Molecular Systems Biology 5; Article number 265; doi:10.1038/msb.2009.15
Citation: ***Molecular Systems Biology*** **5:265**

www.molecularsystemsbiology.com

molecular systems biology

Predicting and controlling the reactivity of immune cell populations against cancer

Kfir Oved[1,5], Eran Eden[2,5], Martin Akerman[1], Roy Noy[1], Ron Wolchinsky[1], Orit Izhaki[3], Ester Schallmach[3], Adva Kubi[3], Naama Zabari[3], Jacob Schachter[3], Uri Alon[2,4], Yael Mandel-Gutfreund[1], Michal J Besser[3,*] and Yoram Reiter[1,*]

[1] Department of Biology, Technion Israel Institute of Technology, Haifa, Israel, [2] Department of Molecular Cell Biology, Weizmann Institute, Rehovot, Israel, [3] Ella Institute for Melanoma Research and Treatment, Sheba Medical Center, Tel-Hashomer, Israel and [4] Department of Complex Systems, Weizmann Institute, Rehovot, Israel

* Corresponding authors. R Yoram, Department of Biology, Technion Israel Institute of Technology, Haifa 32000, Israel. Tel.: +972 4 8292785; Fax: +972 4 829379; E-mail: reiter@tx.technion.ac.il or MJ Besser, Ella Institute for Melanoma Research and Treatment, Sheba Medical Center, Tel-Hashomer 52621, Israel. Tel.: +972 3 5304999; Fax: +972 3 5304922; E-mail: michal.Besser@sheba.health.gov.il

[5] These authors contributed equally to this work

Received 15.10.08; accepted 11.2.09

Heterogeneous cell populations form an interconnected network that determine their collective output. One example of such a heterogeneous immune population is tumor-infiltrating lymphocytes (TILs), whose output can be measured in terms of its reactivity against tumors. While the degree of reactivity varies considerably between different TILs, ranging from null to a potent response, the underlying network that governs the reactivity is poorly understood. Here, we asked whether one can predict and even control this reactivity. To address this we measured the subpopulation compositions of 91 TILs surgically removed from 27 metastatic melanoma patients. Despite the large number of subpopulations compositions, we were able to computationally extract a simple set of subpopulation-based rules that accurately predict the degree of reactivity. This raised the conjecture of whether one could control reactivity of TILs by manipulating their subpopulation composition. Remarkably, by rationally enriching and depleting selected subsets of subpopulations, we were able to restore anti-tumor reactivity to nonreactive TILs. Altogether, this work describes a general framework for predicting and controlling the output of a cell mixture.
Molecular Systems Biology 28 April 2009; doi:10.1038/msb.2009.15
Subject Categories: bioinformatics; molecular biology of disease
Keywords: decision tree algorithms; heterogeneous cell population; subpopulation signature; systems immunology; tumor immunology

Introduction

The collective output of heterogeneous cell populations is dictated by a complex network of interactions that span the molecular, cellular, and environmental levels (Benoist *et al*, 2006; de Visser *et al*, 2006). A major question is whether simple rules can be applied to predict and control the output of such cell populations, made of different cell types. One approach to address this question is to study the behavior of individual subpopulations composing the heterogeneous cell mixture in isolation. However, the behavior of each subpopulation is often altered in the presence of other cell types and factors. An alternative approach is to search for a set of explanatory genes and markers that correlate with the output of a cell mixture using high throughput technologies such as microarrays (Van't veer *et al*, 2002; Segal *et al*, 2004). Although microarrays are effective for analyzing homogeneous populations of cells, they loose their predictive power when applied to heterogeneous populations due to large variability and averaging effects. One useful approach for quantitatively measuring the percentage of different subpopulations characterized by specific receptors or markers in a cell mixture is flow cytometry (Perfetto *et al*, 2004). Flow cytometry was used earlier for the purpose of analyzing immune subpopulations to predict the clinical outcome and tumor metastasis in breast and colorectal cancer patients (Kohrt *et al*, 2005; Pages *et al*, 2005). Following this line of studies, we use multiple subpopulation flow cytometry measurements to ask whether the overall behavior of a heterogeneous population can be predicted and controlled by the composition of its subpopulation constituents.

As a system model for a heterogeneous population of cells, we use tumor-infiltrating lymphocytes (TILs). TILs are an

example of an immune population composed of different lymphocytic subpopulations that are derived from a tumor mass and have specificity and potential reactivity against the tumor. The phenomenon of circulating, as well as infiltrating lymphocytes in cancer patients was documented in a wide variety of solid tumors including primary brain tumors, epithelial cancers, and others (Vose *et al*, 1977; Romero *et al*, 1998; Lee *et al*, 1999). Presently, TILs are used in a clinical protocol called adoptive cell transfer for metastatic melanoma treatment. This treatment consists of TIL isolation from a tumor mass, their functional evaluation as determined by the levels of Interferon-γ (IFN-γ), expansion and re-injection back into the patient (Dudley *et al*, 2003). Because of their potential reactivity against tumors, TILs are extensively studied (Puisieux *et al*, 1994; Zhang *et al*, 2003; Terabe and Berzofsky, 2004; Bronte *et al*, 2005). Yet, despite their clinical importance, little is known about the underlying composition and cellular interactions that determine the degree of TIL reactivity and consequentially on how to control their reactivity.

In this study, we measure the subpopulation frequencies of TIL populations originating from different melanoma terminal patients. We show that none of the individual subpopulations examined in this study is capable of accurately predicting the collective TIL response to cancer cells. However, by using a simple computational model, one can generate a set of rules that accurately predict the degree of TIL reactivity in terms of its subpopulation constituents. Guided by these rules, we were able to control the reactivity of TILs by rational manipulation of their subpopulation composition. This enabled us to turn nonreactive TILs into reactive ones and vise versa (the workflow is summarized in Figure 1A).

Results

Different TILs show a wide range of anti-tumor reactivity levels even when extracted from the same patient

Metastatic melanoma tumors originating from the lymph nodes, lung, liver, spleen, and skin were surgically removed from 27 different patients. This study was approved by the Israeli Ministry of Health (Approval No. 3518/2004, ClinicalTrials.gov Identifier NCT00287131), and informed consent was obtained from all patients. At least two TILs were extracted from each patient tumor resulting in 91 different TILs (see Materials and methods; Supplementary Excel S1).

As a first step, the reactivity of the TILs was determined by measuring IFN-γ secretion, which is a major criterion for determining T-cell activity, after incubation with autologous melanoma. By using the clinical threshold of 200 pg/ml IFN-γ,

