Phenotyping Immune Cells in Tumor and Healthy Tissue Using Flow Cytometry Data

Ye Chen
Department of Computer Science
Purdue University
chen1385@purdue.edu

Ryan D Calvert
Department of Nutrition Science
Purdue University
ryan@purdue.edu

Ariful Azad
Lawrence Berkeley National Lab
azad@lbl.gov

Bartek Rajwa
Bindley Biosciences Center
Purdue University
brajwa@purdue.edu

James Fleet
Department of Nutrition Science
Center for Cancer Research
Purdue University
fleet@purdue.edu

Timothy Ratliff
Dept. Comparative Pathobiology
Center for Cancer Research
Purdue University
tlratliff@purdue.edu

Alex Pothen
Department of Computer Science
Purdue University
apothen@purdue.edu

ABSTRACT

We present an automated pipeline capable of distinguishing the phenotypes of myeloid-derived suppressor cells (MDSC) in healthy and tumor-bearing tissues in mice using flow cytometry data. In contrast to earlier work where samples are analyzed individually, we analyze all samples from each tissue collectively using a representative template for it. We demonstrate with 43 flow cytometry samples collected from three tissues, naive bone-marrow, spleens of tumor-bearing mice, and intra-peritoneal tumor, that a set of templates serves as a better classifier than popular machine learning approaches including support vector machines and neural networks. Our "interpretable machine learning" approach goes beyond classification and identifies distinctive phenotypes associated with each tissue, information that is clinically useful. Hence the pipeline presented here leads to better understanding of the maturation and differentiation of MDSCs using high-throughput data.

KEYWORDS

Flow cytometry, templates, template-based classification, myeloid-derived suppressor cells (MDSC)

ACM Reference Format:
Ye Chen, Ryan D Calvert, Ariful Azad, Bartek Rajwa, James Fleet, Timothy Ratliff, and Alex Pothen. 2018. Phenotyping Immune Cells in Tumor and Healthy Tissue Using Flow Cytometry Data. In *ACM-BCB'18: 9th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, August 29-September 1, 2018, Washington, DC, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3233547.3233583

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only.
ACM-BCB'18, August 29-September 1, 2018, Washington, DC, USA
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5794-4/18/08…$15.00
https://doi.org/10.1145/3233547.3233583

1 INTRODUCTION

Can algorithms distinguish the phenotypes of myeloid-derived suppressor cells (MDSC) in a tumor environment from MDSCs present in healthy tissues in mice using flow cytometry (FC) data? We answer this question in the affirmative. We consider three different tissues in mice, and identify distinct immune cell populations that are present in each tissue by first clustering cells in each flow cytometry sample and then registering these cell populations across the samples to form meta-clusters. We build templates, i.e., collections of meta-clusters in the samples to provide a summary representation of each tissue, and then classify new samples to the closest template. Central to our approach is a combinatorial measure of dissimilarity between two clustered samples computed using the flowMatch algorithm [3] from a mixed edge cover of minimum weight in a bipartite graph.

Inflammation and cancer affect the formation of cells in the bone marrow by altering myelopoiesis toward immunosuppressive cells and reducing maturation of innate immune cells. The heterogeneous collection of immature myeloid cells that can potentially suppress the T-cell response at the tumor are designated MDSCs [5]. In mice, monocytic MDSCs (M-MDSC; $CD11b^{+}Ly6C^{hi}Ly6G^{-/lo}$) and granulocytic MDSCs (G-MDSC; $CD11b^{+}Ly6C^{int}Ly6G^{hi}$), are two MDSC subtypes that can be separated using the markers above via flow cytometry (FC) [5]. Since MDSCs are associated with tumor progression and increased metastasis, these cells are candidate targets for immune therapy.

In this study, we consider MDSC cells from three tissues: naive bone marrow (NVBM), tumor-bearing spleen (TBSP), and intraperitoneal tumors (IPTM). We have shown that the MDSC cells from TBSP are a less mature phenotype that lack immediate T-cell suppressive ability but they gain this function during a 3-day suppression assay [4, 8]. In contrast, tumor MDSCs possess immediate T-cell suppressive ability [4, 8].

