Statistical Methods.

Figures 1A and 1B have identical structures but, of course, they differ in what was measured. In each case there were three comparisons. The least standard of them was comparing subjects, before and after treatment. This comparison was complicated because while 11 subjects were measured both before and after treatment, 15 subjects were measured only before; and 23 only after. We compared all subjects measured before with all subjects measured after. Thus, the specification is that all subjects before are independent and identically distributed. Likewise, all subjects measured after are assumed independent and identically distributed. However, 11 subjects in each of the two groups were measured both times. Their measurements are assumed independent of those from the other 38 subjects, but their before and after treatment measurements are not independent. We denote the (population) mean value of each before measurement by μ_B and the corresponding variance as σ_B^2. Parameters for individual measurements taken after treatment are μ_A and σ_A^2. A principal "null hypothesis" is that $\mu_A = \mu_B$. The respective variances are examples of "nuisance parameters"

Sampling distributions of ordinary two-sample t-tests are predicated not only on the assumption that all data are normally distributed, but additionally on the assumption that the two variances are equal. There is the additional assumption regarding normality of both sets of distributions. For data in Figures 1A and 1B there are inferences on comparisons of (26) before and (34) after treatment measurements with (the 34) control measurements. These comparisons involve ordinary two-sample t test statistics. Always, control measurements are assumed independent of each other, and of measurements taken before or after treatment. They are assumed to have common mean value μ_C and common variance σ_C^2. Thus, when before measurements are compared with control, the null hypothesis is that $\mu_B = \mu_C$. Obviously analogous statements can be made regarding a comparison of measurements taken after treatment and controls.

The assumption of normality matters little to any of the three cited comparisons owing to the "central limit effect" (Bickel and Doksum, Section A.2.7, page 464). This is

because there are $11 + 15 = 26$ measurements before, $11 + 23 = 34$ after, and 34 control measurements. Each comparison involves a difference in sample mean values divided by an estimate of the standard deviation of that difference. Our estimate of μ_B is the mean of all measurements taken before treatment, and is denoted by $\bar{x}_B$; it is computed from all available measurements, 26 for data cited. Our estimate of σ_B^2 is the estimated sample variance of these 26 measurements, write s_B^2. Estimates $\bar{x}_A$ and s_A^2, respectively of μ_A and σ_A^2 are defined analogously from the 34 measurements taken after treatment. The correlation between before and after measurements on the same individual is denoted by ρ_{AB}. In our application regarding data in Figures 1A and 1B, the estimate computed from data, r_{AB}, is based only 11 subjects.

Discussion thus far and simple algebra entail that

$$\text{Var } (\bar{x}_A - \bar{x}_B) = \frac{\sigma_B^2}{26} + \frac{\sigma_A^2}{34} - 2(\frac{11}{(26)(34)})\rho_{AB}\ \sigma_A\sigma_B.$$

It follows from Slutkys theorem (Bickel and Doksum, Section A.2.7, page 460) that when the null hypothesis $\mu_A = \mu_B$ is true, then

$$z = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_B^2}{26} + \frac{s_A^2}{34} - 2(\frac{11}{(26)(34)}\ r_{A,B}\ s_A s_B)}}$$

has, at least approximately, a standard normal distribution. By computing the probability that a standard normal Z exceeds the observed computed z in absolute value, one has an approximate attained significance value (p-value) for a two-sided test of the null hypothesis that $\mu_A = \mu_B$.

