonstrate the following enriched (FDR < 0.01) GO terms (between green and blue) in the top 5: immune cell processes, antibacterial peptide production, antibacterial humoral response, antimicrobial humoral immune response mediated by antimicrobial peptide, extracellular matrix disassembly, and antimicrobial humoral response, perhaps suggestive of tertiary lymphoid structure and leukocyte trafficking heterogeneity in these patient clusters. *Only the subset of BCCs that did not display a “normal-like” or “SCC-like” gene expression pattern were kept in the analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

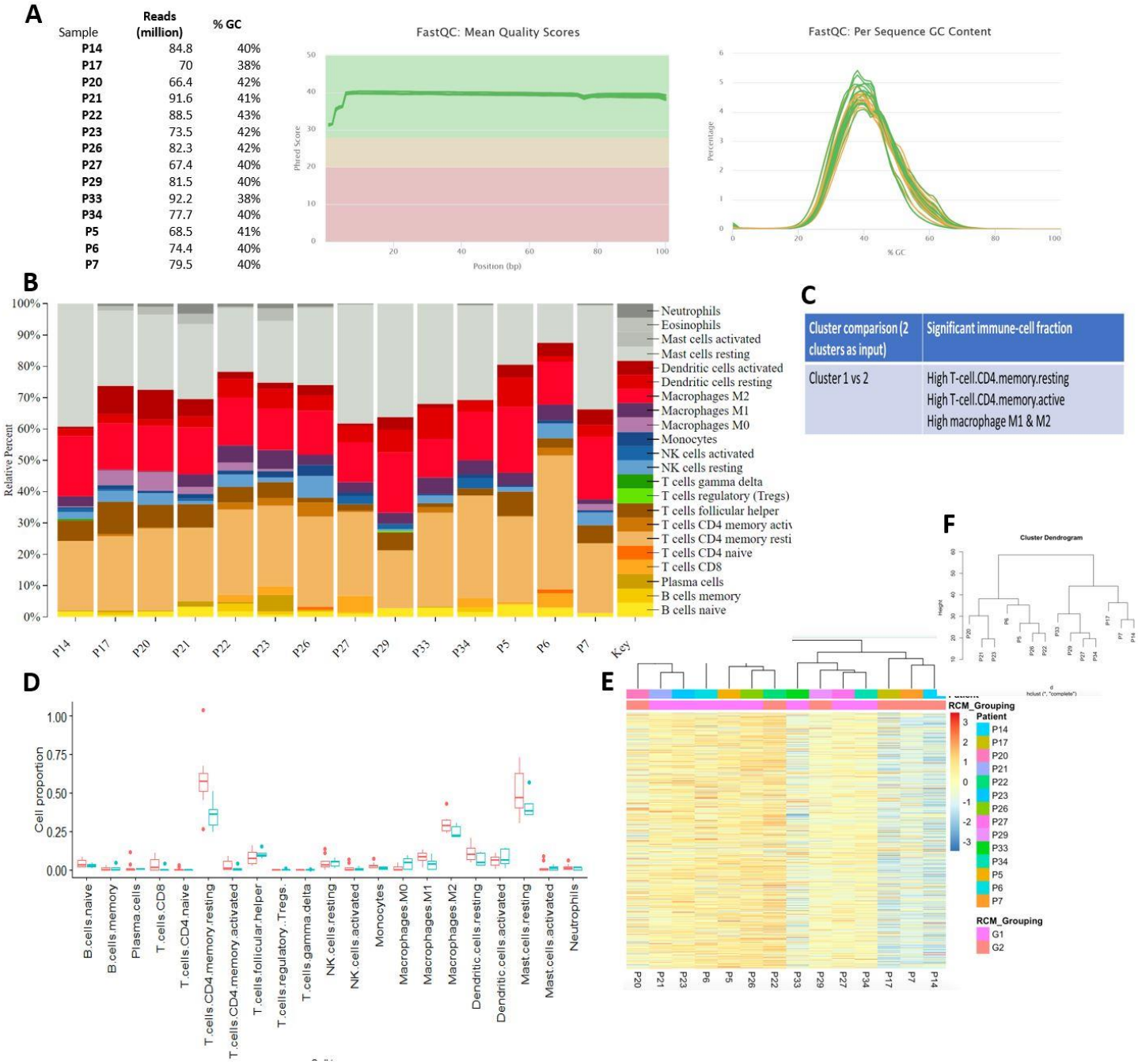


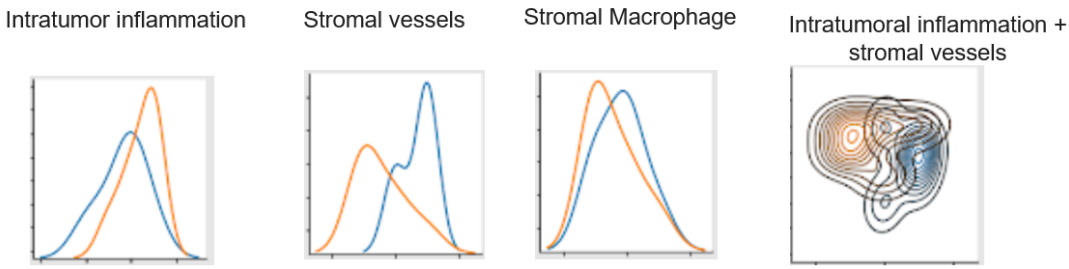

**Fig S5. Gene expression analysis using bulk RNA-seq demonstrates heterogeneity in gene expression**

**A.** Overall QC analysis suggests high quality and reliable transcriptomic data which was subsequently tested using different analytical tools; **B.** CIBERSORT immune profiling was performed to explore differences in immune sub-fractions. Results on 14 patients indicate presence of resting mast cells, M2 macrophages and T-follicular helper cells and CD4 memory resting T cells in all patients although their relative proportions vary. Plasma cells, CD8 T cells, neutrophils and M1 macrophages were seen in some patients; **C, D.** Based on RCM classification of red and blue phenotypes, statistically significant immune cells differences between these groups included CD4 memory resting and memory active cells, M1 and M2 macrophages; **E, F.** HCA on Nanostring Pancancer Immune Panel Genes followed by dendrogram

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

indicates presence of four phenotypes of which P7, P14 and P17 ($Vasc^{high}Inf^{low}$) were found to have lower expression of immune genes and P5, P6, P21, P23, P26 ($Inf^{high}$) showed higher expression of immune panel genes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



