Published in final edited form as:

Integr Biol (Camb). 2017 April 18; 9(4): 328–338. doi:10.1039/c7ib00018a.

Deconvolution of heterogeneous wound tissue samples into relative macrophage phenotype composition *via* models based on gene expression[†]

Nicole M. Ferraro^a, Will Dampier^{a,b,c}, Michael S. Weingarten^d, and Kara L. Spiller^a

^aSchool of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA

^bDepartment of Microbiology and Immunology, Drexel University College of Medicine, Philadelphia, PA, USA

^cCenter for Molecular Virology and Translational Neuroscience, Institute for Molecular Medicine and Infectious Disease, Drexel University College of Medicine, Philadelphia, PA, USA

^dDepartment of Surgery, Drexel University College of Medicine, Philadelphia, PA, USA

Abstract

Macrophages, the primary cell of the innate immune system, act on a spectrum of phenotypes that correspond to diverse functions. Dysregulation of macrophage phenotype is associated with many diseases. In particular, defective transition from pro-inflammatory (M1) to anti-inflammatory (M2) behavior has been implicated as a potential source of sustained inflammation that prevents healing of chronic wounds such as diabetic ulcers. In order to design effective treatments, an understanding of the relative presence of macrophage phenotypes during tissue repair is necessary. Inferring the relative phenotype composition is currently challenging due to the heterogeneous nature of the macrophages themselves and also of tissue samples. We propose here a method to deconvolute gene expression from heterogeneous tissue samples into the composition of two primary macrophage phenotypes (M1 and M2). Our final method uses gene expression signatures for each phenotype cultivated *in vitro* as input to a predictive model that infers sample composition with an average error of 0.16, and whose predictions fit known compositions prepared *in vitro* with an R^2 value of 0.90. Finally, we apply this model to describe macrophage behavior in human diabetic ulcer healing using clinically isolated ulcer tissue samples. The model predicted that non-healing diabetic ulcers contained higher proportions of M1 macrophages compared to healing diabetic ulcers, in agreement with numerous studies that have implicated a dysfunctional M1-to-M2 transition in the impaired healing of diabetic ulcers. These results show proof of concept that the model holds utility in making predictions regarding macrophage behavior in heterogeneous samples, with potential application as a wound healing diagnostic.

[†]Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ib00018a

Correspondence to: Kara L. Spiller.

1. Introduction

Chronic wounds represent a severe clinical problem in the United States. Substantial progress has been made to develop new treatment strategies to promote healing of these wounds, but they still remain a critical issue. In order to understand how to induce chronic wound healing, it is necessary to understand the underlying pathology that prevents healing. A considerable body of work has shown that impaired wound healing is related to a defective inflammatory response.¹ In particular, macrophages, as the primary cell of the innate immune system, transition from primarily pro-inflammatory (M1) to primarily anti-inflammatory (M2) during the course of normal wound healing.²⁻⁴ Dysregulation in this sequence of macrophage activation has been implicated in a variety of pathological repair processes, including chronic diabetic ulcers,⁴⁻⁶ venous ulcers,⁷ myocardial infarction,⁸ and spinal cord injury.⁹ Thus, a method for characterizing macrophage behavior in human tissue is necessary in order to understand their contribution to disease and repair processes, and to develop targeted treatments.

Previously, in a clinical study focusing on chronic diabetic ulcers, we showed that the relative expression of a panel of genes associated with the M1 macrophage phenotype compared to genes associated with the M2 phenotype could be used to predict wound healing, in that the M1/M2 ratio increased for all non-healing wounds but decreased for all healing wounds, which holds clinical value as a diagnostic.⁵ However, the M1/M2 ratio was calculated from gene expression from whole wound tissue because isolation of macrophages is not feasible from the small tissue samples obtained in the clinic. Because the chosen gene markers were not specific to macrophages, the source of the dysregulation could not be identified for use in potential future designs of new treatments and diagnostics. Therefore, the goal of this study was to develop a method to characterize macrophage behavior in heterogeneous wound tissue samples.

The most common methods of characterizing macrophage behavior are based on immunohistochemical analysis (immunostaining and flow cytometry). Immunostaining methods are time-consuming and only semi-quantitative, and the utility of flow cytometry is impeded by the requirement of immediately processing clinical samples that are often too small for cell isolation. Techniques based on gene expression are more promising because small tissue samples are adequate to obtain sufficient quantities of RNA, and computational deconvolution can be used to provide information about the cell source of expression without losing the potential for system-level information. Moreover, gene expression methods are less restricted in the number of markers used in the analysis. However, most deconvolution techniques rely on whole transcriptome sequencing,^{10,11} limiting their widespread application. Thus, the goal of this study was to develop a method of deconvolution to predict the relative composition of macrophage phenotype in heterogeneous tissue samples using a small number of genes as input.

In this study, we began by identifying markers of the M1 and M2 phenotypes of macrophages by comparing lists generated by multiple laboratories using *in vitro* culture of primary human macrophages. Then, we eliminated genes that are highly expressed by other cell types present in the wound environment. We then generated training data for a predictive

model by assessing the expression of those markers in mixed samples of M1 and M2 macrophages of known compositions that were generated *in vitro*. The accuracy of the resulting models was tested on an additional dataset consisting of gene expression data from samples that included M1 and M2 macrophages as well as fibroblasts and endothelial cells to represent the noise in heterogeneous wound tissue samples. Finally, the models were applied to predict macrophage phenotype composition *in vivo* using heterogeneous tissue samples isolated from the wounds of patients with diabetic foot ulcers.

2. Methods

2.1. Experimental design

We performed several steps to develop predictive models based on a small number of genes as input (Fig. 1). We first sought to determine a list of genes that are differentially expressed (DE) between human M1 and M2 macrophages (prepared *in vitro*) and also expressed at low levels by other cell types present in wound tissue. The gene list was further refined through correlation analyses with mixed samples containing known compositions of M1 and M2 macrophages. The expression of the selected genes in mixed samples was then used as training data for predicting sample composition in mixed samples of M1 and M2 macrophages as well as fibroblasts and endothelial cells. Several predictive models were implemented, including linear regression, logistic regression, support vector regression, and neural net models (using the glmnet, nnet, and e1071 packages in R). The accuracy of the resulting models was improved using a feature selection process based on each feature's contribution to the overall variance between samples. The best performing model was applied to gene expression data generated from human wound tissue samples isolated from patients with diabetic foot ulcers to predict the relative proportion of M1 and M2 macrophages. Finally, it is important to note that while we refer to M1 and M2 macrophages as discrete cell populations for the sake of simplicity, they are more likely to exist on a spectrum *in vivo*, including hybrid M1–M2 phenotypes as well as phenotypes that are entirely distinct from the typical M1 and M2 phenotypes.¹² Thus, when applied to heterogeneous samples obtained from the *in vivo* environment, the models developed here are more likely to predict the relative M1–M2 character of the sample in terms of their relationship to *in vitro*-polarized macrophages, as opposed to the actual number of macrophages of a particular phenotype.