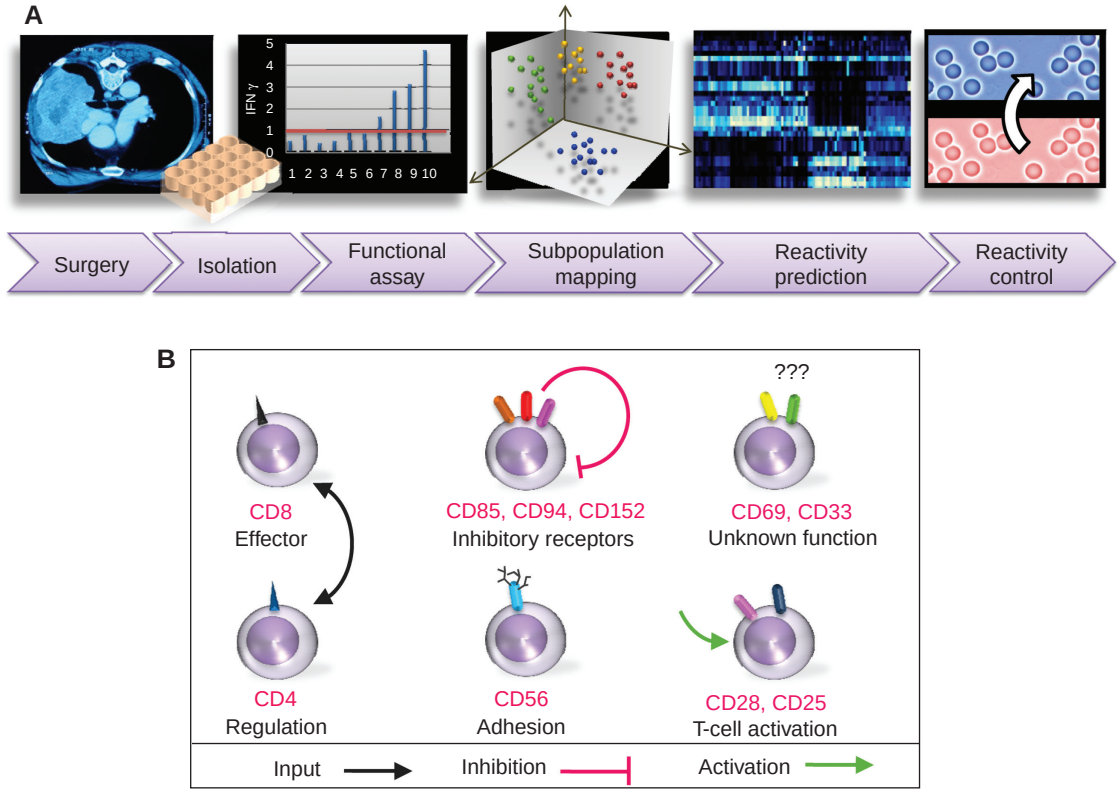

Figure 1 A schematic workflow of TIL characterization, analysis, and reactivity control. (**A**) TILs were extracted from surgically removed tumor mass originating from metastatic melanoma patients. Each TIL was characterized by functional evaluation of IFN-γ secretion levels followed by subpopulation fraction measurements using flow cytometry. This information was combined into a multiparametric model for prediction and rule-based description of TIL reactivity. Following this analysis, specific subpopulations were rationally selected for enrichment and depletion thus enabling control of TIL reactivity against melanoma. (**B**) Different cell surface receptors define specific T-cell subpopulations with distinct functional states. Some of these receptors are mutually exclusive (e.g. a mature T-cell will show either a CD8 or a CD4), whereas other receptors may appear simultaneously on the same cell.

39 TILs cultures were classified as reactive and 52 as nonreactive. Importantly, different TILs from the same patient and tumor mass, produced different reactivity levels varying from a potent to a null response (see Materials and methods).

TILs are composed of a wide range of subpopulations

The immune subpopulation compositions of the TIL cultures were characterized using standard multicolor flow cytometry by counting the number of cells in each subpopulation and calculating its fraction out of the entire population. The markers used for subpopulation mapping included combinations of triple staining from the pool of the following surface receptors: CD3, CD4, CD8, CD25, CD28, CD33, CD56, CD69, CD85, CD94, CD152 (Figure 1B) and the intracellular cytotoxic proteins, perforin and granzyme B. Each triple staining of three different receptors X, Y and Z resulted in 6 single staining (X^+, X^-, Y^+, Y^-, Z^+, Z^-), 12 double staining (e.g. X^+Y^-), and 8 triple staining (e.g. $X^+Y^+Z^-$). The single, double, and triple staining produces a hierarchy of subpopulation characterization ranging from general to more specific subpopulations. Overall, we measured 102 subpopulations for each of the TILs (Supplementary Table SI; Supplementary Excel S1). Each TIL is represented as a vector of subpopulation frequencies, for example, see Supplementary Figure S1. A quality control filtering procedure was used to omit subpopulations whose frequency was near the technical sensitivity limitation of flow cytometry (see Materials and methods), yielding a final dataset in which each TIL is characterized by 33 features: 31 distinct subpopulations, perforin and granzyme B (Supplementary Table SII; Supplementary Excel S2). We observed a wide range of frequency distributions for the different subpopulations (Supplementary Figure S2).

Individual subpopulations are partially predictive of TIL reactivity

To study whether any individual subpopulation can be used to differentiate between reactive and nonreactive TILs, we performed TIL reactivity prediction based on individual subpopulation frequencies (see Materials and methods). The

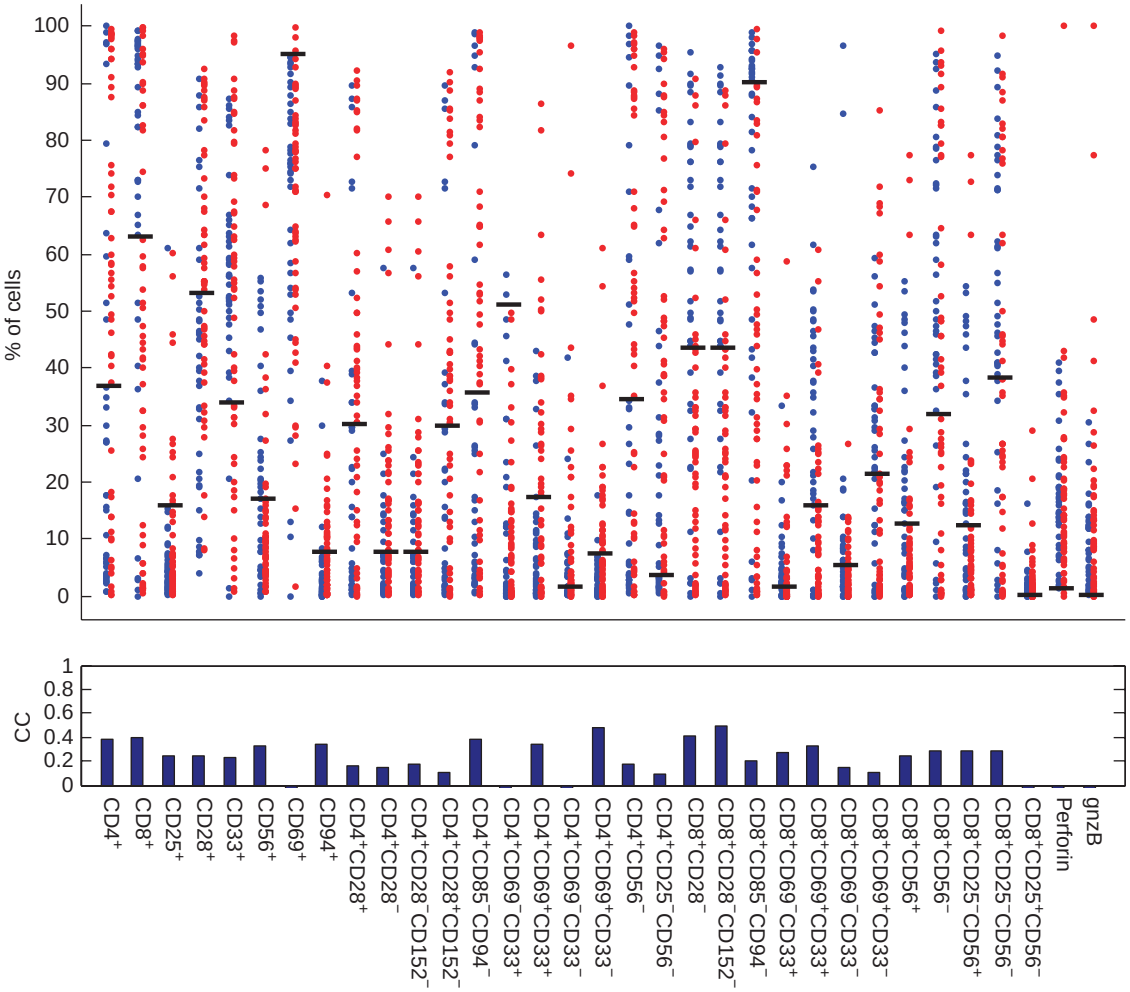

Figure 2 Individual subpopulations are partially predictive of TIL reactivity. For each subpopulation, blue and red dots indicate 39 reactive and 52 nonreactive TILs. The *y*-axis is the percentage of cells that belong to a specific subpopulation. The black horizontal bars indicate the optimal cutoff for classifying reactive and nonreactive TILs. The MCC classification accuracy of each subpopulation is shown at the bottom.

prediction accuracies of different subpopulations ranged between 0 and 0.49 in terms of Matthews correlation coefficients (MCCs) with sensitivity of up to 80% and a specificity of 68% (Figure 2; Materials and methods). These limited classification accuracies suggest that the frequency of any individual subpopulations is a limited predictor of the cell mixture collective output.