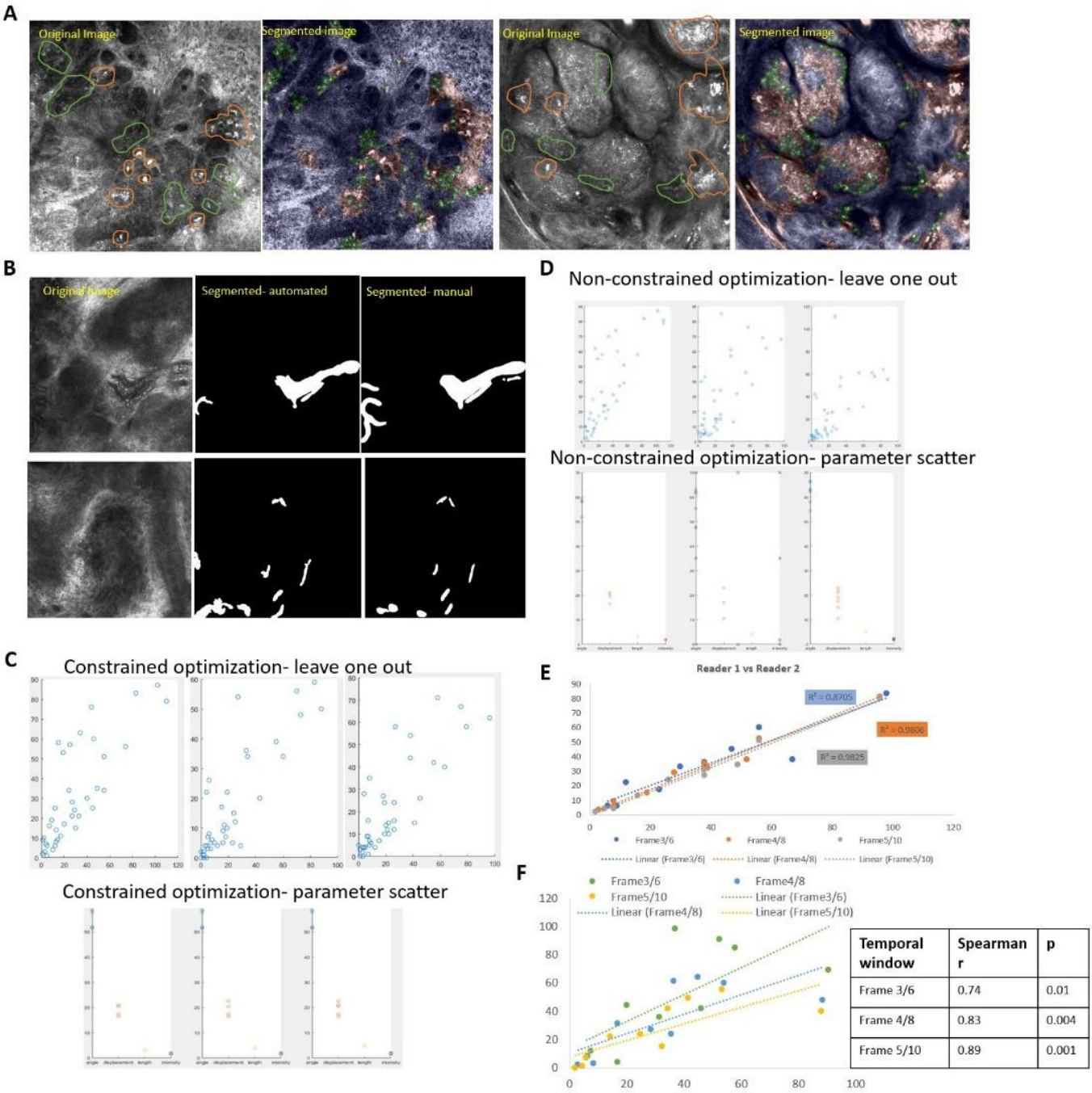

```
SPECIFICITY CRITERION
[2, 0, 3, 9, 10, 13, 16, 14, 5, 6, 4, 7, 18, 8, 17, 1, 15, 11, 12, 19]
[0.85714286 0.85714286 0.85714286 0.85714286 0.85714286 0.85714286
 0.85714286 0.85714286 0.71428571 0.71428571 0.71428571 0.71428571
 0.71428571 0.71428571 0.71428571 0.71428571 0.71428571 0.71428571
 0.71428571 0.71428571]
[0.66666667 0.66666667 0.66666667 0.66666667 0.66666667 0.66666667
 0.66666667 0.66666667 0.66666667 0.66666667 0.66666667 0.66666667
 0.66666667 0.66666667 0.66666667 0.66666667 0.66666667 0.5
 0.5        0.5       ]
[-15.06238189 -15.06238189 -15.06238189 -15.06238189 -15.06238189
 -15.06238189 -15.06238189 -15.06238189 -11.32251495 -11.32251495
 -11.32251495 -11.32251495 -11.32251495 -11.32251495 -11.32251495
 -11.32251495 -11.32251495  -8.42164879  -8.42164879  -8.42164879]

AIC CRITERION
[2, 5, 0, 4, 6, 7, 11, 13, 1, 16, 18, 8, 19, 3, 12, 14, 10, 15, 9, 17]
[0.85714286 0.71428571 0.71428571 0.71428571 0.71428571 0.71428571
 0.71428571 0.71428571 0.71428571 0.71428571 0.71428571 0.71428571
 0.71428571 0.57142857 0.57142857 0.57142857 0.57142857 0.57142857
 0.57142857 0.71428571]
[0.66666667 0.83333333 0.83333333 0.83333333 0.83333333 0.83333333
 0.83333333 0.83333333 0.83333333 0.83333333 0.83333333 0.83333333
 0.83333333 0.66666667 0.66666667 0.66666667 0.5        0.5
 0.33333333 0.5       ]
[-15.06238189 -15.06238189 -15.06238189 -15.06238189 -15.06238189
 -15.06238189 -15.06238189 -15.06238189 -15.06238189 -15.06238189
 -15.06238189 -15.06238189 -15.06238189  -8.42164879  -8.42164879
  -8.42164879  -6.05146855  -6.05146855  -4.04750971  -8.42164879]
```