2.2. Gene list determination

In order to compile genes that are DE between M1 and M2 macrophages, several independently generated datasets of *in vitro*-polarized human macrophages were used. The datasets included were RNA-sequencing (RNA-seq) results generated by our laboratory (temporary ID: PRJNA339309),¹³ the dataset publicly available as GSE57614,¹⁴ and the dataset publicly available as GSE36537.¹⁵ Brief experimental details for each of the datasets are listed in Table 1.

For the RNASeq dataset, differential expression was determined using the voom-limma package in R.¹⁶ Thresholds for false discovery rate adjusted *p*-value and log base 2 fold-change (log FC) were established before the analysis as 0.05 and 2, respectively. The *p*-value

threshold was chosen based on standard measures of statistical significance and was corrected for multiple testing using the Benjamini–Hochberg procedure.¹⁷ The threshold for $\log FC > 2$ was chosen to select for a larger difference in expression between phenotypes. For the microarray datasets, the limma R package¹⁸ was used to assess DE genes using the same thresholds, independently of the published results associated with each dataset. This selection process determined which genes had the potential to function as classifiers between macrophage types.

Then, because no studies have yet isolated and sequenced macrophages from human wounds, a negative selection process was used to remove those genes that were expressed at high levels in other wound cell types, including fibroblasts, keratinocytes, melanocytes, basal epidermal cells, T cells and endothelial cells from skin tissue samples. This process was conducted manually *via* searching available data from the Human Protein Atlas, the GeneCard resource, which compiles information from NCBI, HGNC, ENSEMBL, and UniProtKB,¹⁹ as well as primary literature. Genes were removed if the Human Protein Atlas coded medium-high expression of that gene in wound cell types. If that was not the case, GeneCards and associated resources were consulted to identify any other wound cell types as reported expressers of that gene. If neither method eliminated the gene, a literature search was conducted, and if any previous work reported that the gene was a marker for the cell types likely to be found in the wound environment, it was removed.

Then, in order to assess whether the remaining markers were upregulated in wound cell types responding to an inflammatory stimulus, publicly available datasets were used to look for DE genes between baseline and cell types upon inflammatory activation, focusing on keratinocytes isolated from a model of epidermal injury (GSE30355),²⁰ dermal fibroblasts activated *via* serum (GSE21648),²¹ neutrophils isolated from sepsis patients (GSE3037),²² endothelial cells isolated from dermis of type 2 diabetes patients *versus* normal (GSE38396),²³ naïve blood T-cells stimulated with inflammatory cytokines (GSE39596),²⁴ and CD-40 activated B-cells (GSE54017).²⁵ If the fold change between the activated and baseline cells was significantly greater than 1.5 ($p < 0.05$), that gene was eliminated. This lower $\log FC$ threshold of 1.5 was chosen in this analysis compared to the previous step ($\log FC > 2$) in order to increase the specificity of the selected genes for macrophages. The differential expression of remaining genes was also evaluated between activated macrophages and dendritic cells differentiated from monocytes (GSE33273).²⁶

Finally, the final list of selected genes was evaluated for expression levels in publicly available data sets of human macrophages isolated *ex vivo* from various tissues. To our knowledge no study has yet isolated and sequenced human macrophages from wound tissue. Therefore, we analyzed a range of data sets describing macrophages isolated from adipose, lung, and tumor tissues (GSE4350,²⁷ GSE40885,²⁸ GSE8823,²⁹ GSE42495,³⁰ GSE13896³¹). Specific sample collection and data processing methods for each dataset can be found in their individual publications. The first quartile (25%) of the expression values was used to determine if the level of expression of each gene constituted significant expression by the macrophages. Any gene expressed below this level was considered not expressed.

2.3. Generation of mixed macrophage samples for training data

Phenotypically heterogeneous macrophage samples of known compositions were generated *in vitro* to serve as training data for the predictive models. Primary human monocytes were purchased from the University of Pennsylvania Human Immunology Core following their isolation *via* negative selection from peripheral blood mononuclear cells. Monocytes were cultured in 20 ng ml⁻¹ MCSF for 5 days, with a media change at day 3, prior to the addition of polarizing cytokines from days 5 to 7 (100 ng ml⁻¹ IFN γ and 100 ng ml⁻¹ LPS for M1; 40 ng ml⁻¹ IL4 and 20 ng ml⁻¹ IL13 for M2), as previously described.^{32,33} The macrophages were cultured at 10⁶ cells per ml and 10⁶ cell per well in multiple wells of a 24-well ultra low attachment plate and then combined to minimize experimental variability. M1 or M2 macrophages were lysed and combined in differing volumetric combinations ranging from 0 to 100% M1 cells because we have previously shown that the cell concentration for each phenotype is not different after just two days of culture.³⁴ The training data contained, in terms of M1 sample composition, five 100%, four 80%, five 75%, four 70%, two 60%, ten 50%, four 40%, four 30%, six 25%, four 20%, four 10%, and eight 0% points, for a total of 60 samples in the training data set, with three donors serving as biological replicates and two technical replicates per donor for each sample.

RNA extraction was conducted using an RNAqueous[®]-Micro kit (Thermo Fisher) according to the manufacturer's instructions. The RNA quantity and quality were measured on a Nanodrop ND1000. The 260/280 absorbance ratio was ~2 for all samples. The samples were stored at -80 °C until cDNA synthesis. The DNase I amplification grade kit (Thermo Fisher) was applied to the samples, according to the manufacturer's instructions. Reverse transcription was then performed using the High Capacity cDNA synthesis kit (Thermo Fisher), according to the manufacturer's instructions. Quantitative RT-PCR analysis was then performed with 20 ng cDNA per each reaction and the Fast SYBR[®] Green PCR Master Mix (Thermo Fisher). All primers were synthesized by Thermo Fisher (Table S1, ESI[†]). A melt curve analysis was used to assess the stability of the primers, and any dimerized primers were removed from analysis, which included one gene, CTNNAL1. Any samples with an expression value for GAPDH that differed from the mean of that set by more than 2 cycles were eliminated. The expression of each gene was normalized to the expression of GAPDH (2^{-Ct}) and log-normalized prior to analysis.

