

SCIENTIFIC REPORTS



OPEN

High-throughput RNA sequencing reveals distinct gene signatures in active IgG4-related disease

Brandon W. Higgs¹, Yanying Liu², Jianping Guo², Yinong Sebastian¹, Chris Morehouse¹, Wei Zhu¹, Limin Ren², Mengru Liu², Yan Du², Guangyan Yu³, Lingli Dong⁴, Hong Hua⁵, Pan Wei⁵, Yi Wang⁶, Zhengang Wang⁷, Yihong Yao¹ & Zhan-Guo Li²

We aimed to characterize the molecular differences and effects from prednisone treatment among IgG4-related disease with salivary gland lesions (RD-SG), without SG lesions (RD-nonSG), and IgG4-related retroperitoneal fibrosis (RF). RNA sequencing was conducted on blood from 25 RD-SG, 11 RD-nonSG, 3 RF and 10 control subjects. Among these, 8 RD-nonSG and 12 RD-SG patients were subjected to treatment with prednisone and/or glucocorticoid-sparing agents. Six RD patients had a longitudinal time point. The mRNA levels of *IgG4* and *IgE*, genes specific for Th2 cells, eosinophils, and neutrophils were over-expressed in RD-SG and RD-nonSG. A B-cell signature was suppressed in patients group versus controls, while Th1, Th2, Treg, and eosinophil gene signatures were increased in patients without treatment. Interestingly, Tfh genes and B cell signature were decreased at flare disease state. Prednisone treatment led to increased neutrophil, but decreased Treg signatures. Serum IgG4 levels correlated with the eosinophil and neutrophil gene signatures in RD-SG patients, and with a B cell signature in only RD-nonSG patients. *IgG4*, *IgE*, and cell-specific signatures are regulated in patients, suggesting the imbalance of immune and inflammatory cells in IgG4-related disease. Prednisone treatment selectively modulates Treg, eosinophil, and neutrophil signatures.

IgG4-related disease (IgG4-RD) is a systemic disorder involving a spectrum of multiple indications, distinguished by *often* elevated levels of serum IgG4, infiltration of IgG4⁺ plasma cells into target tissues, and diffuse swelling, mass formation, or fibrosis of affected organs¹. This disease affects men approximately two-fold more often than women and age at diagnosis ranges from 50 to 70 years². Most patients do respond to steroids initially, although relapse is observed in up to 47% cases^{3–5}. Various histopathological features are shared among different IgG4-RD indications, which challenge diagnosis, although certain syndromes have organ-specific involvement⁶. Some examples include: Mikulicz's disease affecting the salivary and lacrimal glands, autoimmune pancreatitis affecting the pancreas, Riedel's thyroiditis affecting the thyroid, and Morbus Ormond or retroperitoneal fibrosis (RF), affecting tissue in the retroperitoneum, to name a few².

Beyond the evidence of certain genetic risk factors^{7–11}, IgG4-RD is mechanistically thought to be activated by the innate response to pathogens that mimic self-antigen, leading to an autoimmune response^{2,6}. Type 1 helper T cells (Th1) are thought to support innate immune response to infection, which then shifts to type 2 helper T cells (Th2) involvement with increases in expression of IL-4, IL-5, and IL-13 mRNA and protein in both the affected tissue and peripheral compartments^{12–15}. Th2 adaptive response can affect Th1 response, thus this Th1/Th2 balance is important in regulation¹. Regulatory T cells (Tregs) are also activated, with accumulation of CD4⁺CD25⁺ T cell infiltrates and abundance of IL-10, FOXP3, and TGF- β ^{12,16,17}. The increase of these cytokines promotes

¹MedImmune, Gaithersburg, MD, 20878, USA. ²Department of Rheumatology and Immunology, Peking University People's Hospital, Beijing, 100044, China. ³Department of Oral and Maxillofacial Surgery, Peking University School of Stomatology, Beijing, 100081, China. ⁴Department of Immunology and Rheumatology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430000, China. ⁵Department of Oral Medicine, Peking University School of Stomatology, Beijing, 100081, China. ⁶Department of Radiology, Peking University People's Hospital, Beijing, 100044, China. ⁷Department of Rheumatology and Immunology, Beijing Tongren Hospital, Beijing, 1000730, China. Brandon W. Higgs, Yanying Liu and Jianping Guo contributed equally to this work. Correspondence and requests for materials should be addressed to Y.Y. (email: yaoy@medimmune.com) or Z.-G.L. (email: zgli99@aliyun.com)

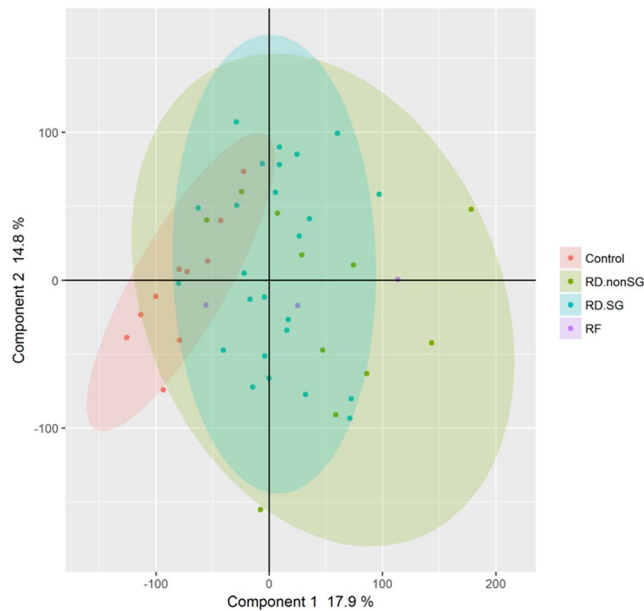


Figure 1. Principal components analysis plot of RD-SG, RD-nonSG, RF, and control subjects using the whole transcriptome.

eosinophilia in the serum or tissue, high levels of IgG4-producing plasma cells, elevated production of IgE, and fibrosis, with inflammatory cell infiltrates ultimately causing organ damage⁶.

