Supplemental Data

Foxp3 Transcription-Factor-Dependent and -Independent Regulation of the Regulatory T Cell Transcriptional Signature

Jonathan A. Hill, Markus Feuerer, Kaley Tash, Sokol Haxhinasto, Jasmine Perez, Rachel Melamed,1 Diane Mathis, and Christophe Benoist

Supplemental Experimental Procedures

Microarray Analysis

Sliding window analysis, "SignatureDepth" GenePattern tool (Fig.1 B, C)- This algorithm developed in S+ (code available upon request) aims to generate better estimates of the number of significant differences that contribute to a Gene Signature, by utilizing information from related determinations of the signature. We define here a Signature as the set of transcripts whose expression in two conditions (A vs B) is significantly affected by a perturbant (e.g. a genetic mutation, a drug treatment) or differs between two cellular differentiation states. The algorithm utilizes two or more determinations of the A/B signature (each averaged from several replicates), thus A/B, A'/B', A"/B", etc (in this instance, A=splenic Treg cells, B=splenic Tconv cells).

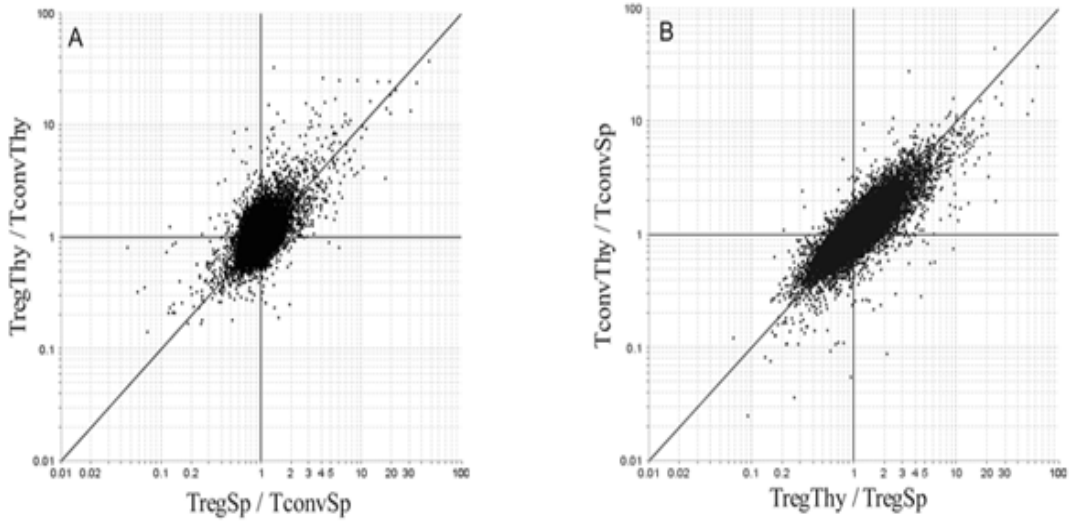
Probes are first ranked according to FoldChange in one of the A/B pairs (hereafter "ranker" pair). A simple graphic analysis then plots, in a series of consecutive sliding windows (window width =50 probes), the mean FoldChange for these ranked genes in the other A/B pairs (hereafter rankee pairs), which gives a first visual impression of the depth of the signature (Fig. 1B). For a determination of significance, a "bias" is then calculated for each of these windows for all rankee pairs; the bias is defined as the departure from a random distribution of FoldChanges (25 probes yielding a FC >1, 25 probes with a FC<1). The probability of each of these observed biases occurring by chance is then determined by generating randomly sampled windows and counting the number of times a particular bias occurs in windows out of 10^4 to 10^6 samplings. These probabilities are then plotted for the whole gene set (Fig. 1C), together with a smoothed fit (computed with Friedman's super smoother). The intercept of this smoothed line with p=0.05 used to define the boundary FoldChange for over- or under-represented probes. An estimate of the number of affected genes is then generated by counting, in the rankee pairs, the number of genes with FC>1 minus those with FC<1 (since, at low FoldChange values, one expects a numbers of false-positives, which is approximated by the number of genes with FC<1; the opposite calculation is performed for under-expressed genes). These numbers are then displayed on the graph.

Foxp3 Index- The Foxp3 Index was calculated from the FoldChange of mean expression values; where wtFoldchange = wtActCD4TGF/wtActCD4, and sfFoldchange = sfActCD4TGF/sfActCD4.

Foxp3 Index = [(wtFoldchange – sfFoldchange) / (maximum Foldchange of either wt or sf)] * Sign of wtFoldchange

Preprocessed dataset- The dataset was preprocessed to remove genes with background level signals, by removing probes whose expression values did not exceed 20 arbitrary units in any of the conditions across the entire datagroup. A probe was kept if the expression value was above 50 arbitrary units in ≥ 3 of the 64 individual microarray experiments. This filtering was conducted using the "Preprocess Dataset" module from "Genepattern".

Supplemental Figure 1



(A) Fold-change, fold-change plot showing the expression value differences between Treg and Tconv cells from the thymus (y-axis), or the periphery (x-axis); (B) FcFc plot showing expression value differences between Treg cells from the thymus and periphery (y-axis), or Tconv cells from the thymus and periphery (x-axis).