This process was repeated in two separate batches with two different donors. A third dataset was created from samples containing not only M1 and M2 macrophages, but also human umbilical vein endothelial cells (Lonza) and human dermal fibroblasts (Lonza). These cells were cultured in their respective growth media using the EGM-2 BulletKit with growth factors (Lonza) for endothelial cells and the FGM-2 BulletKit with growth factors (Lonza) for fibroblasts, according to the manufacturer's instructions, for two days, before being detached *via* Trypsin/EDTA (Lonza), according to the manufacturer's instructions. The cells in each flask were lysed, and then added to the samples that varied in M1/M2 composition. Macrophages have been reported as a primary component of a wound environment,^{35,36} further detailed in review,³⁷ and so 75% of the sample contained varying proportions of M1

[†]Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ib00018a

and M2 macrophages, as indicated, while the remaining 25% of the total sample volume consisted of non-macrophage cells (12.5% each of fibroblasts and endothelial cells). Our goal in these experiments was not to recapitulate the wound environment, but rather to assess the robustness of the models when adding noise from expression originating in other relevant cell types.

Gene expression values in each dataset were normalized to the dataset's mean using the z-score. The expression of each gene in datasets containing only M1 and M2 cells in varying compositions was correlated with known sample composition (represented as the fraction of M1 cells in the sample) using the Pearson product-moment correlation coefficient to further refine the list of markers.

2.4. Model development

A variety of models were developed to determine an optimal method for predicting sample composition. Linear and logistic regression models were used as simple methods for interpolation. Two machine learning models were implemented to improve on predictive ability: neural networks (NN)³⁸ and support vector machines (SVM).³⁹ All models were built using data from one dataset representing one donor and tested on data from additional donors. Then, the models were trained using data from both donors and testing on a third dataset containing the mixture of macrophages, fibroblasts, endothelial cells to simulate heterogeneous tissue.

Linear and logistic regression models were implemented using the glm package in R, and the NN and SVM methods were implemented using the e1071 package. The NN was constructed with varying sizes from 1 to 9 to determine the best parameters, to result in a network with 15 weights, 5 input nodes (gene expression values), 2 hidden layers, and one output node. A tuning process was incorporated to optimize performance. The tuning processed varied the hidden layer size between 1 and 9, based on the number of input features. The weight decay value for the NN was 0.1, which allowed variance of the weight values to optimize their performance, with a maximum iteration value of 1000.

The resulting SVM used a radial kernel transformation with an error threshold of 0.1 and included 74 support vectors out of 87 training points. Values of gamma, the influence of a single variable on performance, and cost, influence of errors, for the support vector machine were 0.2 and 1, respectively. Tuning was used to optimize these parameters, and they were determined to ensure all variables began with equal influence, and that errors did not have a large influence on performance.

In order to assess the accuracy of each model, that model's predictions were correlated with known sample composition to produce a correlation coefficient measure. Leave-one-out-cross-validation (LOOCV) was used in which the model was trained using all but one sample and tested on that remaining sample when assessing the goodness of fit for the training data. The average error for the models' predictions of test data was calculated by summing the difference between each prediction and known value and dividing by the number of test samples.

To further improve accuracy, a forward feature selection process was applied to the best performing models, with tenfold validation at each step. The sample set was split into 10 groups and the models were trained on nine groups and tested on the tenth for all combinations to determine the overall best combination of features to include. The genes were added in order of decreasing contribution to the variance between samples from two donors. In other words, the models started with only using expression data from CCL26, and then were added in the following order: ALOX15, PLCB1, XAF1, ASPHD2, RAMP1, WNT5B, CLEC4G, and CCL19. This order was chosen because the genes that contributed most to the variance between samples, determined by principle component analysis, should be most effective at predicting the relative composition of those samples.⁴⁰ Because NN results can depend on the initial random weights assigned to each node, these changes were minimized both through tuning and in creating ten networks for each model and averaging the error at each step. Standard deviation was assessed to determine if there was large variation in the networks based on initial values. The best model was then selected based on the tuning parameters that led to minimal differences in predictions between trainings.

2.5. Application to clinical samples of human diabetic ulcers

Subjects with chronic diabetic foot ulcers were recruited from the Drexel University Wound Healing Center and provided written consent in accordance with the study protocol approved by the Drexel University Institutional Review Board. Exclusion criteria were evidence of peripheral vascular disease (ankle brachial index <0.75) and those who presented with signs and symptoms of osteomyelitis or systemic infection, as in our previous study.⁴¹ Participants underwent standard treatment, including weekly or biweekly debridement with a sharp scalpel, offloading, and application of moist wound dressings, as determined by the physician. Tissue samples were collected during debridement at each visit until complete wound healing, amputation, or the end of the study (up to 38 weeks from the collection of the first sample). At the completion of the study, subjects' ulcers were classified as "healing" if they closed within 12 weeks of sample collection or "non-healing" if they failed to close within 12 weeks or if the affected limbs were amputated, which is an established, objective method of ulcer classification.⁴² In this study, all healing subjects achieved ulcer healing within 3 weeks of sample collection, and none of the non-healing subjects achieved healing for the duration of the study (18–38 weeks from collection of the first sample). Clinical information for the subjects included in the analysis is shown in Table 2.

Ulcer tissue samples were stored using RNeasy lysis buffer at -80°C until RNA extraction. RNA was extracted by homogenization in Trizol followed by purification using the RNeasy Mini kit (Qiagen) as in our previous study.⁴¹ Gene expression analysis for the genes of interest was conducted using a custom NanoString code set (NanoString Technologies, Seattle, WA) according to the manufacturer's instructions. Samples were removed from downstream analysis if gene counts of any of the 7 included housekeeping genes (EIF2B2, GAPDH, PPIA, RPL37A, TFRC, UBE2D2, VCP, VSP29) were lower than background levels. The 20 remaining samples included in the analysis were from 8 subjects (Table 2). Gene expression values were normalized to GAPDH expression and also normalized to the mean using the z-score because these transformations were performed on the RT-PCR data for training of the models. These values served as input for the predictions of sample composition in the final

model. An unpaired *t*-test was used to compare model predictions between healing and non-healing ulcers at a level of significance of $p < 0.05$.