We make the following contributions in this paper:

(1) **New insights into MDSC phenotyping:** We use new MDSC data to classify samples in three different tissue types using all myeloid cells, M-MDSCs only, and G-MDSCs only. We also identify the impact of cell types on correct classification and phenotyping.
(2) **New computational cytometry approach:** We apply a machine learning (ML) algorithm developed in earlier work, flowMatch [3], to classify MDSCs and allow human-interpretable identification of cell populations in each tissue as meta-clusters. The template-based classification is more accurate than traditional ML techniques such as support vector machines (SVM) or neural networks. The flowMatch software is available as a Bioconductor package (DOI: 10.18129/B9.bioc.flowMatch).
(3) **Automation of FC analysis:** We present a fully automated pipeline to classify samples and identify MDSC phenotypes that includes: data transformation, gating and clustering, cluster matching and template formation, and finally, sample classification and phenotype discovery. This is the first fully automated phenotypic analysis of MDSCs.

2 METHODS

2.1 Description of dataset

The data were drawn from an archive of FC data-sets of ~60 individual experiments conducted over a period of 2 years and representing 94 biological samples. Samples were generated using reagents and methods described previously by our group [4].

Inclusion criterion for samples in the FC file dataset: Each data file was inspected using FlowJo v.10 software (FlowJo LLC; Ashland, OR). Samples that met the following criteria were included in the final dataset: (a) a minimum of 10,000 recorded events, (b) available data for FSC-A, SSC-A, FSC-H, FSC-W, SSC-H, SSC-W, CD11b, Ly6C and Ly6G, (c) use of a common fluorescent compensation protocol, and (d) complete separation of the fluorescent signals for each fluorochrome pair (APC, PE, and PE-Cy7). The quality of each sample was defined with traditional biplots: i.e. (a) FSC-A x SSC-A, (b) PE-Cy7-CD11b x SSC-A, (c) PE-Ly6G x APC-Ly6C [4]. Only samples where positive and negative populations of cells were clearly defined on a specific axis were included in the final dataset (i.e. FSC and SSC (neg: <30, pos. >40 on linear scale) and each fluorochrome (neg: 0-102, pos. 103-104 on log scale)). The final dataset for testing included 10 NVBM samples, 13 TBSP samples, and 20 IPTM samples.

2.2 Automated gating and clustering

In most FC studies, the subset of cells for downstream analysis is selected by manual gating enabled by visual inspection of bivariate plots. Alternatively, the important cell populations can be identified by automated clustering algorithms.

Automated gating. First, we use the biexponential/logicle function [11] to perform heteroskedasticity correction. For this purpose, we employed the R library `flowCore` and estimated the parameters of the transformation running the `estimateLogicle` function.

We automated the initial gating steps selecting viable, single cells, and removing doublets, cellular debris, and clumps (Figure 3). Myeloid cell selection and MDSC subtype selection were done based on fluorochrome intensity. Both M-MDSC and G-MDSC subtype selection used the known location of these cell types within the Ly6C and Ly6G biplot [12]. Our approach to automated gating followed a similar procedure to the one offered by the `flowDensity` R-Bioconductor package, and is based on recognizing regions of high local densities and significant curvatures [9].

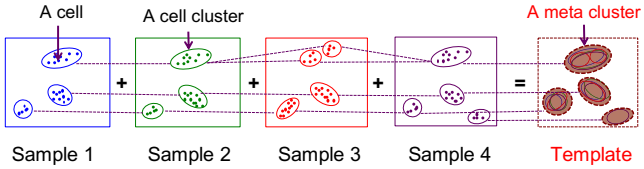


Figure 1: Creating a template from four FC samples. The union of registered clusters forms a meta-cluster, and a collection of meta-clusters defines a template.

Automated clustering. After a series of gating steps, we reduced the dimensionality of the samples from 9-D to 3-D. We clustered each of these gated samples to identify subtypes of cells, using the k-means algorithm. To identify the number of clusters k_{opt}, we used the minimum value of the `S_Dbw` index [7]. In our experience, [3] the S_Dbw index works better with high-dimensional samples than some specialized implementations of k-means, such as flowMeans [1]. However, the S_Dbw index tends to favor k-values which are too high when used with preselected MDSCs, so in this case, we used a combination of three cluster validation criteria: average Silhouette width, Caliński-Harabasz index, and Dunn index. We assume that cells in a p-dimensional cluster are normally distributed, and hence can be statistically summarized by two distribution parameters $\boldsymbol{\mu}$, the p-dimensional mean vector, and Σ, the $p \times p$ covariance matrix.