A combination of subpopulations accurately predicts TIL reactivity

To examine whether the incorporation of multiple subpopulations can significantly improve the reactivity prediction accuracy, we applied a support vector machine (SVM) model (see Materials and methods) (Noble, 2006). The prediction accuracy of the SVM model was MCC=0.74 (sensitivity of 91% and specificity of 88%) compared with an MCC=0.49 achieved by the best individual subpopulation. These results emphasize the importance of combining different subpopulation fractions for accurately predicting cell mixture reactivity, rather than looking at a specific subpopulation. These results are in accordance with the 'multiplayer' nature of the immune system (de Boer and Perelson, 1991; Benoist *et al*, 2006; Frankenstein *et al*, 2006; Noble, 2006). The misclassifications may be explained by flow cytometry sensitivity limitations, important subpopulations that were not measured, and the inherent stochasticity of the system. The fact that a high accuracy of prediction can be achieved by the SVM indicates that there is an underlying pattern connecting a combination of subpopulations and the collective TIL reactivity. However, the SVM model does not lend itself easily to biological interpretation. In the next two sections, we turn to investigate the underlying biological rational and subpopulation interplay, which govern TIL reactivity.

Reactive and nonreactive TILs have distinct subpopulation signatures

The usage of differential gene expression signatures has become a well-established method for distinguishing between various cellular states and different pathological conditions (Golub *et al*, 1999). We extend this concept to heterogeneous cell populations, by using a similar notion of a 'subpopulations signature.' Unsupervised hierarchical clustering was applied on the subpopulation signatures as shown in Figure 3, where each column corresponds to a TIL culture and the rows represent subpopulations. Two significant clusters emerge, each representing a profile of CD4 and CD8 enriched subsets. These two markers represent regulatory and cytotoxic T-cell subpopulations, respectively (Figure 1B). Interestingly, the two clusters also separate between nonreactive and reactive TILs (Fischer exact $P<10^{-3}$). This suggests that TIL reactivity

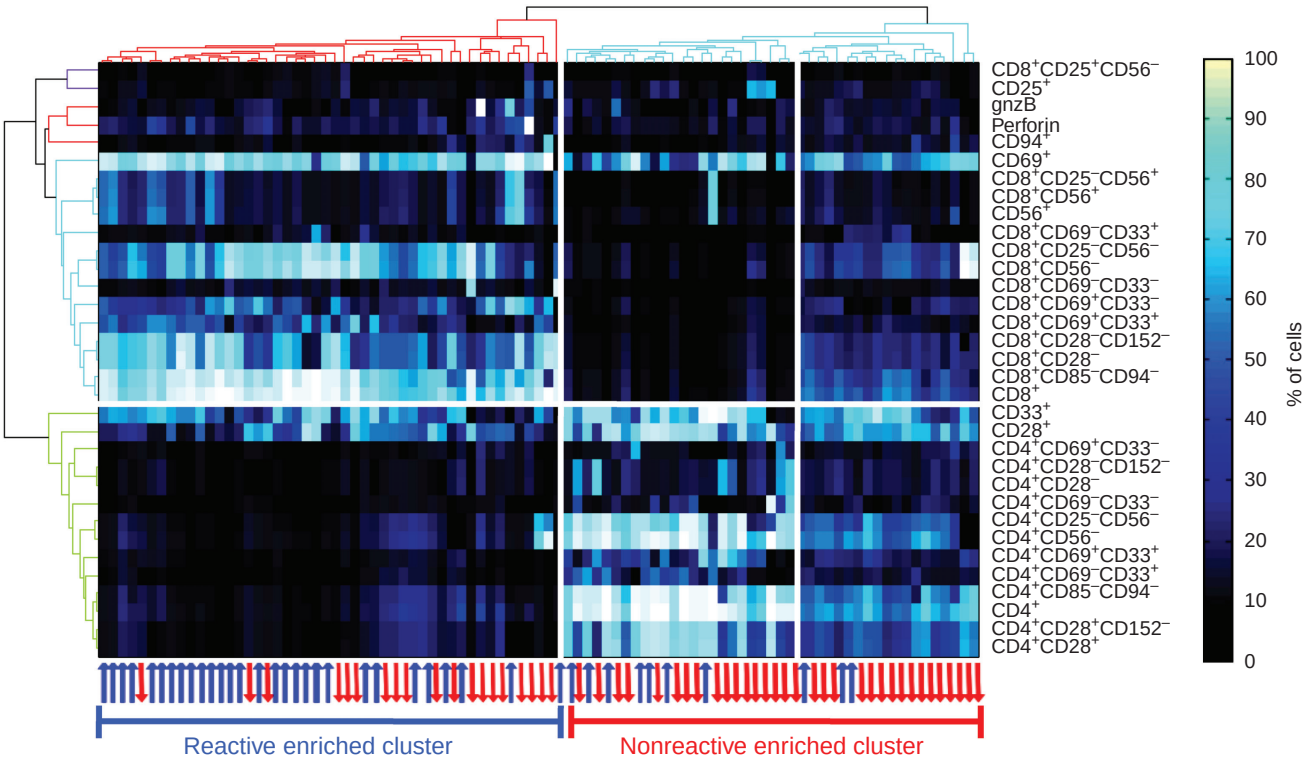

Figure 3 Reactive and nonreactive TILs exhibit distinct subpopulation signatures. Columns and rows correspond to TILs and subpopulations, respectively. Colors indicate the fraction of cells belonging to each subpopulation in each TIL. Unsupervised clustering was used on the rows and columns (see Materials and methods). The red and blue arrows represent nonreactive and reactive TILs, respectively. Two main clusters emerge characterized by $CD4^+$ and $CD8^+$ overabundant subpopulations. Interestingly, although the clustering procedure did not take into account TIL reactivity, the emerging clusters do separate nonreactive from reactive TILs ($P<10^{-3}$).