3. Results

3.1. Gene list determination

Publicly available data from three experiments involving *in vitro* polarized macrophages were used to establish specific markers for each phenotype. 92 M1 genes and 26 M2 genes overlapped between the three datasets (Fig. 2a and Table S2, ESI[†]). Of these, 8 M1 genes and 8 M2 genes were found to be primarily expressed by macrophages, *i.e.* not highly expressed by other wound tissue cell types (Fig. 2b and Table S3, ESI[†]). Of these, 6 M1 genes and 7 M2 genes were found to be not upregulated by the other wound cell types upon their activation with inflammatory stimulus with $\log_2 \text{FC} > 1.5$ between activated and baseline cell types (Table S4, ESI[†]), but one gene had to be removed because of a bad RTPCR primer (CNNTAL1), resulting in a final list of 12 genes (Table 3). Only 2 of the final 12 genes (PLCB1 and WNT5B) were expressed at similar levels by macrophages and dendritic cells (Table S5, ESI[†]). These genes were not removed from the analysis because removal would have resulted in 6 M1 genes and only 4 M2 genes. Finally, we confirmed that the selected genes are expressed by human macrophages isolated *ex vivo* from human tissues using publicly available data sets (Table S6, ESI[†]).

We next analyzed whether or not the selected genes correlated with increasing proportions of M1/M2 macrophages in mixed samples (Fig. 3). The genes were clustered based on their similar expression trends across samples, and the samples were clustered based on their expression across genes, represented on the horizontal and vertical axes, respectively. As expected, the M1 genes clustered together, as did the M2 genes. Samples with 100% M1 cells clustered furthest from samples with 100% M2 cells, with mixed samples clustering between these extremes.

The correlation coefficient for the expression of each gene with sample composition was calculated using mixed macrophage samples from two donors (Table 4). An example of the correlation analysis is shown in ESI[†], Fig. S1. Most significant genes (with the exception of CLEC4G for samples from donor 2) followed the expected trends. Any genes that failed to correlate significantly with sample composition ($p > 0.05$) when both donors were combined (60 samples total) were excluded from further analysis, and these were APOL1, CLEC4D, and NOD2.

3.2. Model development

Data from each donor were used to generate four potential models to predict sample composition of a heterogeneous population of macrophages from the same donor. The models were evaluated using average predictive error, calculated *via* leave-one-out-cross-validation (LOOCV), and goodness of fit (R^2). NN and SVM models performed well when trained and tested on the same donor, with average error around 0.01–0.03 and R^2 around 0.99 (Table 5). However, when training the models on data from one donor and testing on the other, the accuracy and goodness of fit decreased. Not surprisingly, the results were

similar when the models were tested on simulated heterogeneous samples from a third donor that also contained fibroblasts and endothelial cells in that the accuracy decreased from the LOOCV results, even when the models were built on data from the combination of two donors (Table 6).

In order to enhance the predictive ability of the models by focusing on genes that are most important in determining sample composition, feature selection was applied to the three best-performing models in terms of average predictive error: linear regression, NN, and SVM (Fig. 4). Genes were added as features in the order of their contribution to the variance of the training samples from principle component analysis, and that order was ALOX15, CCL26, PLCB1, XAF1, ASPHD2, CCL19, WNT5B, RAMP1, and CLEC4G. The linear regression and NN models achieved maximum accuracy when using data from seven genes: XAF1, ASPHD2, and CCL19 (M1 genes), and ALOX15, CCL26, PLCB1, WNT5B (M2 genes), while the SVM model was most accurate using data from four genes: XAF1 (M1 gene) and ALOX15, CCL26, PLCB1 (M2 genes).

Following feature selection, both the NN and linear regression models achieved similar performance, with average errors around 0.16 when models were trained using seven genes. The predictions of both models trained using data from two donors *vs.* known sample compositions of the heterogeneous dataset (Fig. 5) showed high correlation between the predictions and known values, with linear goodness of fit values around 0.9. The SVM model had slightly reduced performance, with an average error of 0.181 and an R^2 of 0.859. Because of the higher R^2 values and the versatility of the NN model in allowing extension to other applications, the NN model was used in further analysis.

3.3. Application to clinical samples of diabetic ulcers

The final model was applied to gene expression data obtained from heterogeneous tissue samples collected from 8 patients with diabetic ulcers. The model predicted higher proportions of M1 macrophages in non-healing ulcers compared to healing ulcers (Fig. 6).

4. Discussion

Wound healing is a complex process involving a plethora of cell types whose behavior can affect the wound's ability to heal. In order to design effective treatments in impaired wound healing situations such as chronic wounds, an understanding of which cells are dysregulated at various time points is essential. Because macrophages are a primary component of the inflammatory response, we focused on isolating their behavior within the heterogeneous environment of a wound. The final model developed in this study predicted that non-healing diabetic ulcers exhibited higher proportions of M1 macrophages relative to M2 macrophages compared to healing ulcers in a preliminary study of 8 subjects. Numerous studies in animal models and humans have employed traditional immunohistochemical techniques to show that diabetic ulcers contain higher levels of M1 macrophages compared to healthy wounds, and that impaired healing is associated with a failed transition of macrophages from M1 to M2.^{43–45} Together, these results suggest that a deconvolution model based on a small number of genes may hold value in determining the relative levels of M1 and M2 macrophages in a human diabetic ulcer. With further development, this model may be

clinically useful in a diagnostic to determine if a patient's ulcer is on a healing or non-healing trajectory, which is currently difficult in clinical practice.⁴⁶

Many studies regarding wound healing involve whole tissue samples.^{47,48} While a general understanding of the genes present in these tissues can provide important insight into the physiology of the wound, targeted treatment strategies will require an understanding of which cells to target. For example, ongoing work in our laboratory is focused on developing biomaterials that could release cytokines to influence macrophage behavior.⁴⁹ It has been shown previously that chronic wounds undergo sustained inflammation,⁵⁰ and that M1 macrophages are a potential source for this inflammation.^{47,51} Our model provides a potential method for understanding the macrophage-specific behaviors in the context of wound tissue, which can not only inform the targets for these therapies, but also demonstrate whether the cells are responding to the proposed treatments. Previous work has described the use of expression of a refined set of genes to make predictions regarding cell behavior; for example, the ability to classify types of tumor cells based on gene expression has been used to predict an optimal treatment strategy for individual cancers.^{52–54} Similarly, understanding macrophage behavior in wounds could lead to a diagnostic that predicts whether that wound is responding to treatment or whether a change in treatment strategy should be considered.⁵ It has been previously estimated that an accurate diagnostic to predict wound healing trajectory would save approximately \$12 000 per ulcer per patient.⁵⁵ The work presented here demonstrates the potential feasibility of this approach in wound healing, but it could be applicable to other pathological scenarios, such as tumor analysis or other types of complex injuries.