2.3 Registering cell clusters across samples

After clustering each FC sample independently, we match phenotypically-similar clusters across samples, a process commonly known as *cluster registration*.

To match cell clusters, we used a robust variant of a graph matching algorithm called the Mixed Edge Cover (MEC), that allows a cluster in one sample to be matched with zero, one, or more clusters in the second sample [2]. This is a different approach from those taken by other FC algorithms utilizing meta-clusters [6, 10]. The algorithm to compute an optimal MEC was developed in our prior work [2]. Here, we briefly summarize it for completeness.

The MEC algorithm initially creates a bipartite graph from a pair of samples, where vertices in each part represent clusters from a sample. A pair of vertices (clusters) is connected by an edge whose weight is computed by the Mahalanobis distance between a normally distributed pair of clusters. Next, a modified minimum-weight edge cover is computed on the bipartite graph. The resultant solution matches clusters across samples, while possibly leaving a small number of clusters unmatched. For each unmatched cluster we pay a penalty λ, which ensures that the number of such clusters remains small. The cost of a mixed edge cover is the sum of weights of all matched edges and the penalties due to the unmatched clusters. An optimal MEC can be computed in $O(k^3 \log k)$ time where k is the maximum number of clusters in a sample [2].

2.4 Algorithms to create templates

We previously designed an algorithm to create a template from a group of similar samples [3]. First, we create a template from all samples of the same class (Figure 1) by registering clusters in these samples using the MEC algorithm, and merging the matched clusters into meta-clusters (i.e., groups of phenotypically similar clusters that have smaller distances among themselves relative to their distance to other clusters). In Figure 1, the matched clusters are shown joined by dashed lines, and these are merged to form meta-clusters in the rightmost subfigure. Similar to the statistical representation of a cluster, we model a meta-cluster by a Gaussian distribution along with its mean and covariance matrix. Finally, a collection of meta-clusters defines a template of a class of samples as shown in Figure 1. The same statistical representation of a sample and a template with a Gaussian mixture model enables seamless matching of clusters and meta-clusters via the MEC algorithm.

To simplify the registration of clusters across many samples, the algorithm iteratively merges the most similar pair of samples. The similarity is computed by the cost of the optimum mixed edge cover. The pair of samples considered in the current iteration is merged to create an intermediate template. The algorithm for creating templates organizes samples in a binary tree, and merges a pair of samples and intermediate templates that are the most similar in the current round.

2.5 Template-driven classification

Given a group of FC samples belonging to m classes, we build m templates, $T_1, T_2, \ldots, T_m$, one for each class. A new sample S is compared with each of the templates via the MEC algorithm and is predicted to belong to the class whose template it is most similar (least dissimilar) to. This approach is a template-driven nearest neighbor classification, which is more robust than sample-driven nearest neighbor classification. The template-based classification requires m MEC computations, one with each template, making it significantly faster than nearest-neighbor classification that requires a new sample to be compared with all existing samples.

Fully automated classification pipeline. The key to a successful automation pipeline (Figure 2) lies in the parameter optimization steps, since the parameters depend entirely on the data distribution of input FC samples. Two parameters that significantly influence the classification result are: the number of clusters (k) in k-means, and the unmatch penalty (λ) in the MEC algorithm. The variable k defines how many distinct cell populations we believe are reasonable in a given sample, and λ determines whether two clusters should be matched or left unmatched in a pair of samples. We described our strategy to select k in Section 2.2. Two approaches are recommended for selecting λ. First, vary the unmatch penalty and find the inflection point when the number of meta-clusters in the template decreases dramatically and thereafter remains constant [2]. Second, use a multiplier (> 2) of the maximal standard deviation (estimated by the square root of the trace of the covariance matrix) of clusters in the samples. The idea behind this approach is to unmatch the clusters if they do not significantly intersect with each other in multi-dimensional feature space.

Template evaluation and classification quality score. Due to the small dataset (43 samples) we use the Leave One Out Cross Validation (LOOCV) technique to evaluate the prediction accuracy of the templates. In turn, we make each sample the test sample, create templates using all other samples, and then classify the test sample to the nearest template.