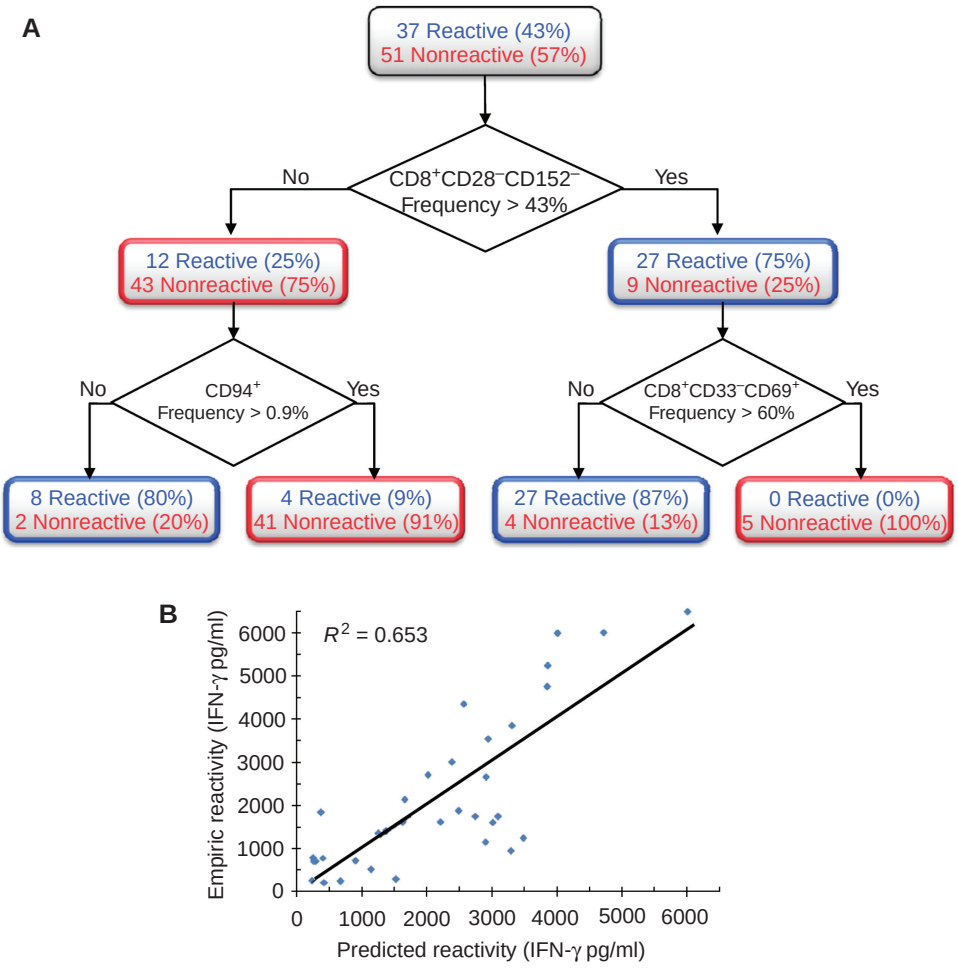

Figure 4 Simple rules based on subpopulation frequencies can predict TIL reactivity. (**A**) A decision tree algorithm was used to generate a simple set of four rules for classifying TIL functionality (see Materials and methods). Each rule is a path from the tree root (top) to one of the leaves (bottom). (**B**) IFN-γ levels of reactive TILs can be described as a function of two subpopulation fractions with positive and negative weights. Each dot is a reactive TIL. The *y*-axis is the empirical IFN-γ measurements and the *x*-axis is the theoretical IFN-γ levels calculated using the following model:

$$\mathrm{IFN}-\gamma(\mathrm{pg/ml}) = 63\cdot(\mathrm{CD8^{+}CD28^{-}}) - 50\cdot(\mathrm{CD8^{+}CD69^{+}CD33^{-}}) + 253.$$

Overall, IFN-γ levels of reactive TILs can be described to a large extent as a balance between two opposing subpopulations with positive and negative effects.

against melanoma is largely dictated by its subpopulation composition. We also observed that the nonreactive cluster is further divided into two subclusters, both of which are enriched with nonreactive TILs that have distinct profiles. The first is mostly CD4, whereas the other is a mixture of CD8 and CD4 subpopulation derivatives, suggesting CD4 dominance over CD8.

Reactivity and nonreactivity can be explained by simple subpopulation-based rules

To further simplify the subpopulation signature, we used a decision tree algorithm that scans through a large space of subpopulation-based rules and finds an optimal set of rules that can be used to predict TIL reactivity (see Materials and methods). This approach produced a simple set of four rules for distinguishing between reactive and nonreactive TILs (Figure 4A). The rules are based on the following questions:

Q1: Does $CD8^{+}CD28^{-}CD152^{-}$ subpopulation constitute more than 43% of the entire TIL population?
Q2: Does $CD94^{+}$ subpopulation constitute more than 0.4% of the entire TIL population?
Q3: Does $CD8^{+}CD33^{-}CD69^{+}$ subpopulation constitute more than 60% of the entire TIL population?

The four rules are:

R1: If Q1 is false and Q2 is false then the TIL is classified as reactive.
R2: If Q1 is false and Q2 is true then the TIL is classified as nonreactive.
R3: If Q1 is true and Q3 is false then the TIL is classified as reactive.

R4: If Q1 is true and Q3 is true then the TIL is classified as nonreactive.

The accuracy of the rule-based predictions is MCC=0.69 (sensitivity of 87% and specificity of 83%). Although these rules yield an accuracy that is slightly reduced compared with the SVM model (MCC=0.74), their advantage is that they are simple and amenable to interpretation. These rules highlight three subpopulations, namely: $CD8^+CD28^-CD152^-$, $CD94^+$, and $CD8^+CD69^+CD33^-$. The first emphasizes the role of the CD28 and 152 receptors in determining the TIL reactivity in addition to CD8. Specifically, our observation that reactive CD8 T-cells lack both co-stimulatory CD28 receptor and the co-inhibitory receptor CD152 on their surface is in agreement with current knowledge. CD28 is down regulated and internalized after proper T-cell activation (Linsley *et al*, 1993; Eck *et al*, 1997; Alegre *et al*, 2001). The absence of CD152 receptor on reactive TILs is in accordance with its potent co-inhibitory role. The second subpopulation is marked by CD94, an inhibitory receptor expressed in low levels on T-cells (Leibson, 2004). Its inhibitory function may explain why higher levels are correlated with nonreactive TILs. The third subpopulation ($CD8^+CD69^+CD33^-$) is characterized by the CD69 and lack of CD33 receptor staining. Little is known about the specific function and ligands of these two receptors and in particular on the role of CD33 in T cells (Nakamura *et al*, 1994; Shiow *et al*, 2006). These findings suggest that CD33 expression on the surface of T cells is positively correlated with anti-tumor T-cell functionality. In addition, the expression of CD69, which is an early activation marker in T cells, was surprisingly correlated with nonreactivity.

To test whether subpopulation analysis can be used, not only to classify between reactive and nonreactive TILs, but also to predict the exact level of IFN-γ secretion, we focused exclusively on the reactive TILs. To this end, we performed a linear regression on pairs of subpopulations and IFN-γ levels. By using an equation of the form $IFN-\gamma = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2$, where X_1 and X_2 represent the fraction of the two different subpopulations, we were able to accurately determine the exact levels of IFN-γ with $P < 10^{-4}$ (see Figure 4B). After scanning all possible pair combinations, the pair that yielded optimal results, in terms of IFN-γ secretion was $CD8^+CD28^-$ and $CD8^+CD69^+CD33^-$. Notably, these subpopulations are highly similar to those used for classification of reactive and nonreactive TILs in the decision tree (Figure 4A). This shows that a simple model that takes into account the amount of cells belonging to selected subpopulations can be used to quantitatively infer the degree of cell mixture reactivity.

A TIL's reactivity can be controlled by rational manipulation of its subpopulation composition

Overall, these results indicate that TIL anti-tumor reactivity is too complex to be explained by an individual subpopulation. Yet, the combination of a few subpopulations-based rules and simple formulas can explain the reactivity to a large extent. These observations raise the conjecture whether one could use these rules to control reactivity of TILs by manipulating their subpopulation fractions. Specifically, we hypothesized that nonreactive TILs can be turned into reactive ones by depleting nonreactive-associated subpopulations and vice versa; that reactive TILs can be turned into nonreactive ones by depleting reactive-associated subpopulations. To test the first hypothesis, we used specific antibodies to selectively deplete the subpopulations CD4, CD28, CD85, CD94, and CD152 that we found to be associated with nonreactivity. The experiments were performed on a fresh cohort of 12 nonreactive TIL cultures that originated from four melanoma patients (Supplementary Table SIII) and were not part of the 91 TIL samples used for the subpopulation signature elucidation. First, we characterized the subpopulation frequencies of each TIL and its anti-tumor reactivity. All of the 12 fresh TILs were nonreactive with IFN-γ levels below the 200 pg/ml clinical threshold. We then depleted the inhibitory-related subpopulations in each TIL using specific antibodies and magnetic bead negative selection (see Materials and methods; Supplementary Excel S3). After 36 h of recovery, both original and manipulated TILs were challenged with autologous melanoma for 12 h followed by supernatant IFN-γ measurement. Remarkably, 9 of the 12 originally nonreactive TILs became reactive after manipulation (Figure 5A; Supplementary Table SIII). The IFN-γ level of the nine reactive TILs showed a dramatic increase with levels exceeding the 200 pg/ml threshold and ranging between ~300 and ~4000 pg/ml. Overall, the reactivity before and after manipulation yielded up to 106-fold increase in IFN-γ secretion. Two of the three TILs that retained a nonreactive state after manipulation also exhibited an increase in IFN-γ levels, but remained under the threshold of 200 pg/ml. The degree of postmanipulation reactivity of TILs was independent of their subpopulation profiles before manipulation. For example, TILs No. 2 and 9 shared similar subpopulation profiles before manipulation (Figure 5B), but different profiles and reactivity levels (4020 and 295 pg/ml) after manipulation (Figure 5A; Supplementary Table SIII).