In this study, we identified markers of two main macrophage phenotypes, M1 and M2. The genes that were identified, ASPHD2, CCL19, and XAF1 for M1, and CCL26, ALOX15, PLCB1, and WNT5B for M2, could potentially provide insight on differential roles between the phenotypes in regulation of tissue growth. For example, XAF1 regulates apoptosis and has been implicated in anti-tumor behavior,⁵⁶ while WNT5B regulates tissue development and may promote tumor growth.⁵⁷ In addition, both CCL19 and CCL26 are involved in T cell trafficking. While we did not explore the functional role of these genes in the context of wound healing, their identification as differentiators between the macrophage phenotypes suggests that dysregulation of their expression could impact macrophage behavior within the wound environment.

These markers were used in an NN model to predict the relative macrophage composition in whole tissue without the need to first isolate these cells. The utility of these genes in predicting sample composition of a mixed cell population was optimal when both training and test data originated from the same donor, which would be useful for *in vitro* experiments in which reference samples could be generated to aid in characterization of an experimental population. Even when testing on other donors, the model's performance was comparable in terms of goodness of fit and average predictive error to flow cytometry, the gold standard to characterize heterogeneous populations, which has been shown to have predictive error around 0.15 and R^2 values around 0.9.^{58,59} Overall, the model can be applied in both *in vivo* and *in vitro* situations to improve understanding of the macrophage behavior in tissue

samples. With further optimization, this model may allow for both high-throughput and cost effective analysis of heterogeneous samples for use in research and clinical applications.

The strategy to select genes for use in our models revealed several interesting and surprising findings. First, we found an overall lack of consistency in gene markers of macrophage cell types reported with the same label (M1 vs. M2) in three different datasets, with only 92 M1 genes and 26 M2 genes consistently identifying the phenotypes. The inconsistencies may derive from differences in polarization procedures, which differed slightly in dose and timing across datasets (Table 1). These findings demonstrate the importance of including multiple datasets when identifying cell type markers.

Secondly, that the model's accuracy improved following feature selection was also surprising, given the relatively strict filtering process applied to the original selection of genes. A model to make predictions using a small set of genes regarding immune cell behavior has previously been implemented to classify dendritic cells as steady-state or inflammatory.⁶⁰ This model also produced increased accuracy with a smaller subset of the genes than originally identified, and the authors proposed that this smaller set represented the genes most involved in consistently regulating activation of dendritic cells. In our study the improved accuracy with fewer genes may also indicate genes that are more consistently involved in macrophage polarization. The three genes that decreased accuracy were the genes that least strongly correlated with sample composition and thus contributed less to the variance between samples, and could indicate that while these genes are differentially expressed between the macrophage cell types, they are not equally expressed across the cells classified within one cell type. In addition, we expand on these previous findings by incorporating the ability to characterize cells in a whole-tissue context, as well as the ability to obtain continuous predictions of composition.

This study did have several limitations. First, because no studies have isolated and sequenced human macrophages from the wound environment, and because there is no established consensus as to what constitutes an M1 and M2 macrophage *in vivo*, we are unable to definitively determine if the selected genes are good markers of M1 and M2 macrophages *in vivo*. This represents the major limitation of our study, and future work will focus on thoroughly characterizing macrophages isolated from the human wound environment. Secondly, only two main macrophage subtypes were included in the analysis. While the M1 and M2 phenotypes as defined in this study have been strongly implicated in wound healing,^{5,61} other phenotypes may also be important, including those stimulated by IL10,^{61,62} phagocytosis of apoptotic cells,⁶³ or crosstalk with mesenchymal stem cells,⁶⁴ among many others. Future studies will focus on incorporating additional macrophage phenotypes that are important in wound healing into the predictive model. In addition, it would be useful to optimize the model *in vitro* using additional relevant cell types, such as neutrophils, dendritic cells, T cells, fibroblasts, and keratinocytes, in more realistic proportions, because the heterogeneous samples that we used to assess the model, while providing insight on its robustness to noise, did not accurately simulate the complexity of a wound in terms of the multitude of cell types present. Another limitation of the strategy more generally is the use of *in vitro*-polarized macrophages to define macrophage populations *in vivo*. While this strategy has been used to show that macrophages from the

lungs of patients who are smokers were closely related to macrophages stimulated *in vitro* with glucocorticoids,⁶⁵ the differences between macrophages in the *in vitro* and *in vivo* environments are not well understood. Future studies should focus on clearly delineating these differences between the *in vitro* and *in vivo* environments, and also incorporate *ex vivo* samples from the wound environment. A technical limitation of our study was that the training data for the models were generated *via* RT-PCR while the clinical samples were analyzed using NanoString. This difference in platform could affect the predictive ability. These limitations notwithstanding, this study provides proof of concept on the use of targeted gene expression data to build models that predict changes in macrophage behavior in a heterogeneous environment.

5. Conclusions

This study showed that predictive models built from carefully selected genes and training data from mixed samples of M1 and M2 macrophages can be used to predict general trends in the relative phenotypic composition in heterogeneous wound samples. Future studies will focus on incorporation of a more diverse macrophage phenotype spectrum and comparison of the *in vitro* and *in vivo* microenvironments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NHLBI R01 HL130037 and the Drexel-Coulter Translational Research Partnership. The authors are grateful to Brandon Marcinkiewicz, Carly Deussenberg, Jane McDaniel, Lori Jenkins, and Lisa Flowers-Morrison for help with patient enrollment and sample collection.