Recently, studies have utilized transcript profiling in labial salivary glands (LSGs) to identify distinguishing molecular features between IgG4-RD and Sjögren's syndrome (SS), a disease with common phenotypic elements^{18–20}. Among other findings, active involvement of Th2- (*IL-4*, *IL-5*, and *IL-21*), T follicular helper cell (Tfh)- (*BCL-6* and *CXCR5*) and T-reg- (*IL-10*, *FOXP3*, *CCL18*, and *TGF- β 1*) related transcripts in patients with IgG4-RD was observed. These data showed how elevated levels of such cytokines and chemokines can induce IgG4 plasma cell infiltration, high IgG4 levels in the periphery, and impact tissue fibrosis in the LSG of IgG4-RD patients¹⁹. However, no studies to date have assessed the differences in molecular pathways or cell populations among IgG4-related disease with salivary gland lesions (RD-SG), without SG lesions (RD-nonSG), and IgG4-related retroperitoneal fibrosis (RF), in the peripheral blood, as well as the effects of corticosteroids on these signaling pathways.

In this study, we used whole transcriptomic sequencing to identify and distinguish both cell and pathway-associated activation in the blood of healthy subjects or those with RD-SG, RD-nonSG, or RF. A large cohort of patients was transcript profiled at a relative baseline time point, with two patients providing additional post baseline flare specimens. To better understand the possible mechanism(s) implicated in the treatments of IgG4-RD, we evaluated the effects of prednisone on the molecular pathways most relevant to disease activity. Additionally, cell-specific gene signatures linking the B and T cell axes were assessed to elucidate cellular involvement, as well as the correlation with *IgG4* mRNA levels across the three diseases.

Results

Transcriptome profiles in patients with RD-SG, RD-nonSG, or RF and healthy controls using principal components analysis. Principal components analysis (PCA) was used to elucidate the whole transcriptome profile among the three diseases in relation to healthy subjects (Fig. 1). Though the plot displayed an overlap in disease and control cohorts, there was an apparent difference between the control subjects and disease subjects. Specifically, along the x-axis (principal component 1), controls (red) were the leftmost cohort, followed by the other disease groups. More relevant was the smaller within-disease variability that was apparent in the control and RD-SG (blue) compared to the RD-nonSG cohort (green). The RF cohort (purple) was very small ($n = 3$), thus the distribution of these points were difficult to interpret.

***IgG4* and *IgE* are the most over-expressed transcripts in RD-SG and RD-nonSG patients, and suppressed by prednisone in RD-SG patients.** *IgG4* and *IgE* were identified as two of the most over-expressed transcripts in both RD-nonSG and RD-SG compared to the control cohort (Supplementary Table 1). These two cohorts were stratified by patients who were currently being treated with prednisone. All four patient cohorts had significantly higher mRNA expression of *IgG4/Total IgGs* and *IgE* ($p \leq 0.001$ for all cohorts; Fig. 2A,B). RD-SG patients treated with prednisone had significant suppression of *IgG4/Total IgGs* ($p = 0.01$) and *IgE* ($p = 0.003$) mRNAs compared to those not treated, while RF patients showed difference in *IgE* from controls, though the sample size was small ($p = 0.04$). *IgG4/Total IgGs* and *IgE* mRNAs were highly correlated across the diseases (Fig. 2C; $\rho = 0.66$, $p < 9.78 \times 10^{-6}$).

A linear model was constructed to identify genes across the transcriptome most correlated with *IgG4* mRNA levels. This approach was used to 1) adjust for transcripts modulated by prednisone treatment, and 2) distinguish

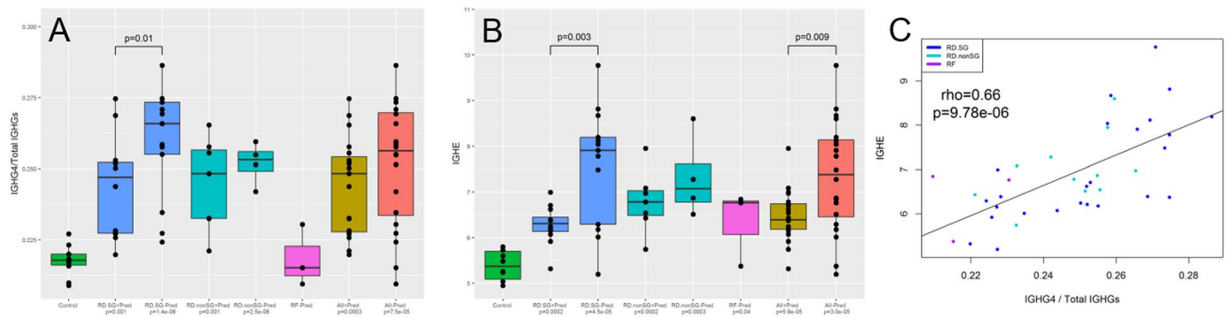


Figure 2. Distribution of *IgG4* and *IgE* mRNAs. **(A)** Expression of *IgG4* scaled by all *IgG1*, *IgG2*, and *IgG3* transcripts and **(B)** *IgE* across control subjects, RD-SG patients on prednisone treatment, RD-SG patients not on prednisone treatment, RD-nonSG patients on prednisone treatment, RD-nonSG patients not on prednisone treatment, RF patients, all patients on prednisone treatment, and all patients not on prednisone treatment. **(C)** Correlation between *IgE* and *IgG4*/Total *IgG* mRNAs for all three diseases. P-values under each disease group indicate comparisons to control and are adjusted by age. Pred + = currently treated with prednisone; Pred - = not currently treated with prednisone.

transcripts unique to one of RD-SG, RD-nonSG, or RF cohorts. Among the top 50 positively and negatively correlated transcripts with *IgG4* ($p < 0.01$), 39/100 were associated with RD-SG ($p < 0.01$), 28/100 with RD-nonSG ($p < 0.01$), and 3 with RF ($p < 0.01$), with 23 being shared between RD-SG and RD-nonSG and 2 associated with RF shared with RD-SG and RD-nonSG cohorts (*CCL23* was unique to RF; Supplementary Table 2). This indicates that similar transcripts correlate with *IgG4* in all three diseases. Among these top 50 most positively correlated genes with *IgG4*, regardless of disease (and adjusting for prednisone treatment), the most activated biological categories were immunoglobulin (*IGHG1*, *IGHG3*, *IGHV1-69*, *IGHV3-30*, *IGKV3-20*, *IGKV3D-15*, *IGLV2-11*, and *IGLV3-21*), followed by nuclear division/mitosis/replication (*CDK1*, *CDC20*, *CDC6*, *DLGAP5*, *MCM10*, *RRM2*, *KIF4A*, and *TOP2A*).