As negative controls, we tested the specificity and spontaneous release of IFN-γ secretion by incubating the TILs with unrelated melanoma, culture media (Figure 5A), or mock antibodies (data not shown). In all controls, IFN-γ levels remained unchanged and below threshold, indicating the specificity of the procedure and excluding the possibilities of IFN-γ secretion increase that is due to spontaneous release or a nonspecific activation.

Next, we tested the complimentary hypothesis that reactive TILs can be turned into nonreactive ones by depletion of reactive-associated subpopulations. To test this, we used five fresh TILs that were not used in the earlier samples and all of which were highly reactive. After depletion of their reactive-associated subpopulations by using the CD8, CD56, and CD33 specific antibodies, all the TILs showed a dramatic decrease in reactivity ranging from ~10 to ~460-fold (Supplementary Figure S5A). Negative controls for testing specificity and excluding spontaneous release of IFN-γ secretion were performed by incubating the TILs with unrelated melanoma, culture media (Supplementary Figure S5A), or mock antibodies (data not shown). In all cases, no significant change in reactivity or subpopulation composition was observed.

The fact that nonreactive TILs can be transformed into reactive ones and vice versa after a manipulation of their

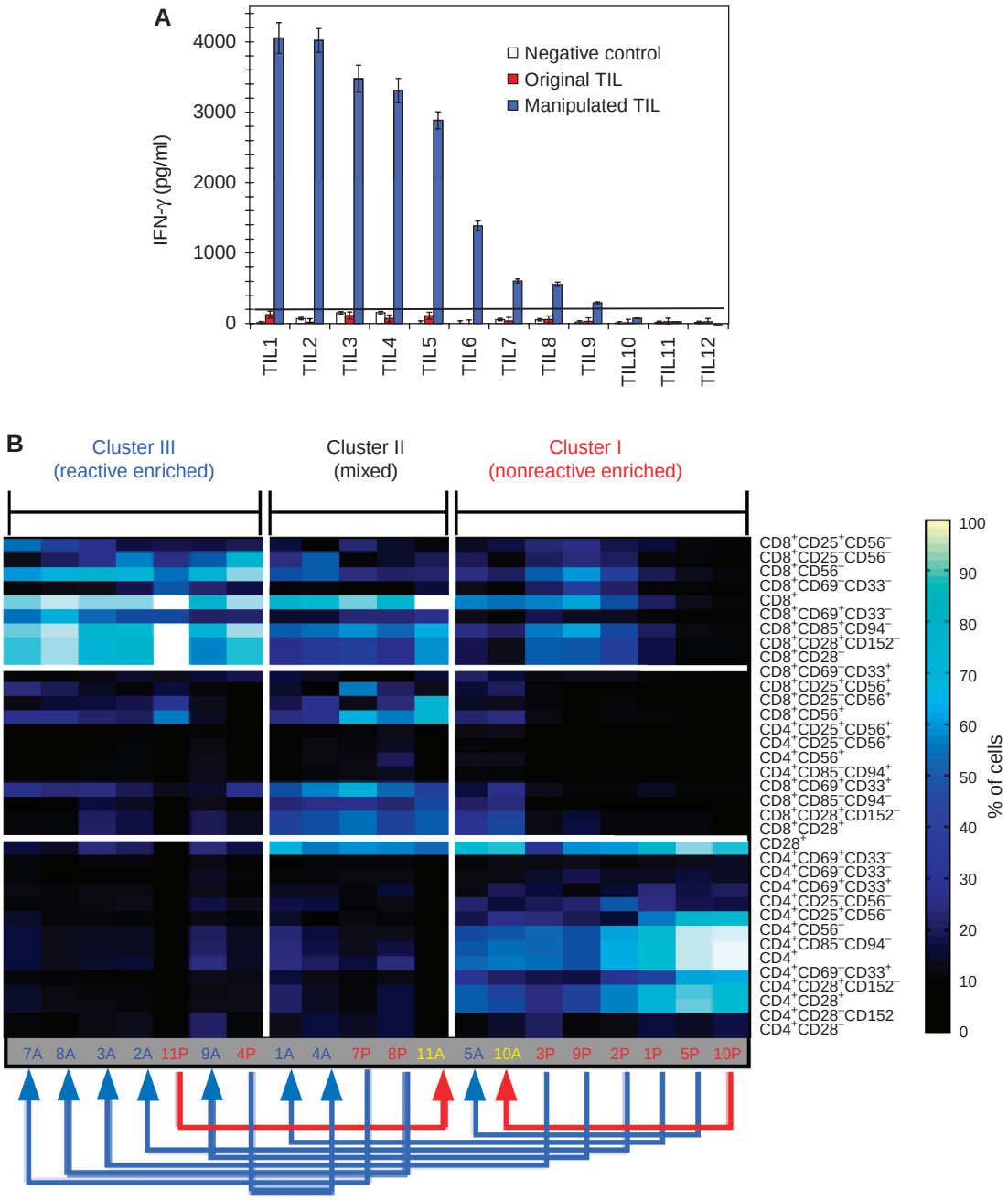

Figure 5 Rational subpopulation manipulation can change TIL anti-tumor reactivity and is accompanied by a shift in subpopulation signature. (**A**) IFN-γ levels of 12 TILs before (red bars) and after (blue bars) rational subpopulation depletion and enrichment are compared. Nine of the original nonreactive TILs show significant increase in IFN-γ levels with up to 106-fold increase observed for TIL #2. Incubation of TILs in the control experiments with culture media or unrelated melanoma (white bars) indicates that the increase in IFN-γ secretion does not occur spontaneously and is tumor (HLA) specific. (**B**) The shift in reactivity can be explained in terms of a shift in subpopulation signature. The subpopulation fractions of 10 TILs before and after subpopulation manipulation are shown. The rows and columns correspond to different subpopulations and TILs, respectively. Two ways unsupervised clustering was performed on the rows and columns. The 10 nonreactive TILs prior to the manipulation are designated by a red color and the letter 'P.' Ten TILs after manipulation are designated by the letter 'A' with blue and yellow corresponding to reactive and nonreactive, respectively. Eight of the nonreactive TILs before manipulation became reactive, seven of which also showed a shift from a nonreactive subpopulation signature to a reactive one as indicated by the blue arrows going from right to left. The two TILs that remained nonreactive after manipulation exhibited either a minor change or a negative change in subpopulation signature as indicated by the red arrows. (**C**) The transformation of a nonreactive TIL to a reactive one can be described as a path between two points in the subpopulation space. In order to visualize the TILs positions in the multidimensional subpopulation space, we applied PCA, which is a method for dimensionality reduction (see Materials and methods). This enabled us a simple 2D visualization of the different TILs. The *x* and *y*-axes are the principal components capturing 49 and 24% of the total variance in the data. The *x*-axis captures a shift from $CD8^+$ and $CD28^-$ enriched subpopulation to $CD4^+$ and $CD28^+$ subpopulations, whereas the *y*-axis reflects a combination of additional subpopulations (see Supplementary Figure S4 for subpopulations coefficients defining each of the principal components). The figure shows a subspace region that is overpopulated with reactive TILs. The change in reactivity can be visualized as a path from a nonreactive TIL to a TIL that resides in the reactive subspace (e.g. see dotted arrow).