Readers note that regarding data in Figures 1A and 1B, we are testing three separate hypotheses regarding pair-wise equality of the parameters μ_A, μ_B, and μ_C. We should not employ a standard one-way analysis of variance, especially but not only because the three sets of observations that comprise $\bar{x}_A, \bar{x}_B$ and $\bar{x}_C$ are not independent. However, an overall p-value for the combined tests $\mu_A = \mu_B;\ \mu_A = \mu_C;\ \mu_B = \mu_C$ can be accomplished

by applying the usual Bonferroni inequality (Bickel and Doksum, pages 288 and 439), and simply adding the p-values. While testing $\mu_A = \mu_B$ can be accomplished by using the statistic z we have derived, testing the null hypothesis that μ_B, alternatively μ_A equals μ_C can be accomplished by a two-sample t-test. Assumptions underlying the ordinary test may fail because of inequality of variances, in which case the Behrens-Fisher test, which is analogous to z (see http://sekhon.berkeley.edu/stats/html/t.test.html) can be used. One might also apply the two-sample Wilcoxon statistic for testing $\mu_B = \mu_C$ or $\mu_A = \mu_C$ (see http://stat.ethz.ch/R-manual/R-patched/library/stats/html/wilcox.test.html). Neither is entirely appropriate since in both cases, the sampling distributions are not identical up to a possible shift of one of them. However, readers please note that all three tests of both null hypotheses have exceedingly small p-values. So it is immaterial which test is employed, and by any reasonably criterion we "reject" not only the two just cited, but even all three null hypotheses.

When it comes to data other than those of Figures 1A and 1B, issues described already are both simpler and more complex. They are simpler because in testing $\mu_A = \mu_B$ all six subjects that figure in computing $\bar{x}_A$ and $\bar{x}_B$ were measured both before and after treatment. Therefore, z reduces to a simple one-sample t statistic. However, because it has only five degrees of freedom, for the null central t-distribution to apply reasonably accurately, the six differences, after minus before by subject, should appear to have normal distributions. We assessed normality by inspecting so-called q-q plots (http://sekhon.berkeley.edu/stats/html/qqnorm.html) and also by application of the Wilk-Shapiro statistic (http://sekhon.berkeley.edu/stats/html/shapiro.test.html). When distributions appear normal to the extent we can infer that, then one-sample t might be trusted. When they are not, then we prefer the one-sample Wilcoxon statistic (URL already given), though any test of its assumption of symmetry about its center is not testable from so little data. Entirely analogous comments apply to testing $\mu_B = \mu_C$ and $\mu_A = \mu_C$ from cited data other than those of Figures 1A and 1B. Here two-sample t has 10 degrees of freedom,

so we are less concerned about normality of data than when we test $\mu_A = \mu_C$. Further, when distributions appear non-normal, we have both Behrens-Fisher-Welch t and the two sample Wilcoxon statistic for testing. Testing normality is, obviously, suspect with only six observations, and we are somewhat perplexed which inferences to draw when competitive p-values give somewhat different signals.

Reference

Bickel, P.J. and Doksum, K.A. (1977), *MATHEMATICAL STATISTICS: Basic Ideas and Selected Topics*, Holden-Day, Inc., San Francisco.

Supplemental Table 1. Demographic and clinical characteristics of GCA patients and control donors

Items	Con	Untreated GCA	Treated GCA
No. of subjects	34	26	34
Sex, female/male	26/8	19/7	28/6
Age, mean+/-SD years	72+8	77+7	73+8
Ethnicity, %			
Cauc	97	100	97
African American	3	0	3
Headache, %		75	
Jaw claudication, %		50	
Scalp tenderness, %		50	
Ischemic optic neuropathy, %		17	
Fever, %		25	
Weight loss, %		33	
Fatigue, %		75	
Anemia, %		50	
Polymyalgia rheumatica, %		42	
ESR+/-SD		81+20	31+7
CRP+/-SD		80.7+24	15.7+5
Platelets+/-SD		350+143	293+75

Supplemental Table 2. Reagents and antibodies utilized in cell culture and staining