**Figure S6. Modeling of imiquimod response highlights importance of stromal features in improving predictive power of linear regression models.** Top row- linear separability curves for intratumor inflammation, stromal vessels and stromal macrophages (vessels and macrophages were predicted as priority labels in AIC and specificity prioritizing models) and combined plot for stromal vessels + intratumoral inflammation shows maximum separability (orange-responders, blue-non-responders).

Bottom row: Models based on prioritizing Specificity or AIC as criteria for selecting features associated with response in linear regression models. For specificity, sequential modeling using 13 factors gave optimum performance while for AIC models the best performance was seen at 8 features.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




| Temporal window | Spearman r | p |
|---|---|---|
| Frame 3/6 | 0.74 | 0.01 |
| Frame 4/8 | 0.83 | 0.004 |
| Frame 5/10 | 0.89 | 0.001 |

**Figure S7. Automated quantification of TiME and vasculature features on RCM images and correlation with TiME features**

**A.** Exemplar RCM images (original and segmented image) showing segmentation of leukocytes (green), and dendritic and macrophages (orange) using a UNet model trained on 1026 images (926-training, 100-validation) yielding a Dice coefficient of 0.72. This model was used to segment immune cell density on 652 independent images used in TiME analysis. **B.** Exemplar RCM video frames (original, manually segmented, automated segmented) showing segmentation of blood vessels from single field-of-view RCM videos. Comparison of manual and automated segmentations on 23 videos yielded a median Dice coefficient of 0.52 (0.29-0.78). This segmentation was used to derive vasculature features such as