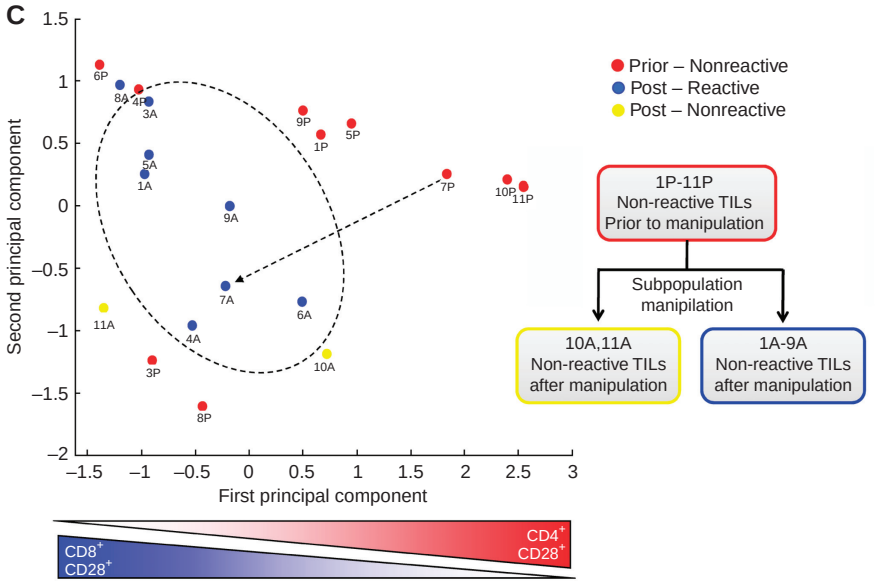

Figure 5 Continued.

subpopulation balance suggests that nonreactivity is largely dictated by simple subpopulation interactions rather than lack of specificity to melanoma cancer epitopes. These results further emphasize the importance of interplay between different subpopulations in determining the TIL collective output.

The shift between nonreactive and reactive states is accompanied by a transformation of the subpopulation signature

To link the change in reactivity with the change in the underlying subpopulation composition, we examined the TIL profiles before and after the manipulation. For this analysis, we used 10 of the 12 TILs that had sufficient cell counts. To rule out the possibility that the shift from nonreactive to reactive state is simply proportional to change in the amounts of the major subpopulations CD4 or CD8, we monitored their amount before and after manipulation (Supplementary Table SIII). The observed average-fold change was 0.7 ± 0.71 for CD4 and 2.5 ± 2.4 for CD8 compared with a 34.8 ± 31.3-fold change of IFN-γ secretion.

A more global comparison of all subpopulations showed that the profiles of 9 of the 10 nonreactive TILs before manipulation were similar to that of the nonreactive TIL profiles as determined by the original 91 sample dataset (compare Figure 5B with Figure 3). Furthermore, the shift from a nonreactive to reactive state was accompanied by a shift from a nonreactive to reactive subpopulation signature. Similarly, a transition from reactive to a nonreactive state was also accompanied by a shift to a nonreactive subpopulation signature. This point is illustrated in Figure 5B and C and Supplementary Figure S5B, C.

Discussion

Predicting and controlling the output of a heterogeneous cell population is a highly challenging task with many biological and clinical implications. To study this question, we chose TILs, a heterogeneous immune cell population composed of different immune subpopulations that have a quantitatively measurable output in terms of their degree of reactivity against tumors. Because of their clinical importance, TILs have been extensively studied, yet little is known about the factors that govern their collective degree of reactivity.

This work describes for the first time an experimental and computational approach that accurately predicts the *in vitro* reactivity of TILs that were extracted from tumors of melanoma patients. This was done by experimentally measuring a unique set of cell surface receptor combinations that define a rich set of subpopulations composing the TIL. Although some of these receptor combinations define well established and highly studied subpopulations, most of the receptor combinations used herein defines previously unstudied subpopulations. Using a quantitative measure of the subpopulation constituents of each TIL, we constructed a computational model that accurately predicts TIL reactivity with 91% sensitivity.

We identified strong correlations in expression of different receptors on individual cells. This co-expression may suggest co-regulation of receptors with related function. For example, the expression levels of the receptors CD85, CD94, and CD152, all of which play inhibitory roles, were significantly correlated (average r=35%), suggesting that these receptors might have common regulation (Supplementary Figure S3). The existence of such correlations simplifies the way one can describe this complex biological apparatus and sets a direction for beginning to understand receptors with an unknown function.

Although in theory the number of possible subpopulation combinations that compose a TIL is large, in practice TILs fall into a few distinct profiles (Figure 3). These defined signatures may imply that under equilibrium different T-cell subpopulations balance one another. We were able to find a simple set of rules that map between a TIL's subpopulation composition and its reactivity (Figure 4A) with an 87% sensitivity. The rules suggest that the degree of reactivity is determined by the balance between subpopulations with effector or inhibitory functions. Furthermore, we show that the same subpopulations can be used, not only to differentiate between reactive and nonreactive TILs, but also to predict the exact level of IFN-γ secretion of the reactive ones (Figure 4B). The rules also highlight the importance of subpopulations with previously uncharacterized functions, such as the subpopulation marked by the CD33 receptor. CD33 is a central myeloid marker that has been identified earlier on the surface of T-lymphocyte (Nakamura *et al*, 1994). Yet, its function and mechanism of action are not well understood. This study shows that CD33-marked subpopulations are correlated with positive anti-tumor reactivity.

This study focuses on predicting and controlling *in vitro* IFN-γ collective output of TILs. It is important to stress that IFN-γ secretion of TILs does not always correlate with clinical response. One interesting future direction of this work is to investigate whether this approach can be used to compose a set of TIL subpopulation-based rules for predicting patient response. Such a prediction is expected to be more difficult than *in vitro* TIL reactivity prediction and will require the incorporation of additional factors such as patient-related background and tumor-related factors.

Our finding that simple subpopulation-based rules can predict TIL reactivity raised the conjecture of whether one might be able to control the overall population output by rational manipulation of its subpopulation constituents. Guided by these rules, we selectively enriched and depleted specific subpopulations of a fresh cohort of nonreactive TILs extracted from melanoma patients. We show for the first time that nonreactive TILs can become anti-tumor reactive after rational manipulation. By manipulating the subpopulation composition in the opposite way, we were also able to transform reactive TILs into nonreactive ones. These results open the way for future studies in which rational subpopulation manipulation of TILs may be used to control their effectiveness in the clinical setting.

Another aspect of this result is that it sheds light on the reason for lack of TIL reactivity *in vitro*. There are three possible explanations why some TILs do not respond against tumors: First, nonreactivity may result from an inappropriate T-cell receptor repertoire that lacks tumor specificity. Second, TILs may not be able to respond because of antigen deletion or down regulation by the tumor. Last, TILs may be actively suppressed by the tumor cells and their microenvironment. Our data support the last explanation of active suppression. We show that by excluding a few defined subpopulations of the TIL itself, it is possible to elicit a potent response in previously nonresponsive TILs. These results support the possibility that the anti-tumor reactivity of T-cells is mainly suppressed due to inhibiting effects of some of its subpopulations constituents and are in accordance with earlier studies that linked between specific immune populations and tumor metastasis or patient prognosis (Chiba *et al*, 2004; Kohrt *et al*, 2005; Pages *et al*, 2005).

In the future, the subpopulation-based framework described herein can also be extended to predict and control the outputs of other types of heterogeneous cell populations in fields such as stem cells, tumor immunology, and tissue engineering.