Reagents/Antibodies	Company	Usage	Concentration or Dilution
LPS	Sigma-Aldrich, St. Louis, MO	Cell culture	1μg/ml
Goat anti-human IL-1β	R&D Systems, Minneapolis, MN	Cell culture	2μg/ml
Goat anti-human IL-6	R&D Systems	Cell culture	4μg/ml
Goat anti-human IL-23p19	R&D Systems	Cell culture	6μg/ml
Purified goat IgG control	R&D Systems	Cell culture	6μg/ml
Mouse anti-human IL-4	R&D Systems	Cell culture	2μg/ml
Mouse anti-human IFN-γ	R&D Systems	Cell culture	2μg/ml
Anti-human CD3 (OKT)	Ortho Biotech, Raritan, NJ	Cell culture	1μg/ml
Anti-human CD45RA microbeads	Miltenyi Biotec, Auburn, CA	Purification	See company instruction
Anti-human CD45RO microbeads	Miltenyi Biotec	Purification	See company instruction
phorbol 12-myristate 13-acetate (PMA)	Sigma-Aldrich	Cell culture	50μg/ml
Ionomycin	Sigma-Aldrich	Cell culture	1μg/ml
Brefeldin A	eBioscience, San Diego, CA	Cell culture	3μg/ml
FITC-conjugated mouse anti-human CD3	BD, San Diego, CA	Cell staining	1:100
PE-conjugated mouse anti-human CD3	BD	Cell staining	1:100
PerCP-conjugated mouse anti-human CD4	BD	Cell staining	1:100
FITC-conjugated mouse anti-human IFN-γ	BD, Pharmingen, San Jose, CA	Cell staining	1:100
APC-conjugated mouse anti-human IL-17	eBioscience, San Diego, CA	Cell staining	1:50
PE-conjugated mouse anti-human Foxp3	BD, Pharmingen	Cell staining	1:25
Mouse anti-human CD3	Dako, Carpinteria, CA	Tissue staining	1:100
Mouse anti-human CD4	Neuromics, Edina, MN	Tissue staining	1:100
Mouse anti-human CD8	Dako	Tissue staining	1:100
Mouse anti-human CD14	Santa Cruz Biotechnology, Santa Cruz, CA	Tissue staining	1:100
Mouse anti-human CD15	Dako	Tissue staining	1:100
Purified mouse IgG control	Innovative Research, Southfield, MI	Tissue staining	1:100
Rabbit anti-human IL-17	Santa Cruz Biotechnology	Tissue staining	1:100
Rabbit anti-human IFN-γ	Santa Cruz Biotechnology	Tissue staining	1:100
Purified Rabbit IgG isotype control	IMGENEX, San Diego, CA	Tissue staining	1:100

Supplemental Table 3. Primer pairs utilized in quantitative RT-PCR assays*

Gene	Sense (5'-3')	Anti-sense (5'-3')
IL-17	AACCGATCCACCTCACCTTGGAAT	TTCATGTGGTAGTCCACGTTCCCA
IFN-γ	ACTAGGCAGCCAACCTAAGCAAGA	CATCAGGGTCACCTGACACATTCA
Foxp3	TTCAAGTTCCACAACATGCGACCC	GCACAAAGCACTTGTGCAGACTCA
IL-1β	AAGTACCTGAGCTCGCCAGTGAAA	TTGCTGTAGTGGTGGTCGGAGATT
IL-6	AGCCACTCACCTCTTCAGAACGAA	AGTGCCTCTTTGCTGCTTTCACAC
IL-23p19	ACTCAGCAGATTCCAAGCCTCAGT	TGGAGATCTGAGTGCCATCCTTGA
IL-12p35	TAACTAATGGGAGTTGCCTGGCCT	AGGGCCTGCATCAGCTCATCAATA
IL-12p40	TCATCAAACCTGACCCACCCAAGA	TTTCTCTCTTGCTCTTGCCCTGGA
MMP-9	TACCACCTCGAACTTTGACAGCGA	GCCATTCACGTCGTCCTTATGCAA
β-actin	ACCAACTGGGACGACATGGAGAAA	TAGCACAGCCTGGATAGCAACGTA

***** cDNA (0.5μl) was mixed with 9μl of double distilled water, 10 μl of SYBR Green Master Mix (2× PCR buffer, 5 mM $MgCl_2$, 0.4 mM each dNTP, 0.05% bovine serum albumin and 1:10000 SYBR Green) (Invitrogen), 0.2μl of each primer (0.1μM) and 0.1 μl of Platinum *Taq* (5 U/μl) (Invitrogen). Amplifications were performed in a Mx3000 PCR machine (Stratagene) under the following cycling conditions: denaturation at 95°C for 5 min, followed by 40 cycles at 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s. For each sample, PCR reactions were performed in triplicate. The level of gene expression was determined by interpolation with a standard curve.