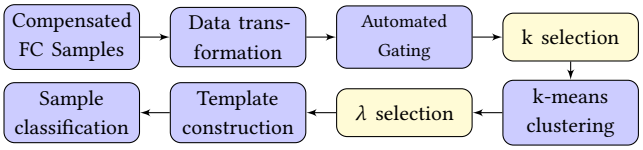


Figure 2: Pipeline of automated template-based classification. Purple blocks denote core steps, and yellow blocks indicate parameter optimization steps.

To evaluate the classification confidence of the samples, we define a classification quality score by the following formula

$$\text{classification_quality_score} = 1 - d_t / \min\{d_a, d_b\}, \qquad (1)$$

where d_t is the dissimilarity of the test sample and the true template of the tissue it belongs to, and d_a and d_b refer to the dissimilarity of the test sample and the other two templates, respectively. (Recall that since we are doing LOOCV, we know the true template for each sample.) Hence a positive score indicates a correct prediction, while a negative score identifies a misclassification. The higher the classification quality score is, the more confidence we have in classifying the sample.

3 RESULTS

3.1 Selecting myeloid cell subsets

To identify different myeloid cell subtypes, we applied a sequence of gates as shown in Figure 3. The boundary of each gate is computed as described in Section 2.2, and we compared these with manual gating patterns across samples to confirm that our method selected commonly accepted ranges for each gate.

When we consider all myeloid cells, we cluster cells using the k-means algorithm after the myeloid gate is applied, but not the M-MDSC and G-MDSC gates (see Figure 3). In contrast, to investigate M-MDSCs and G-MDSCs in isolation, we gate on Ly6C and Ly6G to identify these subtypes and then apply k-means clustering. In both cases, we cluster cells in three dimensional marker space (CD11b, Ly6C and Ly6G).

After we gate and cluster every sample individually, we group all samples collected from a tissue and create a template for that tissue type (NVBM, TBSP, and IPTM). Furthermore, to investigate different cell types in each tissue, we create templates using three different set of cells: all myeloid cells, M-MDSCs only, and G-MDSCs only. Hence we construct nine templates as shown in Figure 4. To compare the templates, we summarize the meta-clusters within each template with respect to the following four aspects: number of meta-clusters, fraction of cells in all myeloid cells, average expression level of markers Ly6G and Ly6C, and average standard deviation of the meta-clusters (Table 1). The average expression level of markers measures the approximate position of meta-clusters in the 2-D template plot, and the average standard deviation estimates the spatial extent of each meta-cluster.

(a) Myeloid cells templates. From the myeloid cells templates, for the three tissues we observed similar meta-cluster distribution patterns that contain the M-MDSC range at top left, the G-MDSC

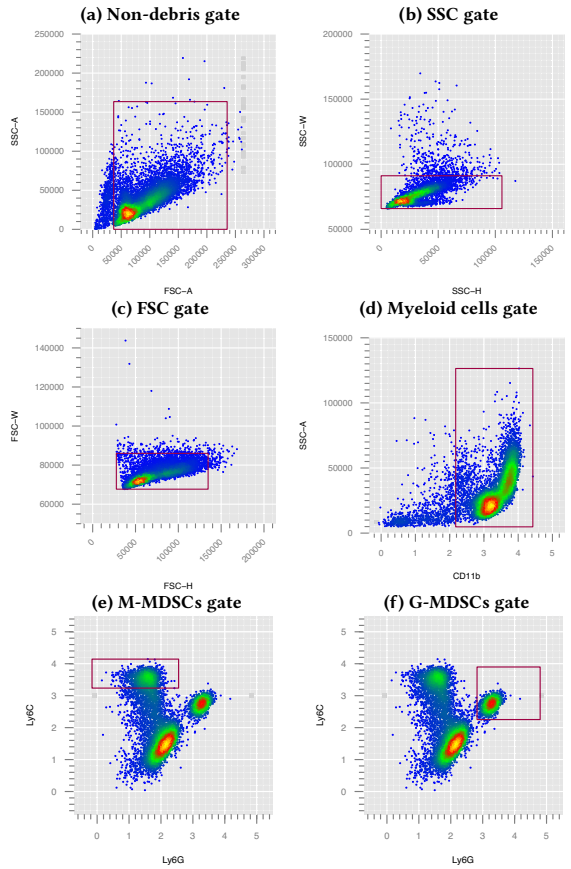


Figure 3: The sequence of gates applied to FC data to select MDSC subtypes. The red rectangle in each plot shows the boundaries for the gates. Gates listed in Subfigures (a) to (d) are applied in sequence to select myeloid cells, followed by individual gating of the M- or G-MDSC subtypes.