Materials and methods

TIL and melanoma cultures

Metastatic melanoma tumors were surgically removed from 27 different patients. The isolation and expansion of TILs performed in this study closely follows an earlier published protocol (Dudley *et al*, 2003). Briefly, tumor tissue from metastatic melanoma patients was surgically removed. The tumors were extracted from lymph nodes, lung, liver, spleen, and skin metastases. Every patient signed an informed consent approved by the Israeli Ministry of Health (Approval No. 3518/2004). Various techniques, such as homogenization, fragmentation, fine needle aspiration, and enzymatic digestion, were used to obtain melanoma cultures. TILs cultures were established by dissecting tumor specimens into 1–2 mm^3 fragments. Each fragment was placed in a different well (we used a 24-well tissue culture plate) with 2 ml of complete medium, comprised of RMPI 1640 (Lonza, Verviers Sprl, Belgium) containing 10% human serum (Gemini Bio-Products, West Sacramento, USA; Blood bank, Magen David Adom, Tel-Hashomer, Israel), 25 mmol/l HEPES pH 7.2 (Lonza), 100 U/ml penicillin (Lonza), 100 μg/ml streptomycin (Lonza), 50 μg/ml gentamycin (Lonza) and 5.5×10^{-5} mol/l 2-mercaptoethanol (Invitrogen Corp., Paisley, UK) and 6000 IU/ml IL-2 (Proleukin, Chiron BV, Amsterdam, The Netherlands). After 4–6 days, a dense TIL carpet was observed around the fragment and 8–12 days after initiation lymphocytes covered the entire well. Each TIL culture was maintained independently and split into two daughter wells to maintain a cell concentration of $0.5–2.0 \times 10^6$ cells/ml. Individual TIL cultures reached a cell number of $30–60 \times 10^6$ cells after 24.5 ± 6.6 (stderr) days.

Measuring IFN-γ release after co-culture with melanoma cells

The activity and specificity of each TIL was determined by an IFN-γ release assay. After washing, TILs were co-cultured with autologous melanoma cells, or HLA-mismatched melanoma lines as negative control, at an effector target ratio of 1:1 (1×10^5 each) in a 96-well plate for 14–16 h. Culture medium contained 10% fetal bovine serum (Invitrogen), 25 mmol/l HEPES pH 7.2, 100 U/ml penicillin, 100 μg/ml streptomycin, and 2 mmol L-glutamine, 1 mM sodium pyruvate (Lonza) in RPMI 1640. Cells were centrifuged, supernatant was removed and the secreted IFN-γ levels were determined by sandwich enzyme-linked immunosorbent assay and CBA cytokine bead array (R&D Systems, BD biosciences). All the assays were done with at least two biological independent replicates, within each of which we performed at least two technical replicates. The clinical threshold of 200 pg/ml IFN-γ was used to classify each TIL as reactive or nonreactive. Importantly, different TILs from the same tumor mass, exhibited a wide spectrum of biological activity ranging from potent to null response (Supplementary Table SI).

Characterization of TIL subpopulation composition using flow cytometry

Various combinations of triple staining from the pool of the following antibodies were used for subpopulation mapping: αhCD4, αhCD25, αhCD28, αhCD56, αhCD69, αhCD85, αhCD94 (DakoCytomation), αhCD152 (Serotec) and αhCD3, αhCD33, αhCD8 (BD biosciences), and αh-perforin and αh-granzyme-B (eBiosciences). For detailed list of all the subpopulations that were measured see Supplementary

Table SI. For flow cytometric analysis of cell surface receptors, 2.5×10^5 cells were washed and re-suspended in PBS containing 0.1% BSA. Cells were incubated on ice with the appropriate conjugated antibodies for 20 min and were subsequently washed three times with cold PBS containing 0.1% PBS. Samples were analyzed using a FACScaliber machine (BD Biosciences). Background staining was assessed using an isotype control antibody.

Dataset filtering procedure

To remove patient interdependence in the dataset, in case two or more TILs originating from the same patient showed a similar reactivity level and subpopulation composition only one representative TIL was kept. Starting from 101 TILs, 10 TILs were excluded resulting in the final set of 91 different TILs (Supplementary Table SI). The remaining TILs exhibit a wide range of functional and phenotypic variations even when extracted from the same patient and tumor mass.

Initially, each of the 91 TILs was characterized by measuring 102 different subpopulations (raw data is given in Supplementary Excel S1). Only subpopulations adhering to the following criteria were kept in the dataset: (1) the average fraction of the subpopulation in all TILs was $>1\%$ and (2) the individual subpopulation-based classification threshold was between 1 and 99% (see details below). These steps eliminated subpopulations whose amount was near the experimental technical sensitivity of the flow cytometry. The final filtered dataset included 33 subpopulations for 91 TILs (Supplementary Table SII and Excel S2).

TIL classification based on individual subpopulations

The individual subpopulations-based classification procedure was performed by finding the optimal subpopulation frequency cutoff between reactive and nonreactive TILs in terms of the MCC on a train set and measuring the performance on a previously unseen test set (results are summarized in Figure 2). The MCC is defined as follows:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FN)(TP+FP)(TN+FN)(TN+FP)}}$$

Where *TP*, *FP*, *TN*, *FN* are true positives, false positives, true negatives, and false negatives, respectively. In general, MCC values range between -1 to $+1$, indicating completely wrong and perfect classification, respectively. An MCC of 0 indicates random classification. It has been shown earlier that the MCC is especially useful for measuring and optimizing classification accuracy in cases of unbalanced class sizes (Pierre Baldi, 2001) (in our case the 39 of the TILs are reactive and 52 are nonreactive). The MCC measure assigns a proportional weight to instances from each set (each instance of the smaller set receives a larger weight). The sensitivity and specificity were computed as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \ \text{Specificity} = \frac{TN}{TN+FP}$$

The classification accuracy results were estimated using a leave-one-out cross-validation, where each TIL was excluded once from the train set and used as a test set. In addition, to rule out interdependence between TILs originating from the same patient, we performed a leave-one-patient-out cross-validation. During each iteration, all the TILs that were extracted from the tumor mass of the same patient were excluded from the train set and used for testing. This yielded similar results as the original leave-one-out cross-validation on the TIL level.

SVM classification

We applied a SVM model to predict whether TILs are reactive or nonreactive based on subpopulation fractions (CJC Burges, 1998). Briefly, each TIL is mapped to a point in a multidimensional space according to its subpopulation fractions. The SVM classifier generates a hyper-surface that separates reactive and nonreactive points. Input data were normalized by linearly rescaling the subpopulation frequencies of each TIL to values between −1 and 1. Classifications were done with a linear kernel and the accuracy was assessed using a leave-one-out procedure. All SVM classifications were performed using the gist-train-svm software http://bioinformatics.ubc.ca/gist/.

TILs and subpopulation clustering

Two ways clustering was performed using an agglomerative hierarchical clustering scheme with average linkage. The analysis was performed using the Matlab statistical and bioinformatics toolbox software (for details see Matlab Bioinformatics and Statistics toolbox, a user's guide). The Spearman correlation between each pair of TILs (i.e. the correlation between the corresponding column vectors of subpopulation frequencies) was used as a distance measure for clustering the columns (see Figure 3). Similarly, the distance between subpopulations was computed using a Spearman correlation on the corresponding rows. The two main clusters were generated by cutting the TIL dendrogram at the top most linkage. Spearman correlation was chosen rather than the alternative distance measures such as Pearson correlation or Euclidean distance, as it is less sensitive to outliers and resulted in a more significant Fischer exact enrichment of reactive and nonreactive TILs in each of the main clusters.

Decision tree algorithm

The decision tree in Figure 4A was generated using the Matlab statistics toolbox software. Briefly, the algorithm iteratively selects a subpopulation and the threshold that best classifies instances of the train set into distinct classes thus generating a tree-like structure. Here, we use the Gini impurity function (Breiman, 1993) as a quantitative measure for subpopulation classification accuracy. First, the tree root is selected by identifying the subpopulation that best separates TILs into reactive and nonreactive. The tree root splits all the TIL samples into two mutually exclusive subgroups with reduced class impurities. Each of the subgroups is then further split in an iterative manner with different subpopulations and thresholds using the same impurity minimization principle. Here, we only allowed splits on nodes that contain more than 10 samples. The resulting tree is then automatically pruned to avoid model over-fitting (for details see Matlab Statistics toolbox, a user's guide). Model accuracy was evaluated using a leave-one-TIL-out and leave-one-patient-out cross-validation (for details see Section: TIL classification based on individual subpopulations). We note that the threshold of 200 pg/ml that is used in the clinic for classifying a TIL reactivity is arbitrary. Therefore, we tested the sensitivity of our model to other thresholds ranging from 150 to 500 pg/ml and found that the prediction accuracy as well as the subpopulations that were identified as good separators remained similar within this threshold range.