Supplemental Table 4. Summary of Statistical Analysis

	Experimental set	Groups	Endpoint	Analysis performed	Comments
Analysis 1	Fig.1 A-C Effect of glucocorticoids on circulating Th1 and Th17 cells	Controls [a] n=34 GCA Patients untreated [b] n=26 treated [b] n=34	Frequencies of cytokine producing T cells	Customized t-like statistics (see supplement Statistical Methods)	[b] 11 patients were enrolled when untreated and analyzed again on therapy. 15 patients were tested only as untreated and 23 patients were only analyzed while on therapy.
Analysis 2	Fig.1 G-H; Fig. 3 A-C Effect of glucocorticoids on plasma cytokines (IL-17, IFN-γ, IL-1β, IL-6, IL-12)	Controls [a] n=6 GCA Patients untreated [b] n=6 treated [b] n=6	Plasma levels of cytokines	One-sample t, two-sample t, Behrens-Fisher-Welch, or one-sample Wilcoxon as appropriate	[b] patients were enrolled before treatment and analyzed again on therapy
Analysis 3	Fig.2 A-B; Fig.4 A-E Effect of glucocorticoids on tissue cytokines (IL-17, IFN-g, IL-1b, IL-6, IL-12p40, IL-12p35, IL-23p19)	Controls [a] n=8 GCA Patients untreated [b] n=8 treated [b] n=8	Transcript levels of tissue cytokines	One-sample t, two-sample t, Behrens-Fisher-Welch, or one-sample Wilcoxon as appropriate	[b] samples from untreated and treated patients were paired
Analysis 4	Fig.3 D-H Effect of glucocorticoids on cytokine production in peripheral monocytes (IL-1b, Il-6, IL-12p35, IL-12p40, IL-23p19)	Controls [a] n=8 GCA Patients untreated [b] n=8 treated [b] n=8	Transcript levels of cytokines	One-sample t, two-sample t, Behrens-Fisher-Welch, or one-sample Wilcoxon as appropriate	[b] samples from the untreated and treated patients were paired
Analysis 5	Fig.5 Effect of glucocorticoids on tissue cytokines in human artery-SCID chimeras.	Control tissues [c] n=9 Tissues untreated [d] n=9 Tissues treated [d] n=9	Transcript levels of tissue cytokines	One-sample t, two-sample t, Behrens-Fisher-Welch, or one-sample Wilcoxon as appropriate	[c, d] artery tissues were engrafted into mice. Arterial wall inflammation was induced or not (control). Arteries with inflammation were treated with steroids or left untreated.
Analysis 6	Suppl. Fig.3 In vitro testing of anti-cytokine antibodies during the induction of Th17 cells	Isotype control [c] n=3 Anti-IL-1b [c] n=3 Anti-IL-6 [c] n=3 Anti-IL-23 [c] n=3	Frequency of TH17 cells	Friedman rank sum test	[c] cells were isolated from untreated GCA patients and co-cultured with either isotype control or anti-IL-