Treg, Th2, eosinophil, and neutrophil gene signatures are over-expressed in RD-SG and RD-nonSG, with a B cell signature suppressed in all diseases. Various cell-specific gene signatures were used to evaluate cell population involvement in the diseases studied here. Interestingly, a Treg gene signature showed significant over-expression in RD-SG and RD-nonSG patients without treatment with prednisone compared to controls (Fig. 3A). Regarding the Th2 cytokine signature, while IL-13 gene showed significant over-expression in only RD-SG patients compared to controls regardless of prednisone treatment, IL-4 gene showed significant elevation in both RD-SG and RD-nonSG patients compared to controls (Fig. 3B and C). The B cell signature was significantly suppressed in the majority of patient cohorts (including RF) (Fig. 3D). The eosinophil and neutrophil gene signatures showed opposite effects between patients with or without prednisone treatment, respectively, in RD-SG. Specifically, compared to patients without prednisone treatment, the eosinophil gene signature was significantly suppressed (Fig. 3E, $p = 0.0001$ in RD-SG), whereas the neutrophil gene signature was significantly elevated in RD-SG patients treated with prednisone (Fig. 3F, $p = 0.01$). Interestingly, a plasma cell gene signature showed no changes in any of the diseases with or without treatment, compared to controls (data not shown).

A molecular characterization of cell-specific gene signatures in RD-SG, RD-nonSG, and RF patients and controls. The molecular characterization of the cell-specific gene signatures, i.e. B, Th1, Th2, Treg, Tfh, and eosinophil cells, were analyzed across case and control cohorts (Fig. 4). From the heatmap, the effect of prednisone on all T cell sub-populations is evident. In general, the gene signatures were down-regulated in all patients treated with prednisone. The pattern seemed more apparent in T cell sub-populations of Th1, Treg, and Tfh in RD-SG patients. A similar pattern was seen in the eosinophil gene signature in both RD-SG and RD-nonSG patients. For the B cell signature, most genes were suppressed in RD-SG and RD-nonSG patients regardless of treatment, compared to healthy controls, though a few patients with active disease showed elevated expression across all genes (red vertical stripes in the B cell signature).

***IgG4* and *IgE* mRNAs, T, B, and eosinophil cell-specific genes/gene signatures differed among RD-SG/RD-nonSG patients with flare or stable status.** In addition to a baseline time point, blood was procured at a second time point from six RD-SG/RD-nonSG patients, two of whom experienced a flare. Though the exact time differences between the relative baseline and post baseline visit were not identical for each patient, the general molecular patterns were consistent for the two flare patients and differed for the four stable patients at the second visit (Fig. 5). For each RD-SG/RD-nonSG patient, there was induction at the flare time point in Th1, Th2, Treg, and eosinophil gene signatures. Similar induction at the flare time point was observed in *IgG4* and *IgE* mRNA levels (Fig. 5A–F). In contrast, the B cell signature, as well as two genes associated with Tfh cells (*BCL6* and *CXCR5*), all showed suppressed profiles at the flare time point (Fig. 5G–I). The control cohort was provided in each plot to indicate relative similarity of expression levels to a normal healthy population.

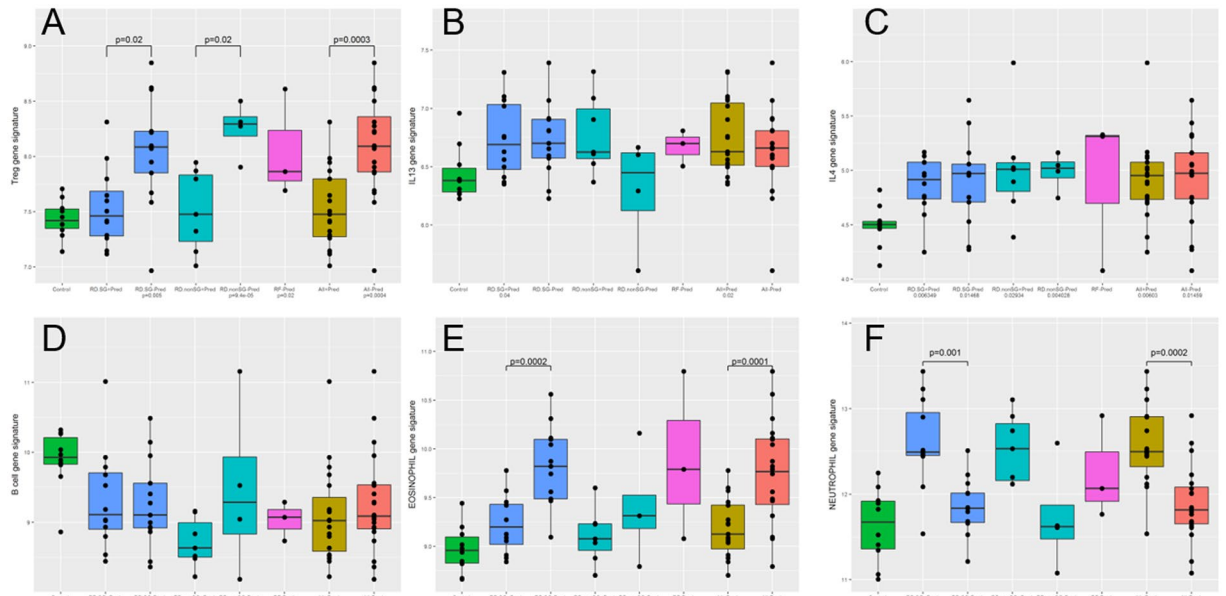


Figure 3. Distribution of cell-specific gene signatures. Expression of gene signatures of (A) Treg, (B) IL-13, (C) IL-4, (D) B cell, (E) eosinophil, and (F) neutrophil across control subjects, RD-SG patients on prednisone, RD-SG patients not on prednisone, RD-nonSG patients on prednisone, RD-nonSG patients not on prednisone, RF patients, all patients on prednisone combined, and all patients not on prednisone. P-values under each disease group indicate comparisons to control and are adjusted by age. Pred+ = currently treated with prednisone; Pred- = not currently treated with prednisone.