References

1. Eming SA, et al. Inflammation in wound repair: molecular and cellular mechanisms. *J. Invest. Dermatol.* 2007; 127(3):514–525. [PubMed: 17299434]
2. Menzies F, et al. Sequential expression of macrophage anti-microbial/inflammatory and wound healing markers following innate, alternative and classical activation. *Clin. Exp. Immunol.* 2010; 160(3):369–379. [PubMed: 20059482]
3. Rodero MP, et al. Skin wound healing modulation by macrophages. *Int. J. Clin. Exp. Pathol.* 2010; 3(7):643. [PubMed: 20830235]
4. Mirza R, et al. Dysregulation of monocyte/macrophage phenotype in wounds of diabetic mice. *Cytokine.* 2011; 56(2):256–264. [PubMed: 21803601]
5. Nassiri S, et al. Relative Expression of Proinflammatory and Antiinflammatory Genes Reveals Differences between Healing and Nonhealing Human Chronic Diabetic Foot Ulcers. *J. Invest. Dermatol.* 2015; 135(6):1700–1703. [PubMed: 25647438]
6. Mirza RE, et al. Blocking interleukin-1 β induces a healing-associated wound macrophage phenotype and improves healing in type 2 diabetes. *Diabetes.* 2013; 62(7):2579–2587. [PubMed: 23493576]
7. Novak ML, et al. Phenotypic transitions of macrophages orchestrate tissue repair. *Am. J. Pathol.* 2013; 183(5):1352–1363. [PubMed: 24091222]
8. Nahrendorf M, et al. The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. *J. Exp. Med.* 2007; 204(12):3037–3047. [PubMed: 18025128]

9. Kigerl KA, et al. Identification of two distinct macrophage subsets with divergent effects causing either neurotoxicity or regeneration in the injured mouse spinal cord. *J. Neurosci.* 2009; 29(43): 13435–13444. [PubMed: 19864556]
10. Shen-Orr SS, et al. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr. Opin. Immunol.* 2013; 25(5):571–578. [PubMed: 24148234]
11. Yadav VK, De S. An assessment of computational methods for estimating purity and clonality using genomic data derived from heterogeneous tumor tissue samples. *Briefings Bioinf.* 2014; 16(2):232–241.
12. Xue J, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity.* 2014; 40(2):274–288. [PubMed: 24530056]
13. Lurier EB, et al. Transcriptome analysis of IL-10-stimulated (M2c) macrophages by next-generation sequencing. *Immuno-biology.* 2017; doi: 10.1016/j.imbio.2017.02.006
14. Derlindati E, et al. Transcriptomic analysis of human polarized macrophages: more than one role of alternative activation? *PLoS One.* 2015; 10(3):e0119751. [PubMed: 25799240]
15. Mehraj V, et al. Monocyte responses in the context of Q fever: from a static polarized model to a kinetic model of activation. *J. Infect. Dis.* 2013; 208(6):942–951. [PubMed: 23801603]
16. Law CW, et al. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014; 15(2):R29. [PubMed: 24485249]
17. Benjamini Y, et al. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B.* 1995:289–300.
18. Smyth, GK. Bioinformatics and computational biology solutions using R and Bioconductor. Springer; 2005. Limma: linear models for microarray data; p. 397-420.
19. Safran M, et al. GeneCards Version 3: the human gene integrator. *Database.* 2010; 2010:pbaq020.
20. Kennedy-Crispin M, et al. Human keratinocytes' response to injury upregulates CCL20 and other genes linking innate and adaptive immunity. *J. Invest. Dermatol.* 2012; 132(1):105–113. [PubMed: 21881590]
21. Enoch S, et al. 'Young' oral fibroblasts are geno/phenotypically distinct. *J. Dent. Res.* 2010; 89(12): 1407–1413. [PubMed: 21098492]
22. Silva E, et al. HMGB1 and LPS induce distinct patterns of gene expression and activation in neutrophils from patients with sepsis-induced acute lung injury. *Intensive Care Med.* 2007; 33(10): 1829–1839. [PubMed: 17581740]
23. Haemmerle M, et al. Enhanced lymph vessel density, remodelling and inflammation is reflected by gene expression signatures in dermal lymphatic endothelial cells in type 2 diabetes. *Diabetes.* 2013; DB_120844. doi: 10.2337/db12-0844
24. Martínez-Llordella M, et al. CD28-inducible transcription factor DEC1 is required for efficient autoreactive CD4⁺ T cell response. *J. Exp. Med.* 2013; 210(8):1603–1619. [PubMed: 23878307]
25. Shimabukuro-Vornhagen A, et al. Inhibition of protein geranylgeranylation specifically interferes with CD40-dependent B cell activation, resulting in a reduced capacity to induce T cell immunity. *J. Immunol.* 2014; 193(10):5294–5305. [PubMed: 25311809]
26. Maliga Z, et al. A genomic toolkit to investigate kinesin and myosin motor function in cells. *Nat. Cell Biol.* 2013; 15(3):325–334. [PubMed: 23417121]
27. Dalmas E, et al. T Cell-derived IL-22 amplifies IL-1 β -driven inflammation in human adipose tissue: relevance to obesity and type 2 diabetes. *Diabetes.* 2014; 63(6):1966–1977. [PubMed: 24520123]
28. Reynier F, et al. Gene expression profiles in alveolar macrophages induced by lipopolysaccharide in humans. *Mol. Med.* 2012; 18(9):1303. [PubMed: 22952057]
29. Kazeros A, et al. Overexpression of apoptotic cell removal receptor MERTK in alveolar macrophages of cigarette smokers. *Am. J. Respir. Cell Mol. Biol.* 2008; 39(6):747–757. [PubMed: 18587056]
30. Hartmann S, et al. Macrophages in T cell/histiocyte rich large B cell lymphoma strongly express metal-binding proteins and show a bi-activated phenotype. *Int. J. Cancer.* 2013; 133(11):2609–2618. [PubMed: 23686423]