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

blood vessel diameter, area and number of vessels used in TiME analysis. **C.** Summary of optimization for leukocyte trafficking done using constrained optimization of parameters: leave-one out and parameter scatter. **D.** Summary of optimization for leukocyte trafficking done using non-constrained optimization of parameters: leave-one out and parameter scatter. **E.** Agreement between manual counts by two independent readers on 10 videos. **F.** High Spearman correlation (0.74-0.89) observed for average manual reader count and automated counts for all 3 temporal windows. Highest correlation was observed for the longest temporal window since correlation is lower for faster events such as rolling in shorter vessels. 3 frame (0.6 s) temporal window selected for TiME analysis to include both fast and slow trafficking events.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447835; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

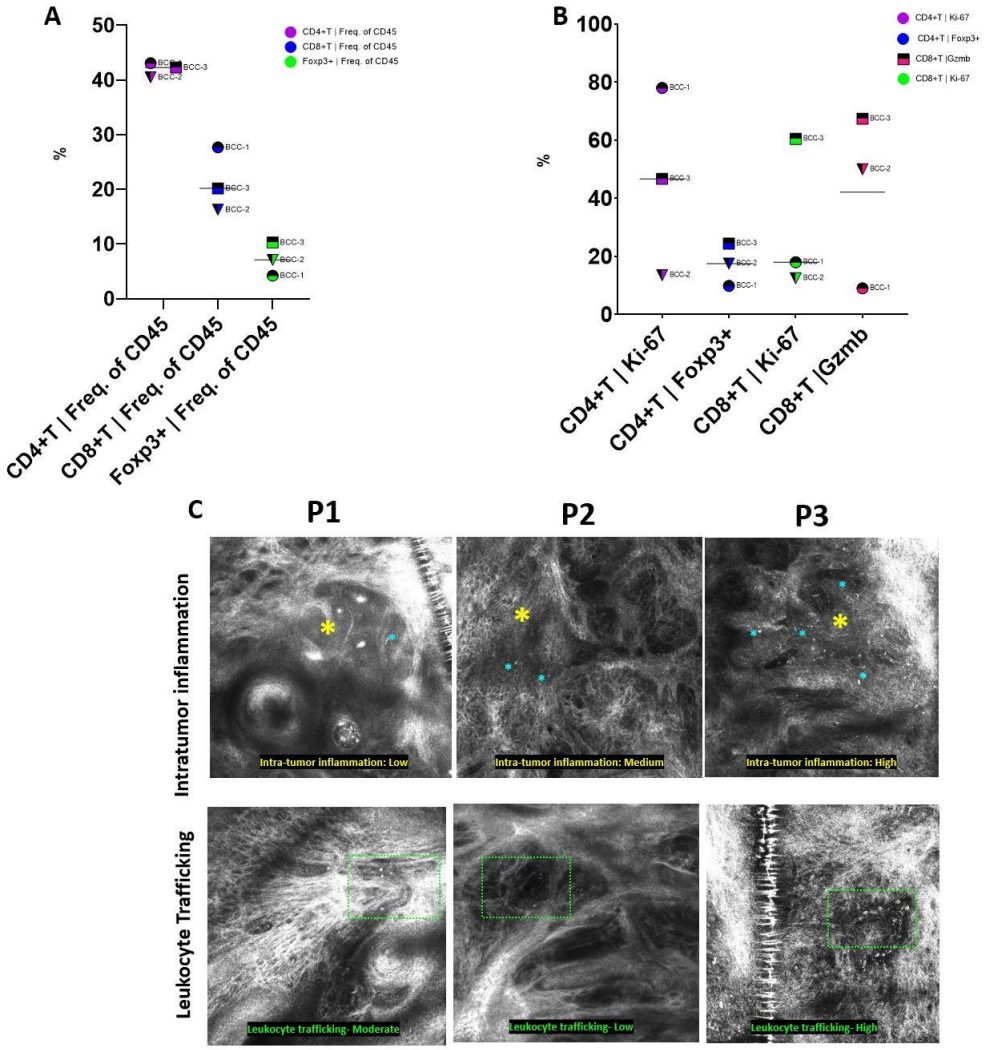



**Fig S8. Flow-based immunophenotyping for understanding T-cell distribution in red and blue RCM phenotypes.** **A.** Higher CD8 and lower Foxp3 T-cell populations in BCC-1 (RCM blue) as compared to BCC-3 (red), **B.** Higher activated and proliferating populations of CD8 T cells in BCC-3 ($Vasc^{high}Inf^{high}$) while BCC-1 ($Vasc^{high}Inf^{low}$) showed higher proliferating CD4 T cells; **C.** Representative RCM images from BCC1, 2 and 3 patients highlighting the differences in intratumoral inflammation, trafficking and vasculature between the $Vasc^{high}Inf^{high}$ (BCC-1) and $Vasc^{high}Inf^{low}$ (BCC-3) patient.