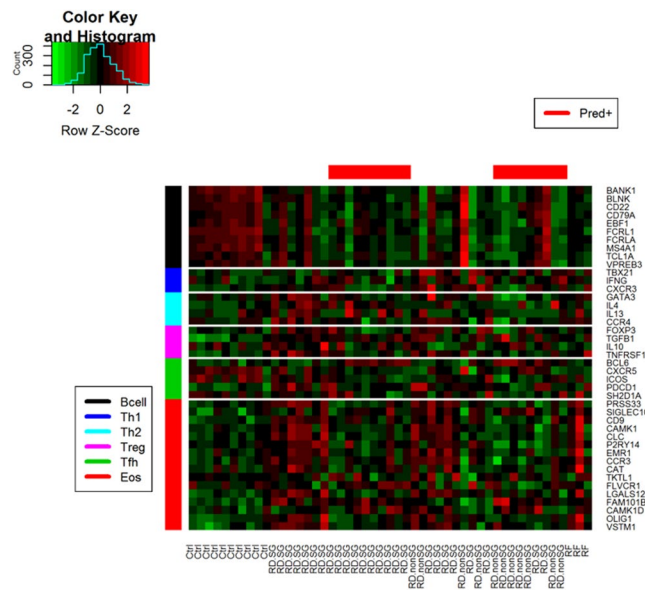


Figure 4. Heatmap of cell-specific gene signatures across the control, RD-SG, RD-nonSG, and RF cohorts. Ctrl = control subjects; Pred+ = currently treated with prednisone.

IgG4 serum levels correlate with cell signatures. Serum levels were correlated with cytokine and cell signatures described previously in RD-SG and RD-nonSG patients. IgG4 correlated with eosinophil and neutrophil gene signatures ($\rho = 0.44$, $p = 0.03$ and $\rho = -0.48$, $p = 0.02$, respectively) in RD-SG patients. In RD-nonSG, IgG4 levels correlated only with a B cell signature ($\rho = 0.74$, $p = 0.01$) (Fig. 6). No other associations between serum levels of IgG4 and gene signatures were observed.

Discussion

We used RNA sequencing to molecularly profile a large cohort of RD-SG, RD-nonSG, and RF patients. We showed that *IgG4* and *IgE* are among the most expressed transcripts in the blood of RD-SG or RD-nonSG patients,