31. Shaykhiev R, et al. Smoking-dependent reprogramming of alveolar macrophage polarization: implication for pathogenesis of chronic obstructive pulmonary disease. *J. Immunol.* 2009; 183(4): 2867–2883. [PubMed: 19635926]
32. Graney PL, et al. *In vitro* response of macrophages to ceramic scaffolds used for bone regeneration. *J. R. Soc., Interface.* 2016; 13(120):20160346. [PubMed: 27466438]
33. Witherel CE, et al. Response of human macrophages to wound matrices *in vitro*. *Wound Repair Regen.* 2016; 24(3):514–524. [PubMed: 26874797]
34. Spiller KL, et al. The role of macrophage phenotype in vascularization of tissue engineering scaffolds. *Biomaterials.* 2014; 35(15):4477–4488. [PubMed: 24589361]
35. Loots MA, et al. Differences in cellular infiltrate and extracellular matrix of chronic diabetic and venous ulcers *versus* acute wounds. *J. Invest. Dermatol.* 1998; 111(5):850–857. [PubMed: 9804349]
36. Leibovich SJ, et al. The role of the macrophage in wound repair. A study with hydrocortisone and antimacrophage serum. *Am. J. Pathol.* 1975; 78(1):71–100. [PubMed: 1109560]
37. Martin P, et al. Inflammatory cells during wound repair: the good, the bad and the ugly. *Trends Cell Biol.* 2005; 15(11):599–607. [PubMed: 16202600]
38. Specht DF. A general regression neural network. *IEEE Trans. Neural Netw.* 1991; 2(6):568–576. [PubMed: 18282872]
39. Gunn SR. Support vector machines for classification and regression. *ISIS technical report.* 1998:14.
40. Song, F., et al., editors. Feature selection using principal component analysis. *System Science, Engineering Design and Manufacturing Informatization (ICSEM), 2010 International Conference on; IEEE.* 2010.
41. Nassiri S, et al. Relative Expression of Proinflammatory and Antiinflammatory Genes Reveals Differences between Healing and Nonhealing Human Chronic Diabetic Foot Ulcers. *J. Invest. Dermatol.* 2015; 135(6):1700–1703. [PubMed: 25647438]
42. Sheehan P, et al. Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial. *Diabetes Care.* 2003; 26(6): 1879–1882. [PubMed: 12766127]
43. Mirza RE, et al. Sustained inflammasome activity in macrophages impairs wound healing in type 2 diabetic humans and mice. *Diabetes.* 2014; 63(3):1103–1114. [PubMed: 24194505]
44. Mirza R, et al. Dysregulation of monocyte/macrophage phenotype in wounds of diabetic mice. *Cytokine.* 2011; 56(2):256–264. [PubMed: 21803601]
45. Sindrilaru A, et al. An unrestrained proinflammatory M1 macrophage population induced by iron impairs wound healing in humans and mice. *J. Clin. Invest.* 2011; 121(3):985–997. [PubMed: 21317534]
46. Margolis DJ, et al. Diabetic neuropathic foot ulcers: predicting which ones will not heal. *Am. J. Med.* 2003; 115(8):627–631. [PubMed: 14656615]
47. Sindrilaru A, et al. An unrestrained proinflammatory M1 macrophage population induced by iron impairs wound healing in humans and mice. *J. Clin. Invest.* 2011; 121(3):985–997. [PubMed: 21317534]
48. Trengove NJ, et al. Mitogenic activity and cytokine levels in non-healing and healing chronic leg ulcers. *Wound Repair Regen.* 2000; 8(1):13–25. [PubMed: 10760211]
49. Spiller KL, et al. Sequential delivery of immunomodulatory cytokines to facilitate the M1-to-M2 transition of macrophages and enhance vascularization of bone scaffolds. *Biomaterials.* 2015; 37:194–207. [PubMed: 25453950]
50. Tarnuzzer RW, et al. Biochemical analysis of acute and chronic wound environments. *Wound Repair Regen.* 1996; 4(3):321–325. [PubMed: 17177727]
51. Khanna S, et al. Macrophage dysfunction impairs resolution of inflammation in the wounds of diabetic mice. *PLoS One.* 2010; 5(3):e9539. [PubMed: 20209061]
52. Onken MD, et al. Gene expression profiling in uveal melanoma reveals two molecular classes and predicts metastatic death. *Cancer Res.* 2004; 64(20):7205–7209. [PubMed: 15492234]

53. Osareh A, et al. Classification and diagnostic prediction of cancers using gene microarray data analysis. *J. Appl. Sci.* 2009; 9(3):459–468.
54. Deb K, et al. Reliable classification of two-class cancer data using evolutionary algorithms. *BioSystems*. 2003; 72(1):111–129. [PubMed: 14642662]
55. Weingarten MS, et al. Diffuse near-infrared spectroscopy prediction of healing in diabetic foot ulcers: a human study and cost analysis. *Wound Repair Regen.* 2012; 20(6):911–917. [PubMed: 23110417]
56. Ng KC, et al. XAF1 expression is significantly reduced in human melanoma. *J. Invest. Dermatol.* 2004; 123(6):1127–1134. [PubMed: 15610524]
57. Kato S, et al. Mesenchymal-transitioned cancer cells instigate the invasion of epithelial cancer cells through secretion of WNT3 and WNT5B. *Cancer Sci.* 2014; 105(3):281–289. [PubMed: 24344732]
58. Young NA, et al. Use of flow cytometry for high-throughput cell population estimates in brain tissue. *Front. Neuroanat.* 2012; 6:27.doi: 10.3389/fnana.2012.00027 [PubMed: 22798947]
59. Šškrajnar Šš, et al. A flow cytometric method for determination of the blood neutrophil fraction in rats. *J. Am. Assoc. Lab. Anim. Sci.* 2009; 48(2):152. [PubMed: 19383211]
60. Tuana G, et al. Classification of dendritic cell phenotypes from gene expression data. *BMC Immunol.* 2011; 12(1):1. [PubMed: 21194488]
61. Spiller KL, et al. The role of macrophage phenotype in vascularization of tissue engineering scaffolds. *Biomaterials.* 2014; 35(15):4477–4488. [PubMed: 24589361]
62. Spiller KL, et al. Discovery of a novel M2c macrophage gene expression signature indicates a major role in human wound healing. *Wound Repair Regen.* 2015; 23(2) A40–A.
63. Fadok VA, et al. Macrophages that have ingested apoptotic cells *in vitro* inhibit proinflammatory cytokine production through autocrine/paracrine mechanisms involving TGF-beta, PGE2, and PAF. *J. Clin. Invest.* 1998; 101(4):890–898. [PubMed: 9466984]
64. Zhang QZ, et al. Human gingiva-derived mesenchymal stem cells elicit polarization of m2 macrophages and enhance cutaneous wound healing. *Stem Cells.* 2010; 28(10):1856–1868. [PubMed: 20734355]
65. Xue J, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity.* 2014; 40(2):274–288. [PubMed: 24530056]

Insight, innovation, integration

Macrophages, the primary cell of the innate immune system, act on a spectrum of phenotypes that correspond to diverse functions. Dysregulation of macrophage phenotype is associated with many diseases. In order to design effective treatments and diagnostics, characterization of the macrophage phenotype composition during tissue repair is necessary, but this task is complicated by the heterogeneous nature of macrophages themselves and also of tissue samples isolated from the wound environment. We describe a method to deconvolute gene expression from heterogeneous samples into the composition of two primary macrophage phenotypes (M1 and M2). This strategy allowed us to infer that non-healing diabetic ulcers contain higher proportions of M1 macrophages than healing ulcers, which may lead to new diagnostics and targeted treatments.