range at the middle right, and other cell populations at left bottom (Figure 4a). However, significant differences exist across templates. As our primary interest is in MDSCs, we analyze the meta-clusters based on the areas: M-MDSC, G-MDSC, and 'Others'. We consider only the meta-clusters which reside entirely within each subtype range as G- or M-MDSC meta-clusters, and those partially involved in the subtype range are considered in the 'Others' category. Because of the large number of meta-clusters identified in each template, only significant meta-clusters are shown in the table and in the template plots. A meta-cluster is *significant* if it exists in more than one sample or has > 1% of myeloid cells. In general, NVBM and TBSP have fewer meta-clusters than IPTM, especially within the 'Others' area. Numbers and patterns of meta-clusters within the M- and G-MDSC range appear similar across tissues, but a slight shift of meta-clusters can be observed: for example, meta-clusters in M-MDSC shift upward (Ly6C increases) from NVBM to IPTM, while TBSP tends to have lower Ly6C and Ly6G expression in G-MDSC.

Although a few common meta-clusters that appear in all tissues are identified in the templates, we also discover many unique meta-clusters that exist in only one tissue. A large number of myeloid cells are selected by the gating procedure, but the M and G subtype gates represent a much smaller fraction of the meta-clusters. Therefore, we consider the gating of the M and G subtypes across samples to investigate if flowMatch would classify the tissues based solely on these subtypes.

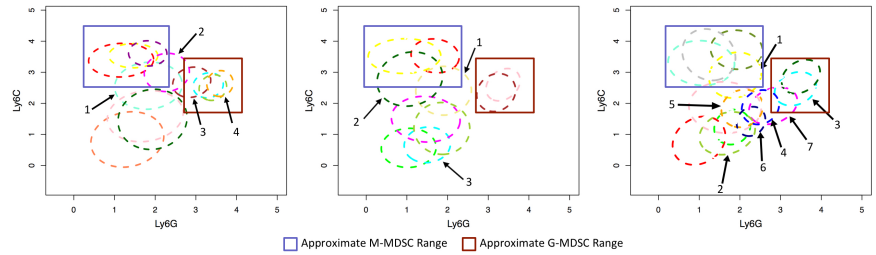
(b) M-MDSC templates. We observe that IPTM is the most distinctive template among all, with the highest number and the broadest distribution of meta-clusters (Figure 4b). Table 1 shows that we observe a large gap in the number of meta-clusters in M-MDSC templates, which provides a good separation between tissues. We include all meta-clusters (including insignificant ones) in the table to avoid losing information with few meta-clusters. The expression level of both Ly6C and Ly6G increase from naive samples to tumor samples, which shows up as an upward and rightward shift of meta-clusters in the template plots. Unlike the Myeloid cells templates, the average spatial size of meta-clusters remains similar across tissues. We observe an increase of fraction of MDSCs from NVBM to TBSP, then followed by a decrease of M-MDSCs in IPTM. The two significant meta-clusters in NVBM template are also found, one each, in the TBSP and IPTM templates, and hence there is no significant unique meta-cluster in the NVBM template.

(c) G-MDSC templates. The G-MDSC templates reveal divergent patterns. Firstly, the location of meta-clusters is changed across tissues: NVBM meta-clusters aggregate at the middle left, TBSP meta-clusters are centered at the bottom right, while in IPTM, they are distributed along the diagonal line, which can be confirmed from the average expression level of markers (Table 1). Secondly, the number of meta-clusters dramatically increases in the IPTM template. G-MDSCs account for over 60% of myeloid cells in NVBM, and decrease from TBSP to IPTM. Additionally all meta-clusters are unique in the corresponding tissue, indicating that G-MDSC templates are more distinctive than M-MDSC templates.

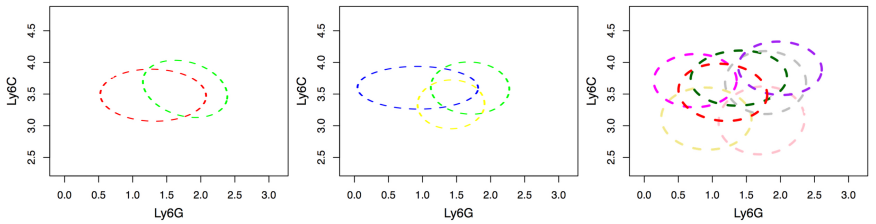
Consistent with other research, it was found that in all cases, there are more (or similar numbers of) MDSCs in NVBM samples compared to the IPTM samples. This is most likely due to NVBM enrichment of myeloid-precursor cells that express the selected markers, whereas IPTM consists of a broader mix of immune cell types found in the tumor microenvironment.