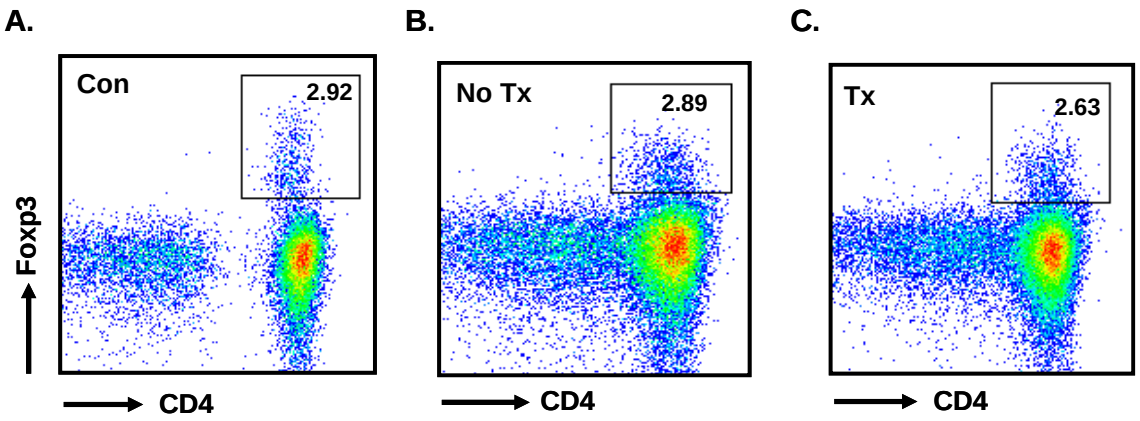
					1b, or anti-IL-6 or anti-IL-23 antibody in tissue culture. Frequencies of induced Th17 cells were measured.
Analysis 7	Table 1 Effect of glucocorticoids on the frequencies of Th17 cells	GCA Patients untreated [b] n=11 treated [b] n=11		One-sample t	[b] 11 patients were enrolled while untreated and were analyzed again while on therapy.

[a] Healthy volunteers served as controls. Age distributions in controls, untreated and treated patients were not significantly different.

Supplemental Table 5. The influence of glucocorticoid therapy on the frequency of circulating Th17 cells in individual patients