Fig. 1.
Experimental design overview.

Fig. 2.

Refinement of gene list based on overlap between three independently generated datasets (a) and by removal of genes expressed in other wound cell types (b).

Fig. 3.
Hierarchical clustering and expression trends of each gene across samples of varying M1 composition.

Fig. 4.

Changes in model predictive performance (average error) based on the number of genes included in training. The NN plot includes error bars as ten networks were created at each stage to account for variation in initial conditions. Mean $\pm$ standard deviation. Some error bars are too small to be seen.

Fig. 5. Predictions from a linear regression model (top) and neural network model (bottom) vs. known sample composition following feature selection.

Fig. 6. Neural network model predictions of the relative composition of M1 macrophage character in human diabetic ulcers. Mean $\pm$ standard deviation. Student's t -test, $*p < 0.05$.

Table 1

Datasets used to determine M1 and M2 markers

Dataset	Time in MCSF	Dose of MCSF	Dose of polarizing cytokines (M1)	Dose of polarizing cytokines (M2)	Time in cytokines before analysis	Method of analysis
PRJNA339309	7 days	20 ng ml ⁻¹	LPS: 100 ng ml ⁻¹ IFN γ : 100 ng ml ⁻¹	IL-4; 40 ng ml ⁻¹ IL-13; 20 ng ml ⁻¹	48 hours	RNA-seq
GSE57614	6 days	50 ng ml ⁻¹	LPS: 100 ng ml ⁻¹ IFN γ : 20 ng ml ⁻¹	IL-4; 20 ng ml ⁻¹	12–24 hours	Microarray
GSE36537	Not specified	Not specified	IFN γ : 20 ng ml ⁻¹	IL-4; 20 ng ml ⁻¹	18 hours	Microarray

Table 2

Subject information

Group	Subjects (n)	Samples analyzed (n)	Outcome	Gender	Age, mean (range)	BMI, mean (range)	Smoking status
Healing	2	4	Healed within 3 weeks of sample collection	50% M	62 (51–73)	31.9 (26.1–37.8)	50% smoker, 50% non- or former smoker
Non-healing	6	16	Amputated or remained open for more than 18 weeks	100% M	60 (57–64)	42.7 (33.1–50.1)	33% smoker, 33% non- or former smoker, 33% unknown

Table 3

Final list of M1 and M2 gene markers and a brief description of the function of the encoded protein based on information listed in NCBI's gene resource

M1 genes		M2 genes	
XAF1	Counteracts an inhibitor of apoptosis	CLEC4G	Glycan binding receptor involved in T cell immune response
CCL19	Involved in immune cell trafficking	CCL26	Involved in immunoregulation and inflammation
APOL1	High density lipoprotein	ALOX15	Maybe be involved in metabolism, iron-iron binding and oxidoreductase activity
ASPHD2	May play a role in dioxygenase activity	PLCB1	Involved in lipid metabolism
CLEC4D	Involved in cell adhesion, signaling, and inflammatory processes	CTNNAL1	May be involved in structural molecule activity
NOD2	Involved in immune response to LPS	RAMP1	Involved in transport of calcitonin receptor-like receptor
		WNT5B	Secreted signaling proteins

Table 4

Correlation coefficients and p -value for each gene across samples of varying composition from two donors, and in the combined dataset

Gene	Donor 1	p -Value	Donor 2	p -Value	Combined	p -Value
M1 genes						
APOL1 ^a	0.3559	0.0743	0.5493	0.0008	-0.1158	0.3781
ASPHD2	0.5952	0.0013	0.6292	0.0001	0.6297	0.0001
CCL19	0.4749	0.0142	0.4022	0.0184	0.3677	0.0039
CLEC4D ^a	-0.0241	0.9069	0.8550	0.0001	0.0859	0.5138
NOD2 ^a	0.1068	0.6035	0.5791	0.0003	-0.0090	0.9454
XAF1	0.7646	0.0001	0.7515	0.0001	0.7522	0.0001
M2 genes						
ALOX15	-0.8118	0.0001	-0.8380	0.0001	-0.6575	0.0001
CCL26	-0.8376	0.0001	-0.8371	0.0001	-0.7584	0.0001
CLEC4G	-0.7202	0.0001	0.3982	0.0197	-0.4035	0.0014
PLCB1	-0.7945	0.0001	-0.7191	0.0001	-0.6235	0.0001
RAMP1	-0.8585	0.0001	-0.2399	0.1718	-0.5988	0.0001
WNT5B	-0.4429	0.0234	-0.2751	0.1154	-0.3760	0.0031

^aSignifies genes that were not used in downstream analysis.

Table 5

Models trained using gene expression data from each donor and tested on itself (LOOCV), or when trained with one donor and tested on the other. D1: donor 1; D2: donor 2

Model	Average error when training and testing on same donor (LOOCV)		R^2 of linear fit for predictions vs. known composition		Average error when training on one donor and testing on the other		R^2 when training on one donor and testing on the other	
	D1	D2	D1	D2	D1	D2	D1	D2
Linear regression	0.157	0.154	0.945	0.848	0.191	0.169	0.577	0.757
Logistic regression	0.207	0.155	0.964	0.850	2.20	1.33	0.634	0.643
NN	0.018	0.012	0.998	0.996	0.224	0.221	0.654	0.501
SVM	0.038	0.035	0.987	0.970	0.179	0.258	0.718	0.634

Models trained using gene expression data from one or two donors and tested on heterogeneous samples consisting of cells from a third donor as well as fibroblasts and endothelial cells. D1: donor 1, D2: donor 2, D1 + D2: both donors

Table 6

Model	Average error when testing on heterogeneous samples			R^2 of linear fit for predictions vs. known composition		
	D1	D2	D1 + D2	D1	D2	D1 + D2
Linear regression	0.145	0.220	0.209	0.891	0.832	0.847
Logistic regression	1.45	1.25	0.911	0.839	0.823	0.730
NN	0.108	0.237	0.210	0.890	0.792	0.824
SVM	0.214	0.240	0.223	0.770	0.818	0.834