3.2 Classifying samples using templates

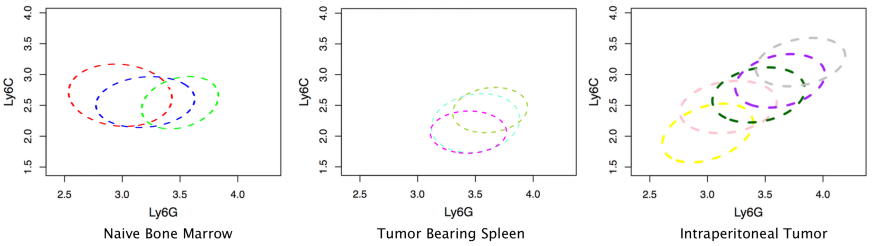
We show the classification quality score of each sample when using different sets of templates in Figure 5. With each of the three template sets: myeloid cells templates, M-MDSC templates, and G-MDSC templates, we observe that the flowMatch algorithm correctly classifies all samples to their true tissue. Compared with the others, G-MDSC templates have the highest average score when classifying NVBM samples (Figure 5c), while regarding TBSP and IPTM, myeloid cells templates perform the best, followed by M-MDSC templates and G-MDSC templates. Surprisingly, M-MDSC templates show higher classification confidence when classifying TBSP compared with G-MDSC templates (Figure 5b). Observing the distribution of meta-clusters, TBSP samples are more distinct in G-MDSCs than in M-MDCSs. One possible reason could be the higher number of IPTM meta-clusters in M-MDSC templates, making it easier to distinguish IPTM from the TBSP. As we expected, myeloid templates have the highest average classification confidence (Figure 5a), because they capture the characteristic distribution of both G-MDSCs and M-MDSCs, and also other myeloid cells.



(a) Meta-clusters in Myeloid Cell Templates.

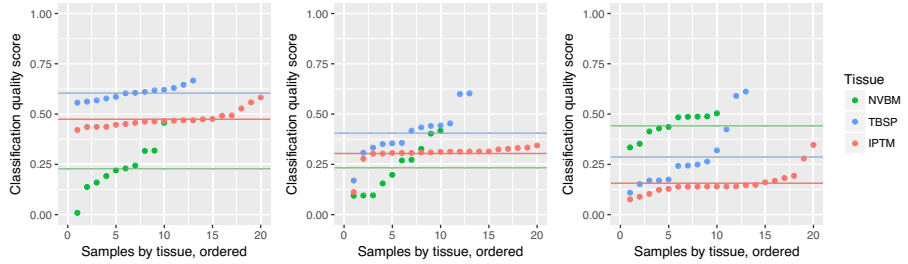


(b) Meta-clusters in M-MDSC Templates



(c) Meta-clusters in G-MDSC Templates

Figure 4: Templates for NVBM (column 1), TBSP (column 2), and IPTM (column 3) illustrated by their meta-clusters. Each meta-cluster is represented by a contour that includes 95% of the cells within it. Templates are built with a) all myeloid cells, b) M-MDSCs only, and c) G-MDSCs only. These are 2-D plots obtained from a collection of 3-D contour plots. (a). Purple and red rectangles indicate the approximate areas of M-MDSC and G-MDSC, respectively. Unique meta-clusters in each template are indicated by arrows with indices specified. b) and c). Distinct meta-clusters are represented by different colors, with matched meta-clusters represented by the same color across templates. Meta-clusters that are present in only one sample are omitted for simplicity.



(a) Myeloid Cells Templates (b) M-MDSC Templates (c) G-MDSC Templates

Figure 5: The classification quality score of samples classified by nearest templates using LOOCV, using G-MDSCs only, M-MDSCs only, or all myeloid cells. Horizontal lines indicate the average classification quality score of each template. Higher score indicates higher confidence in the prediction.