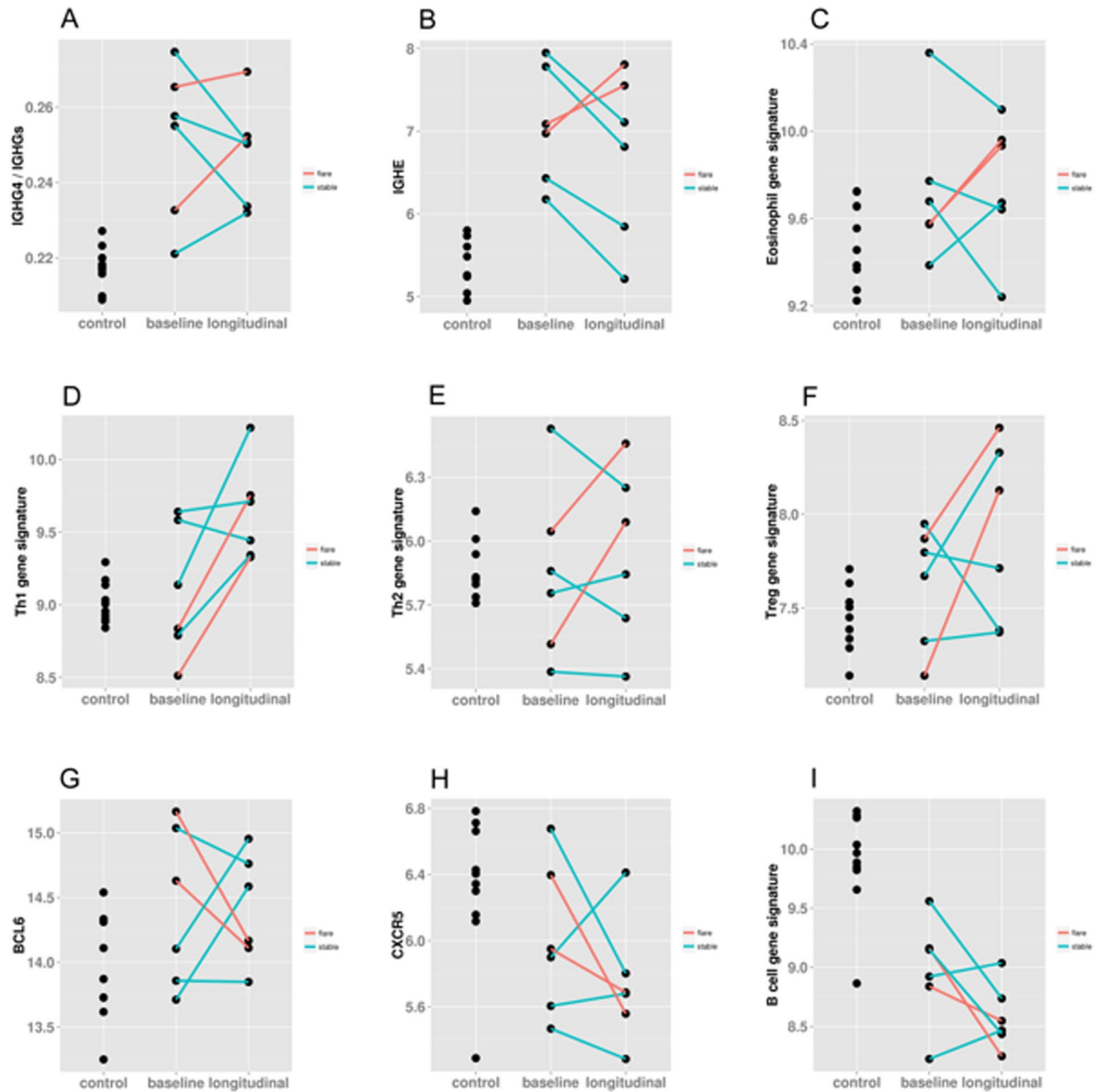


Figure 5. Longitudinal gene or gene signature profiles of two patients at relative baseline and post relative baseline visits when a flare was observed. All plots include the control cohort and all the 6 patients for (A) *IgG4*/Total *IgG*, (B) *IgE*, (C) eosinophil gene signature, (D) Th1 gene signature, (E) Th2 gene signature, (F) Treg gene signature, (G) *BCL6*, (H) *CXCR5*, and (I) B cell gene signature. Pink = patient experienced flare at visit 2; Blue = patient had stable disease at visit 2.

though not RF compared to controls, and both genes are highly correlated with each other. We also demonstrate that prednisone suppresses the levels of these genes in the blood in RD-SG, but not RD-nonSG patients. Reduction in serum *IgG4* protein following steroid therapy has been observed previously in various studies as a result of immune suppression; glucocorticoid treatment is a well-known regimen to help attain remission in MD patients^{21,22}. Among the 25 RD-SG patients, 17 of them in this study only have sialadenitis and dacryoadenitis involvement, which may explain the homogeneity of dramatic reduction in both *IgG4* and *IgE* levels in the blood from prednisone treatment. The higher intra-cohort variability observed in RD-nonSG compared to RD-SG patients in PCA across the whole transcriptome may also support this finding. Th2 gene signatures were generally increased in both RD-SG and RD-nonSG patients, and Treg gene signature was significantly reduced in patients treated with prednisone. Activation of Th2 cytokines and blood eosinophilia in RD patients has been previously reported, suggesting an allergy response mechanism in *IgG4*-RD, and eosinophilia is often treated with corticosteroids to promote cell death and clearance^{23–25}. The increased neutrophil signature in RD-SG patients treated with prednisone may be explained by the well-known phenomena of glucocorticoid-induced granulocytosis, where leukocytes have increased release from the bone marrow, and reduced migration out of the circulation²⁶.

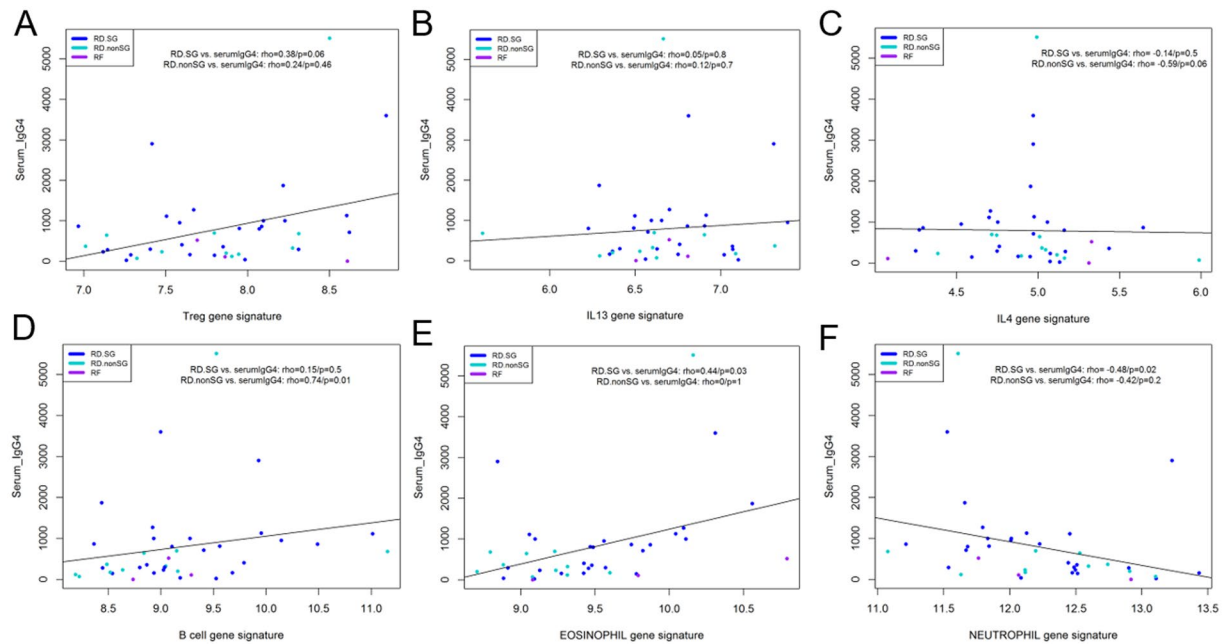


Figure 6. Correlation between serum levels of IgG4 and (A) Treg signature, (B) IL13 signature, (C) IL4 signature, (D) B cell signature, (E) eosinophil signature, (F) neutrophil signature in RD-SG, RD-nonSG, and RF patients.

Within IgG4-RD specifically, a microarray study observed neutrophil-specific genes (DEFA3 and DEFA4) significantly over-expressed in peripheral blood mononuclear cells (PBMCs) of patients on steroid therapy compared to those not²⁷.

Genes associated with mitosis, cell cycle, and replication were most correlated with *IgG4* expression. A previous study evaluating circulating autoantibodies in sera from IgG4-RD patients identified high levels of antibodies against prohibitin in patient subsets of autoimmune pancreatitis, MD, RF, IgG4-RD, and Sjögren's syndrome (not healthy donors)²⁸. The prevalence of anti-prohibitin auto-antibodies in IgG4-RD patients was hypothesized to increase cell proliferation, ultimately driving tissue enlargement. Additionally, a microarray study evaluating labial salivary glands in RD patients identified regulation of cell proliferation among the top enriched biological categories¹⁹. As increased IgG4 is a hallmark of this disease, the association between the phenotype and cell cycle processes is supported by these previous studies.