Enrichment and depletion of specific subpopulations

T-cell depletion was obtained by incubating the TILs with anti-CD4, anti-CD28, anti-CD152, anti-CD85, and anti-CD94 for 20 min. Subsequently, cells were mixed with anti-mouse IgG-coated magnetic beads (Dynal, Lake Success, NY) for additional 10 min followed by magnetic depletion for 5 min. The negative fraction was then washed three times with PBS 0.1% BSA and was incubated for 36 h recovery in 37°C.

Principal component analysis of TILs

Principal component analysis (PCA) is a commonly used method for transforming multidimensional data to lower dimensions at the expense of loosing part of the data variance (IT Jolliffe, 2002). The PCA procedure used here was implemented as part of the Matlab statistics toolbox. Briefly, each TIL, which is defined by a multidimensional vector of its subpopulation fractions, was reduced to a 2D entity using two principal components where each principal component is a linear combination of original subpopulations. The exact contributions of each subpopulation to each principal component are specified in Supplementary Figure S4.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb).

Acknowledgements

We thank Professor Zohar Yakhini from the Technion and Shay Sela from the Hebrew university for critical reading and discussions, Elad Oved from the Technion for competent graphical assistance and figure design. KO conceived and designed the experiments, performed the experiments, analyzed the data, and wrote the paper. EE and MA analyzed the data and wrote the paper. RN and RW made intellectual donation. OI, ES, AK, and NZ contributed reagents/materials/analysis tools. YS is the coordinating physician. UA and YM wrote the paper. MB conceived and designed the experiments. YR conceived and designed the experiments and wrote the paper. We thank Nechemia and Chaya Lemelbaum for their financial support. This work was supported in part by a research grant from the Israel Science Foundation (granted to YR).

Conflict of interest

The authors declare no conflict of interest.

References

Alegre ML, Frauwirth KA, Thompson CB (2001) T-cell regulation by CD28 and CTLA-4. *Nat Rev Immunol* **1:** 220–228

Benoist C, Germain RN, Mathis D (2006) A plaidoyer for 'systems immunology'. *Immunol Rev* **210:** 229–234

Breiman L (1993) *Classification and Regression Trees*. Boca Raton: Chapman & Hall

Bronte V, Kasic T, Gri G, Gallana K, Borsellino G, Marigo I, Battistini L, Iafrate M, Prayer-Galetti T, Pagano F, Viola A (2005) Boosting antitumor responses of T lymphocytes infiltrating human prostate cancers. *J Exp Med* **201:** 1257–1268

Chiba T, Ohtani H, Mizoi T, Naito Y, Sato E, Nagura H, Ohuchi A, Ohuchi K, Shiiba K, Kurokawa Y, Satomi S (2004) Intraepithelial CD8 + T-cell-count becomes a prognostic factor after a longer follow-up period in human colorectal carcinoma: possible association with suppression of micrometastasis. *Br J Cancer* **91:** 1711–1717

CJC Burges (1998) A Tutorial on Support Vector Machines for Pattern Recognition. Boston, MA, USA: Kluwer Academic Publishers

de Boer RJ, Perelson AS (1991) Size and connectivity as emergent properties of a developing immune network. *J Theor Biol* **149:** 381–424

de Visser KE, Eichten A, Coussens LM (2006) Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer* **6:** 24–37

Dudley ME, Wunderlich JR, Shelton TE, Even J, Rosenberg SA (2003) Generation of tumor-infiltrating lymphocyte cultures for use in adoptive transfer therapy for melanoma patients. *J Immunother* **26:** 332–342

Eck SC, Chang D, Wells AD, Turka LA (1997) Differential down-regulation of CD28 by B7-1 and B7-2 engagement. *Transplantation* **64:** 1497–1499

Frankenstein Z, Alon U, Cohen IR (2006) The immune-body cytokine network defines a social architecture of cell interactions. *Biol Direct* **1:** 32

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286:** 531–537

IT Jolliffe (2002) *Principal Component Analysis*. New York, USA: Springer-Verlag

Kohrt HE, Nouri N, Nowels K, Johnson D, Holmes S, Lee PP (2005) Profile of immune cells in axillary lymph nodes predicts disease-free survival in breast cancer. *PLoS Med* **2:** e284

Lee PP, Yee C, Savage PA, Fong L, Brockstedt D, Weber JS, Johnson D, Swetter S, Thompson J, Greenberg PD, Roederer M, Davis MM (1999) Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nat Med* **5:** 677–685

Leibson PJ (2004) The regulation of lymphocyte activation by inhibitory receptors 8. *Curr Opin Immunol* **16:** 328–336

Linsley PS, Bradshaw J, Urnes M, Grosmaire L, Ledbetter JA (1993) CD28 engagement by B7/BB-1 induces transient down-regulation of CD28 synthesis and prolonged unresponsiveness to CD28 signaling. *J Immunol* **150:** 3161–3169

Nakamura Y, Noma M, Kidokoro M, Kobayashi N, Takei M, Kurashima S, Mukaiyama T, Kato S (1994) Expression of CD33 antigen on normal human activated T lymphocytes. *Blood* **83:** 1442–1443

Noble WS (2006) What is a support vector machine? *Nat Biotechnol* **24:** 1565–1567

Pages F, Berger A, Camus M, Sanchez-Cabo F, Costes A, Molidor R, Mlecnik B, Kirilovsky A, Nilsson M, Damotte D, Meatchi T, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, Galon J (2005) Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* **353:** 2654–2666

Perfetto SP, Chattopadhyay PK, Roederer M (2004) Seventeen-colour flow cytometry: unravelling the immune system. *Nat Rev Immunol* **4:** 648–655

Pierre Baldi SB (2001) *Bioinformatics: The Machine Learning Approach*. London, UK: MIT press

Puisieux I, Even J, Pannetier C, Jotereau F, Favrot M, Kourilsky P (1994) Oligoclonality of tumor-infiltrating lymphocytes from human melanomas. *J Immunol* **153:** 2807–2818

Romero P, Dunbar PR, Valmori D, Pittet M, Ogg GS, Rimoldi D, Chen JL, Lienard D, Cerottini JC, Cerundolo V (1998) *Ex vivo* staining of metastatic lymph nodes by class I major histocompatibility complex tetramers reveals high numbers of antigen-experienced tumor-specific cytolytic T lymphocytes. *J Exp Med* **188:** 1641–1650

Segal E, Friedman N, Koller D, Regev A (2004) A module map showing conditional activity of expression modules in cancer. *Nat Genet* **36:** 1090–1098

Shiow LR, Rosen DB, Brdickova N, Xu Y, An J, Lanier LL, Cyster JG, Matloubian M (2006) CD69 acts downstream of interferon-alpha/beta to inhibit S1P1 and lymphocyte egress from lymphoid organs. *Nature* **440:** 540–544

Terabe M, Berzofsky JA (2004) Immunoregulatory T cells in tumor immunity. *Curr Opin Immunol* **16:** 157–162

Van't veer LJ, Dai H, van de Vijver V, He YD, Hart AA, Mao M, Peterse HL, van der KK, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415:** 530–536

Vose BM, Vanky F, Argov S, Klein E (1977) Natural cytotoxicity in man: activity of lymph node and tumor-infiltrating lymphocytes. *Eur J Immunol* **7:** 353–357

Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, Makrigiannakis A, Gray H, Schlienger K, Liebman MN, Rubin SC, Coukos G (2003) Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* **348:** 203–213