Tissue		Myeloid cells templates			M-MDSC templates	G-MDSC templates
		M-MDSC range	G-MDSC range	Others range		
NVBM	Fraction in Myeloid Cells	19.23%	49.27%	28.54%	18.55%	62.12%
	Avg. Expression Level (Ly6G, Ly6C)	(1.48, 3.51)	(3.44, 2.57)	(2.01, 2.02)	(1.25, 3.49)	(3.19, 2.56)
	Avg. Standard Deviation	0.15	0.07	0.29	0.16	0.11
	Number of MC	3	3	6	3	4
TBSP	Fraction in Myeloid Cells	19.87%	25.41%	54.36%	19.14%	25.28%
	Avg. Expression Level (Ly6G, Ly6C)	(1.31, 3.52)	(3.27, 2.47)	(1.47, 1.51)	(1.35, 3.51)	(3.31, 2.50)
	Avg. Standard Deviation	0.19	0.14	0.25	0.14	0.09
	Number of MC	2	2	6	3	3
IPTM	Fraction in Myeloid Cells	15.77%	17.24%	56.78%	14.38%	19.14%
	Avg. Expression Level (Ly6G, Ly6C)	(1.40, 3.53)	(3.43, 2.66)	(1.96, 1.65)	(1.41, 3.57)	(3.27, 2.56)
	Avg. Standard Deviation	0.24	0.15	0.18	0.16	0.09
	Number of MC	3	2	9	10	7

Table 1: Comparison of meta-clusters in three sets of templates shown in Figure 4. Only significant meta-clusters are included for myeloid cell templates, while for G- and M-MDSC templates all meta-clusters are included.

Classifier	All myeloid cells					M-MDSCs					G-MDSCs				
	A	P	R	S	F	A	P	R	S	F	A	P	R	S	F
KNN	0.721	0.787	0.800	0.870	0.728	0.744	0.751	0.782	0.874	0.752	0.930	0.916	0.908	0.952	0.909
SVM-Radial	0.930	0.935	0.924	0.961	0.929	0.837	0.831	0.831	0.914	0.828	0.922	0.970	0.941	0.971	0.953
SVM-Linear	0.837	0.840	0.856	0.913	0.847	0.488	0.473	0.478	0.737	0.472	0.837	0.849	0.874	0.920	0.851
Naive Bayes	0.977	0.861	0.831	0.986	0.978	0.791	0.809	0.764	0.882	0.779	0.930	0.957	0.923	0.957	0.933
Neural Network	0.791	0.814	0.765	0.881	0.781	0.442	0.394	0.418	0.716	0.397	0.814	0.841	0.814	0.894	0.826
FlowMatch	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2: Five Statistical measures (Accuracy, Precision, Recall, Specificity, and F-score) for the performance of classifiers using all myeloid cells, only M-MDSCs, or only G-MDSCs. Each classifier was trained using LOOCV and the final models were selected based on the highest accuracy. All statistical measures except Accuracy are averages due to multi-class classification.

3.3 Other classification approaches

We compare the template-based classification method to five other widely used non-template-based classification approaches, listed in Table 2; these methods are available in the R package `caret`.

(a) Features construction. We cluster each sample by k-means and construct features from the clusters. For each cluster, we select the size (cell count), the mean vector, and the covariance matrix as features, yielding a total of $1 + m + m^2$ features for one cluster, where m is the number of markers. With k clusters, an FC sample is finally represented by a feature vector of length $k * (1 + m + m^2)$. To ensure that the feature vectors have consistent meanings across samples, we label clusters in the ascending order of their mean vector. Additionally, we fix the the number of clusters for this comparison so that each sample has equal number of features without missing values. Using a majority voting rule, we set $k = 10$ in myeloid cells and $k = 2$ in M-MDSCs or G-MDSCs. With all samples having three markers, this yields 130 features for a myeloid cell sample and 26 features for an M-MDSC or G-MDSC sample.

(b) Classification results. The training function implemented in `Caret` automatically selects the best parameter set yielding the highest accuracy. For a fair comparison with flowMatch, we again use LOOCV to evaluate each classifier, and classify samples using myeloid cells, G-MDSCs only, and M-MDSCs only. For each classifier and each set of cells, we report prediction accuracy, precision, recall, specificity and the F-score in Table 2. SVM-Radial and Naive Bayes generally have the highest accuracies and F-scores among these, but at least one sample is misclassified by these methods. The neural network performs the worst in all cases. When classifying samples by non-template-based methods, G-MDSCs appear to be more distinguishing than all myeloid cells and M-MDSCs, while the highest classification accuracy comes from Naive Bayes classifier using all myeloid cells, where only one sample is misclassified.