For RD patients with a longitudinal time point, there was an association between two patients that flared at the second visit (both on prednisone) and induction of Th1, Th2, Treg, and eosinophil gene signatures as well as *IgG4* and *IgE* mRNAs. In contrast, this pattern across these gene signatures/genes was not consistently observed in the four patients with stable disease status at the second visit. That is to say, no stable patient had multiple induced Th1, Th2, Treg, or eosinophil gene signatures at the second time point. The balance between cell-mediated immunity (Th1 cells), humoral immunity (Th2 cells), and maintenance of immune homeostasis (Tregs) and how this correlates with disease pathogenesis or activity have been investigated in rheumatoid arthritis and SLE, though conclusions have varied^{29,30}. Immune-activated over-expression of IL-4, -5, -10, -13, and TGF- β 1 drives eosinophilia and increased IgG4 and IgE levels in IgG4-RD, thus this suggests that at states of increased disease activity, Th1, Th2, Treg, and eosinophil involvement would be greater⁶.

An inverse relationship was observed between the flare visit and both the B cell signature and Tfh genes, where the gene signature/genes were suppressed at the flare visit. Similar to those gene signatures/genes showing induction at the flare visit, there was no consistent pattern of agreement in the four patients with stable disease across these gene signatures/genes at the second time point. At baseline, the B cell signature was significantly reduced in all patients compared to controls and was even more pronounced at the longitudinal flare time point in the two patients, suggesting cell infiltration to the disease tissue from the periphery in increased disease activity states. Tfh cells are located within germinal centers and secrete IL-21, driving differentiation of B cells to produce antibodies, thus the pattern showed by these Tfh-associated genes at the flare time point is consistent with that of the B cell signature³¹. A study in PBMCs of SLE patients showed that flares may be positively correlated with expansion of both Tfh and regulatory B cells through a regulatory feedback mechanism³¹. Another study in SLE found that a peripheral subset of CD27-IgD-CD97+ memory B cells were increased with disease flare, though the entire subset of CD27-IgD- B cells had no correlation with disease activity³². These results are in contrast to the pattern observed in this study at the flare visit in the two patients, though the B cell signature used here is not specific to either regulatory or memory B cells, and as indicated in the study by Jacobi *et al.*³², differences in B cell subsets can greatly vary with respect to phenotype.

No.	Sex	Age	organ involvement												
			sialadenitis + dacryoadenitis	eye	nose	deep lymph nodes	kidney	pancreatitis	liver	cholangitis	prostatitis	thyroid	hypophysis	retroperitoneal fibrosis	inflammatory pseudotumor
C1	F	38													
C2	M	31													
C3	M	30													
C4	M	48													
C5	F	52													
C6	M	32													
C7	M	49													
C8	M	57													
C9	F	48													
C10	M	30													
RD-SG1	M	55	X												
RD-SG2	M	65	X												
RD-SG3	M	44	X												
RD-SG4	F	54	X												
RD-SG5	M	55	X												
RD-SG6	F	50	X												
RD-SG7	M	56	X												
RD-SG8	F	41	X												
RD-SG9	M	54	X												
RD-SG10	F	49	X												
RD-SG11	M	32	X												
RD-SG12	F	51	X												
RD-SG13	F	81	X												
RD-SG14	F	56	X												
RD-SG15	F	34	X												
RD-SG16	M	64	X												
RD-SG17	F	55	X												
RD-SG18	F	52	X						X						
RD-SG19	M	59	X				X	X		X					
RD-SG20	M	49	X					X							
RD-SG21	M	73	X											X	
RD-SG22	M	61	X			X									
RD-SG23	F	56	X			X		X							
RD-SG24	F	60	X			X	X	X			X			X	
RD-SG25	F	49	X					X							
RD-non SG1	M	48		X	X	X									
RD-non SG2	M	59					X							X	
RD-non SG3	M	56						X		X					
RD-non SG4	M	66						X						X	
RD-non SG5	M	80								X	X			X	
RD-non SG6	M	62					X	X		X	X				
RD-non SG7	M	55					X	X			X	X	X		
RD-non SG8	F	68						X	X						
RD-non SG9	M	58						X		X					X
RD-non SG10	M	73						X		X					
RD-non SG11	F	55				X	X								
RF-1	M	65												X	
RF-2	M	54												X	
RF-3	M	48												X	

Continued

No.	serum IgG4 (mg/dl)	pathology		treatment							Longitudinal					
		number of IgG4+ plasma cell/HP	IgG4+ plasma cell/IgG+ plasma cell(%)	Pre-disone	Cyclophosphamide	Ursodesoxycholic acid	Azathioprine	Tamoxifen	Hydroxychloroquine	Methotrexate						
C1																
C2																
C3																
C4																
C5																
C6																
C7																
C8																
C9																
C10																
RD-SG1	288	>10	>40	5 mg qd	0.6/2 m											
RD-SG2	35.2	>10	>40	5 mg qd						0.2 bid						
RD-SG3	281	>10	>40	5 mg qd	0.4/m											
RD-SG4	402	>10	>40	5 mg qd	0.4/2 m											
RD-SG5	710	>10	>40													
RD-SG6	948	>10	>40													
RD-SG7	3600	>30	>40													
RD-SG8	18.5	>10	50	15 mg qd												
RD-SG9	230	ND	ND	7.5 mg qod							7.5 mg qw					
RD-SG10	159	>10	>40	5 mg qd	0.4/m											
RD-SG11	1000	100	50													
RD-SG12	859	100	>40													
RD-SG13	357	>10	>50													
RD-SG14	1870	>10	>40													
RD-SG15	145	>10	>40	5 mg qd												
RD-SG16	155	40	50	2.5 mg qd												
RD-SG17	296	ND	ND	2.5 mg qd						0.2 bid						
RD-SG18	1130	ND	ND													
RD-SG19	800	>10	>40													
RD-SG20	1270	ND	ND									X				
RD-SG21	807	>10	>40									X				
RD-SG22	2900	100	>40	30 mg qd												
RD-SG23	1110	ND	ND	60 mg qd						0.2 bid						
RD-SG24	864	>10	>40													
RD-SG25	1001	>10	>40													
RD-non SG1	5510	>50	50													
RD-non SG2	117	>50	>40													
RD-non SG3	364	>10	>40	20 mg qd		250 mg tid										
RD-non SG4	69.9	>10	>40	10 mg qd	0.4/2w							X				
RD-non SG5	642	ND	ND	2.5 mg qd		250 mg tid	50 mg qd	10 mg bid				X				
RD-non SG6	200	>10	>40	5 mg qd								X				
RD-non SG7	697	50	100	20/17.5 mg qod			50 mg qd					X				
RD-non SG8	680	ND	ND													
RD-non SG9	172	>30	>40	2.5 mg qd												
RD-non SG10	230	>10	>40	10 mg qod												
RD-non SG11	324	>10	>40				50 mg qod									
RF-1	105	>10	>40													
RF-2	520	>10	>40													
RF-3	<0.357	>10	>40													