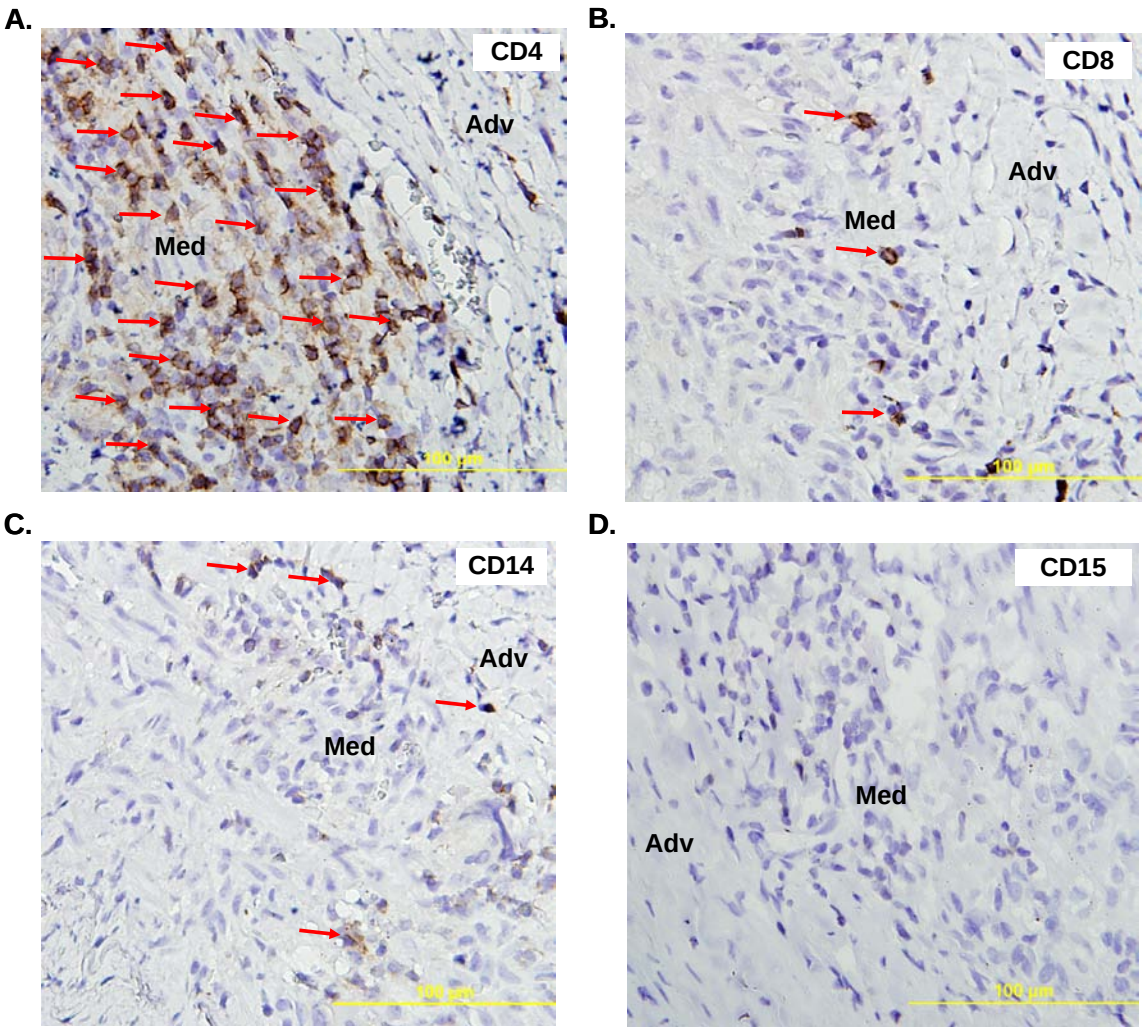
	Untreated	Treated		
Patient	Th17 (%)	Treatment duration (months)	Current prednisone dose (mg/day)	Th17 (%)
No. 1	2.1	2.5	30	0.29
No. 2	2.2	3	30	0.28
No. 3	1.8	3	30	0.65
No. 4	1.6	3.5	20	0.4
No. 5	2.5	3.5	15	0.45
No. 6	1.9	5	9	0.39
No. 7	4.3	6	15	0.53
No. 8	2.9	6.5	8	0.49
No. 9	3.3	7	20	0.42
No. 10	2.2	7	7	0.61
No. 11	1.8	9	15	0.18