In conclusion, non-template-based classifiers perform well in classifying samples using G-MDSCs and all myeloid cells, but none of them achieves the 100% prediction accuracy attained by the template-based method. Additionally, our template-based classifier clearly shows the characteristic cell populations and their distributions, thereby uncovering the underlying biological differences that are useful to clinicians.

4 CONCLUSIONS

This paper presents a new approach to classifying and phenotyping MDSCs by using templates built from FC data of NVBM, TBSP and IPTM. These templates provide a simple, yet effective, classification scheme, as well as a cleaner way to track prototypical changes in terms of their meta-clusters. For example, we can match meta-clusters between the NVBM and IPTM templates and precisely identify changes in different cell types. The ability to identify biologically interpretable features behind a classification decision is the novel and compelling aspect of our method. Methods to identify specific, disease-associated cell phenotypes is important for cancer research and drug discovery. Our results show that a phenotypic difference can be seen between MDSC found at the tumor and peripheral sites. This finding is consistent with unpublished trasncriptomic and functional data from our lab. As such, the template-based classification may aid in the development of drugs designed to target MDSC at a specific tissue or development state.

REFERENCES

[1] N. Aghaeepour, R. Nikolic, H.H. Hoos, and R.R. Brinkman. 2011. Rapid cell population identification in flow cytometry data. *Cytometry A* 79, 1 (2011), 6–13.

[2] A. Azad, J. Langguth, Y. Fang, A. Qi, and A. Pothen. 2010. Identifying rare cell populations in comparative flow cytometry. *Lect. Notes Bioinf. (Proc. WABI)* 6293 (2010), 162–175.

[3] A. Azad, B. Rajwa, and A. Pothen. 2016. Immunophenotype discovery, hierarchical organization, and template-based classification of flow cytometry samples. *Front. Onc.* 6 (2016), 188.

[4] C. C. Bozkus, B. D. Elzey, S. A. Crist, L. G. Ellies, and T. L. Ratliff. 2015. Expression of cationic amino acid transporter 2 is required for myeloid-derived suppressor cell-mediated control of T-cell immunity. *J. of Immunol.* 195 (2015), 5237–5250.

[5] V. Bronte, S. Brandau, S-H. Chen, et al. 2015. Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nature Communications* (2015), 10.

[6] M. Dundar, F. Akova, H. Z. Yerebakan, and B. Rajwa. 2014. A non-parametric Bayesian model for joint cell clustering and cluster matching: Identification of anomalous sample phenotypes with random effects. *BMC Bioinfo.* 15, 1 (Sep 2014), 314.

[7] M. Halkidi and M. Vazirgiannis. [n. d.]. Clustering validity assessment: Finding the optimal partitioning of a data set. In *Proc. IEEE Internat. Conf. Data Mining (ICDM).* 187–194.

[8] J. M. Haverkamp, S. A. Crist, B. D. Elzey, C. Cimen, and T. L Ratliff. 2011. In vivo suppressive function of myeloid-derived suppressor cells is limited to the inflammatory site. *Eur. J. Immunol.* 41, 3 (2011), 749–759.

[9] M. Malek, M. J. Taghiyar, L. Chong, G. Finak, R. Gottardo, and R. R. Brinkman. 2015. flowDensity: reproducing manual gating of flow cytometry data by automated density-based cell population identification. *Bioinf.* 31, 4 (Feb 2015), 606–607.

[10] I. Naim, S. Datta, J. Rebhahn, J. S. Cavenaugh, T. R. Mosmann, and G. Sharma. 2014. SWIFT: Scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, Part 1: Algorithm design. *Cytometry A* 85, 5 (2014), 408–421.

[11] D.R. Parks, M. Roederer, and W.A. Moore. 2006. A new Logicle display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. *Cytometry Part A* 69, 6 (2006), 541–551.

[12] J-I. Youn, S. Nagaraj, M. Collazo, and D. I. Gabrilovich. 2008. Subsets of myeloid-derived suppressor cells in tumor-bearing mice. *J. Immunol.* 181, 8 (2008), 5791–5802.