Table 1. Subject summary: Control, RD-SG, RD-nonSG, and RF.

In summary, we show the importance of the T and B cell axis with molecular profiling across RD-SG, RD-nonSG, and RF as well as features that distinguish these three diseases. Future work seeks to better understand the molecular mechanisms at relapse or recurrence following steroid reduction in these patients.

Methods

39 patients fulfilled the 2011 comprehensive IgG4-RD diagnostic criteria were involved in this study³³. Among them, 26 patients were classified as definite IgG4-RD, 6 patients were classified as probable IgG4-RD and 7 patients were classified as possible IgG4-RD. Blood was procured from 25 RD-SG (ages 32–81; 12 Males), 11 RD-nonSG (ages 48–80; 9 Males), 3 RF (ages 48–65; 3 Males) and 10 control (ages 30–57; 7 Males) Chinese subjects (Table 1). Any organ with the salivary and lacrimal gland was involved in RD-SG patients. Except the salivary and lacrimal gland, other organs were involved in RD-nonSG patients. However, only retroperitoneal fibrosis was found in RF patients. Twenty patients were treated with prednisone (≤ 60 mg), with or without another glucocorticoid-sparing agent (cyclophosphamide, ursodeoxycholic acid, azathioprine, tamoxifen, hydroxychloroquine, methotrexate, and/or mycophenolate mofetil). Six patients had one longitudinal time point: two patients exhibited a flare at this second time point, while four did not. All participants provided written informed consent, in accordance with the Declaration of Helsinki. The study was approved by the Ethical Committee of Peking University People's Hospital.

In this study, stable or active condition was defined for every subject at the first visit. Stable condition was defined as the disappearance of clinical symptoms, normalization or stabilization of serum IgG or IgG4, and resolution of organ manifestations on imaging. Or else, it was defined as active condition. At the longitudinal time point, we defined flare condition as a recurrence of symptoms with the development or reappearance of organ involvement or abnormalities on imaging studies and elevation of serum IgG or IgG4 level.

The IL-13 and IL-4 gene signatures were identified as what have been described previously^{34,35}. The B and plasma cell gene signatures were developed using experiments as described in Streicher *et al.*³⁶. The eosinophil and neutrophil gene signatures were identified from a phase 1 clinical trial in systemic lupus erythematosus (SLE)³⁷. Baseline blood cell counts of SLE patients were correlated with whole genome microarray transcript profiles measured in the blood of the same patients. Th1, Th2, Treg, and Tfh gene signatures were taken from Dong *et al.*³⁸. The genes that compose each gene signature are provided in Supplementary Table 3.

The methods for RNA sequence read mapping and differential expression analysis are provided in the Supplementary Methods.