Supplemental Figure 1



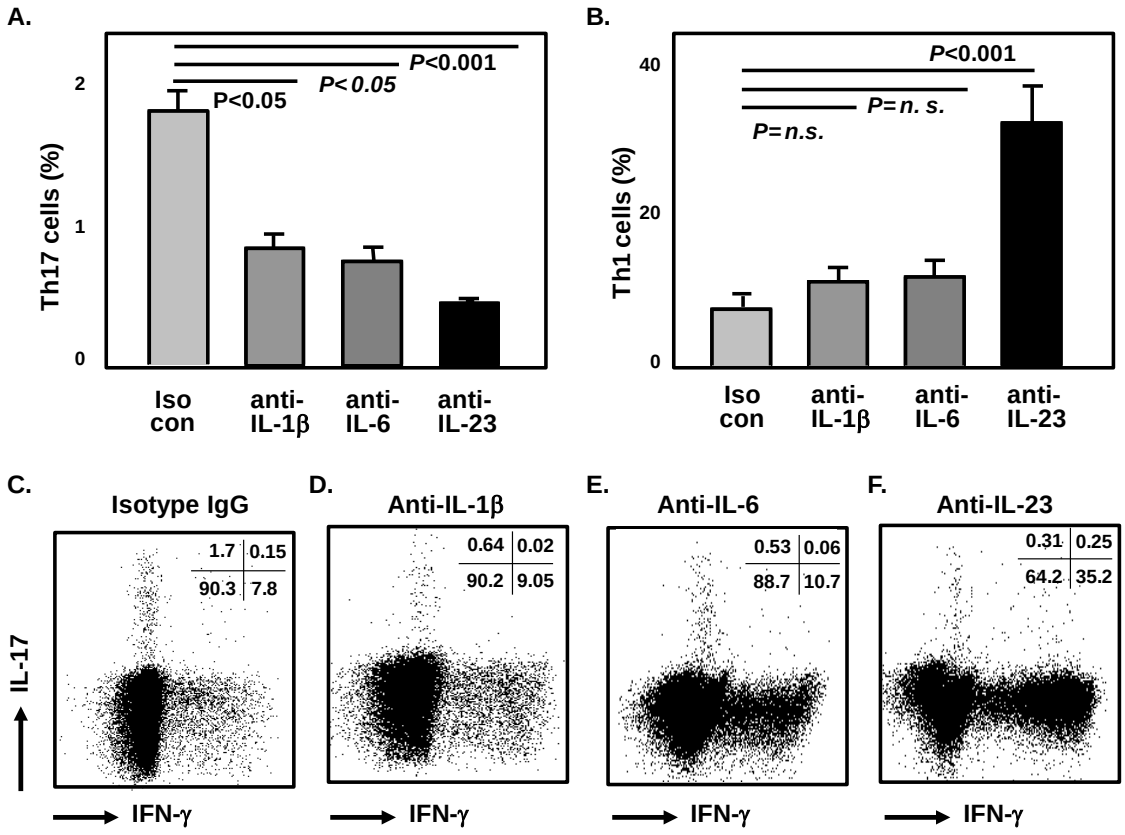
Supplemental Figure 1. ***Foxp3^{+} CD4 regulatory T-cells in GCA.*** PBMC isolated from age-matched control donors (**A**), untreated **(B)** and GC-treated GCA patient **(C)** as described in Figure 1 were stained with FITC anti-CD3, PerCP anti-CD4 and PE anti-Foxp3 antibodies and analyzed by flow cytometry. Representative cytometric dot plots are shown. Frequencies of Foxp3^{+} regulatory CD4 T cells amongst CD3 T cells were similar in controls, untreated and treated patients.

Supplemental Figure 2



Supplemental Figure 2. ***Representation of different cell types in the granulomatous lesions of GCA.*** Paraffin-embedded tissue sections of temporal arteries from GCA patients were de-waxed, and stained with mouse anti-human CD4 **(A)**, mouse anti-human CD8 **(B)**, mouse anti-human CD14 (monocyte marker) **(C),** anti-human CD15 (neutrophil marker) **(D)** antibodies or purified mouse IgG control (not shown). The images presented are representative of five different biopsy samples. Positive cells are marked by red arrows. Similar to anti-human CD15 staining (D), mouse IgG control showed complete negative. Magnification 600× in (A), (B), (C) and (D).

Supplemental Figure 3



Supplemental Figure 3. ***IL-1β, IL-6, and IL-23 promote Th17 cell differentiation in GCA.*** CD45RO$^-$ CD4 naïve T cells and CD14$^+$ monocytes were purified from PBMC of GCA patients. Naïve T cells were co-cultured with monocytes at a 2:1 ratio in the presence of anti-CD3 antibodies and LPS for 4 days, supplemented with goat anti-human IL-1β, anti-human IL-6**,** or anti-human IL-23 neutralizing antibodies, or purified goat IgG isotype control antibody. Cells were stained with PE-conjugated anti-CD3, PerCP-conjugated anti-CD4, FITC-conjugated anti-IFN-γ, and APC-conjugated anti-IL-17 antibodies, and analyzed by flow cytometry as described in Figure 1. Frequencies of Th17 **(A)** and Th1 cells **(B)** were analyzed by flow cytometry, and are presented from three independent experiments. **(C-F)** Representatives of cytometric dot

plots from the co-cultured supplemented with goat anti-human IL-1β (D), anti-human IL-6 (E), or anti-human IL-23 (F) neutralizing antibodies, or purified goat IgG isotype control antibody (C).