References

- Moriyama, M. *et al.* T helpersubsets in Sjögren's syndrome and IgG4-related dacryoadenitis and sialoadenitis: a critical review. *J Autoimmun.* **51**, 81–88 (2014).
- Beyer, G., Schwaiger, T., Lerch, M. M. & Mayerle, J. IgG4-related disease: a new kid on the block or an old acquaintance? *United EurGastroent* **2**, 165–172 (2014).
- Kamisawa, T. *et al.* Standardsteroid treatment for autoimmune pancreatitis. *Gut* **58**, 1504–1507 (2009).
- Hart, P. A. *et al.* Long-termoutcomes of autoimmune pancreatitis: a multicenterinternational analysis. *Gut* **62**, 1771–1776 (2013).
- Sah, R. P. *et al.* Differences in clinical profile and relapse rate of type 1 versus type 2 autoimmune pancreatitis. *Gastroenterology* **139**, 140–148 (2010).
- Stone, J. H., Zen, Y. & Deshpande, V. IgG4-related disease. *N Engl J Med* **366**, 539–551 (2012).
- Kawa, S. *et al.* HLA DRB10405-DQB10401 haplotype isassociated with autoimmune pancreatitisin the Japanese population. *Gastroenterology* **122**, 1264–1269 (2002).
- Park, D. H. *et al.* Substitutionof aspartic acid at position 57 ofthe DQbeta1 affects relapse of autoimmune pancreatitis. *Gastroenterology* **134**, 440–446 (2008).
- Chang, M. C. *et al.* T-cell regulatory gene CTLA-4 polymorphism/haplotype association with autoimmune pancreatitis. *ClinChem* **53**, 1700–1705 (2007).
- Umamura, T. *et al.* Association of autoimmune pancreatitis with cytotoxic T-lymphocyte antigen 4 genepolymorphisms in Japanese patients. *AmJ Gastroenterol* **103**, 588–594 (2008).
- Umamura, T. *et al.* Geneticassociation of Fc receptor-like 3 polymorphismswith autoimmune pancreatitis inJapanese patients. *Gut* **55**, 1367–1368 (2006).
- Zen, Y. *et al.* Th2and regulatory immune reactions are increasedin immunoglobulin G4-related sclerosing pancreatitis and cholangitis. *Hepatology* **45**, 1538–1546 (2007).
- Miyake, K. *et al.* Peripheral CD4 + T cells showing a Th2phenotype in a patient with Mikulicz'sdisease associated with lymphadenopathyand pleural effusion. *Mod Rheumatol* **18**, 86–90 (2008).
- Kudo-Tanaka, E. *et al.* A case of Mikulicz's disease withTh2-biased cytokine profile: possible featurediscriminable from Sjogren's syndrome. *Mod Rheumatol* **19**, 691–695 (2009).
- Kanari, H. *et al.* Role of Th2 cells in IgG4-relatedlacrimal gland enlargement. *Int Arch Allergy Immunol* **152**, 47–53 (2010).
- Detlefsen, S., Sipsos, B., Zhao, J., Drewes, A. M. & Klöppel, G. Autoimmune pancreatitis: expression and cellular source of profibrotic cytokines and their receptors. *Am J SurgPathol* **32**, 986–995 (2008).
- Miyoshi, H. *et al.* Circulating naive and CD4 + CD25highregulatory T cells in patients with autoimmune pancreatitis. *Pancreas* **36**, 133–140 (2008).
- Maehara, T. *et al.* Interleukin-21 contributes to germinal centre formation and immunoglobulin G4 production in IgG4-related dacryoadenitis andsialoadenitis, so-called Mikulicz's disease. *Ann Rheum Dis* **71**, 2011–2019 (2012).
- Tsuboi, H. *et al.* DNA microarray analysis of labial salivary glands in IgG4-related disease: comparison with Sjögren's syndrome. *Arthritis Rheumatol* **66**, 2892–2899 (2014).
- Tanaka, A. *et al.* Th2 and regulatory immune reactions contribute to IgG4 production and the initiation of Mikulicz disease. *Arthritis Rheum* **64**, 254–263 (2012).
- Yamamoto, M. *et al.* Evaluation and Clinical Validity of a New Questionnaire for Mikulicz's Disease. *Int J Rheumatol* **2012**, 283459 (2012).
- Yamamoto, M. *et al.* Beneficial effects of steroid therapy for Mikulicz'sdisease. *Rheumatology (Oxford)* **44**, 1322–1323 (2005).
- Della Torre, E. *et al.* Prevalence of atopy, eosinophilia, and IgE elevation in IgG4-related disease. *Allergy* **69**, 269–272 (2014).
- Chen, H. *et al.* IgG4-related disease in a Chinese cohort: a prospective study. *Scand J Rheumatol* **43**, 70–74 (2014).
- Watson, S. J., Jenkins, D. A. & Bellamy, C. O. Nephropathy in IgG4-related systemic disease. *Am J SurgPathol* **30**, 1472–1477 (2006).
- Nakagawa, M. *et al.* Glucocorticoid-induced granulocytosis: contribution of marrow release anddemargination of intravascular granulocytes. *Circulation* **98**, 2307–2313 (1998).

27. Nakajima, A. *et al.* Decreased Expression of InnateImmunity-Related Genes in Peripheral Blood Mononuclear Cells from Patients withIgG4-Related Disease. *PLoS One* **10**, e0126582 (2015).
28. Du, H. *et al.* Prohibitin Is Involved in Patients with IgG4 Related Disease. *PLoS One* **10**, e0125331 (2015).
29. Chen, J. *et al.* Comprehensive evaluation ofdifferent T-helper cell subsets differentiation and function in rheumatoidarthritis. *J Biomed Biotechnol* **2012**, 535361 (2012).
30. Yang, J. *et al.* Th17 and naturalTreg cell population dynamics in systemic lupus erythematosus. *Arthritis Rheum* **60**, 1472–1483 (2009).
31. Yang, X. *et al.* T follicular helpercells and regulatory B cells dynamics in systemic lupus erythematosus. *PLoS One* **9**, e88441 (2014).
32. Jacobi, A. M. *et al.* Activated memory B cell subsetscorrelate with disease activity in systemic lupus erythematosus: delineation byexpression of CD27, IgD, and CD95. *Arthritis Rheum* **58**, 1762–1773 (2008).
33. Umehara, H. *et al.* Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. *Mod Rheumatol* **22**, 21–30 (2012).
34. Yao, Y. *et al.* Development of PotentialPharmacodynamic and Diagnostic Markers for Anti-IFN- α Monoclonal Antibody Trials in Systemic Lupus Erythematosus. *Hum Genomics Proteomics*, 2009, pii: 374312 (2009).
35. Yao, Y. *et al.* Neutralization ofinterferon-alpha/beta-inducible genes and downstream effect in a phase I trial ofan anti-interferon-alpha monoclonal antibody in systemic lupus erythematosus. *Arthritis Rheum* **60**, 1785–1796 (2009).
36. Streicher, K. *et al.* The plasma cell signature in autoimmune disease. *Arthritis Rheumatol* **66**, 173–184 (2014).
37. Petri, M. *et al.* Sifalimumab, a human anti-interferon- α monoclonal antibody, in systemic lupus erythematosus: a phase I randomized, controlled, dose-escalation study. *Arthritis Rheum* **65**, 1011–1021 (2013).
38. Dong, C. & Martinez, G. J. T cells: the usual subsets. *Nat Rev Immunol* **10**, 9 (2010).

Acknowledgements

This work was supported by grants from Research and Development Fund of Peking University People's Hospital (RS2016-03). We would like to acknowledge all the subjects enrolled in this study.

Author Contributions

B.W.H., Y.Y.L. and J.P.G. carried out most of the experiments and drafted the manuscript. Y.N.S., C.M., W.Z., M.R.L. and Y.D. contributed to experiment implementation. L.M.R., G.Y.Y., L.L.D., H.H., P.W., Y.W. and Z.G.W. collected the clinical data. Y.H.Y. and Z.G.L. participated in the conception and design of the study. All authors contributed to critical revision, and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-17602-9>.

Competing Interests: The work described in this article was supported by MedImmune. All authors affiliated to MedImmune are full-time employees and have stock in AstraZeneca. Other authors have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017