Supplementary Appendix

Supplement to: Baghdassarian H, Blackstone SA, Clay OS, et al. Variant *STAT4* and response to ruxolitinib in an autoinflammatory syndrome. N Engl J Med 2023;388:2241-52. DOI: 10.1056/NEJMoa2202318

This appendix has been provided by the authors to give readers additional information about the work.

Supplementary Appendix 1

Variant STAT4 and autoinflammatory syndrome responsive to ruxolitinib

Table of Contents

Author contributions

Lori Broderick, Hratch Baghdassarian, Sarah A. Blackstone, Rachael Philips, Hirotsugu Oda, John J. O'Shea, Joshua D. Milner, Daniel L. Kastner and Christopher D. Putnam conceived the overall experimental designs.

Lori Broderick, Hratch Baghdassarian, Nathan Lewis, Vivian K. Hua, David R. Murdock, Nobuyuki Horita, Shimul Chowdhury, David Dimmock, Sofia Rosenzweig, Brynja Matthiasardottir, Elaine Remmers, Adam Mark, Roman Sasik, Kathleen M. Fisch, and Kristen Jepsen performed sequencing and analysis for genomic and single cell RNA sequencing studies. Brynja Matthiasardottir and Rachael Philips performed bulk RNA sequencing and analyses. Christopher D. Putnam, Elif Erin, Norman R. Watts and Hirotsugu Oda performed structural modeling and analyses. Lori Broderick, Owen Clay, Rachael McVicar and Yang Liu initiated and maintained primary fibroblast lines. Lori Broderick, Vivian K. Hua, and Owen Clay performed and analyzed wound healing assays. Lori Broderick, Sarah A. Blackstone, and Pallavi Pimpale Chavan performed and analyzed studies involving *in vitro* expression, plasmid construction, flow cytometry, and ELISA. Lori Broderick, Sarah A. Blackstone, and Davide Randazzo performed immunofluorescence staining and microscopy. Suzanne Tucker performed clinical histology and immunohistochemistry. Chi Ma, Massimo Gadina, Daniella M. Schwartz and Joshua Milner performed and analyzed *ex vivo* flow cytometry. Lori Broderick, Suzanne M. Tucker, Natalie Deuitch, Michele Nehrebecky, Anwesha Sanyal, Giffin Werner, Raphaela Goldbach-Mansky, William A. Gahl, Johanna Chang, and Kathryn Torok were responsible for clinical samples.

Lori Broderick, Hratch Baghdassarian, Johanna Chang, Daniel L. Kastner and Christopher D. Putnam wrote the initial draft of the manuscript with contributions from Sarah A. Blackstone, Brynja Matthiasardottir, David Dimmock and Kathryn Torok. All other authors reviewed the final draft and agreed to publish the manuscript.

Detailed case presentations

Family 1

The elder of the two siblings (STAT4 p.Ala635Val) presented with severe oral ulcerations in early childhood leading to inability to protrude his tongue by age 3 years. At age 16, he presented to Rheumatology clinic for a newly developed waxy, pale flat lesion on his chest, with biopsy consistent with morphea. Over a 5-week period, the skin lesions increased in size, to involve the forearms and posterior ears, and subsequently progressed to involve the anterior tibia, as well as the face. His initial laboratory evaluation was notable for mild elevation of CRP with normal IgG and IgM but absent IgA. White blood cell and absolute lymphocyte counts were mildly decreased and subset analysis demonstrated decreased CD4 and CD8 T cells. He was treated with intermittent pulse methylprednisolone 1 g daily x 3 days, methotrexate 25 mg SQ weekly and prednisone 20 mg daily. Despite therapy, he continued to rapidly develop new lesions on his back with ulceration and spreading waxy hypopigmentation, leading to addition of mycophenolate mofetil and PUVA therapy for 6 weeks. The rash continued to progress with ulcerating lesions on the skin over his buttocks, and new erythematous, raised, pruritic rash on his hands and feet, with biopsy suggestive of nodular keloids, with minimal response to laser therapy. Six months later, he developed erythematous nodular lesions on his chest, with persistence of lesions on his hands, wrists, ankles, feet and upper legs, while the waxy hypopigmented lesions on his face, upper arms and lower legs largely subsided. The patient continued to note intermittent aphthous ulcers. In an attempt to taper his daily prednisone and methotrexate, monthly intravenous immunoglobulin was initiated at 2g/kg, with no further spread of the lesions. Four years into his disease course, identification of a mutation in *STAT4* led to introduction of the JAK inhibitor ruxolitinib 5 mg BID, and discontinuation of mycophenolate mofetil. Ultimately, ruxolitinib was increased to 10 mg BID and the monthly IVIG dose decreased to 1 g/kg monthly. During this time, he had notable improvement in weight and in his nodular rash, without development of new lesions. After 11 months of therapy, the rash and oral ulcers had largely resolved, and IVIG was administered only for replacement dosing. His most recent laboratory evaluation is notable for stable white blood cell, neutrophil, and lymphocyte counts, normal inflammatory markers and normal IgG and IgM. He continues to have IgA deficiency. No adverse events have been reported on ruxolitinib. By eighteen months after initiation of ruxolitinib, he had discontinued all other medications, with complete resolution of rash on his chest and significant clearing of both hands and feet.

The younger sibling presented at age 5 with neck limitation, bilateral wrist arthritis, bilateral knee arthritis, and arthritis of the metatarsophalangeal joints. He had significant diffuse muscular weakness, with normal clinical laboratory studies including serum inflammatory markers. Radiographs were notable for diffuse osteopenia without joint space narrowing or erosive disease. He was diagnosed with polyarticular JIA, with initial treatment with prednisone and naproxen. However, due to lack of improvement, he was started on methotrexate. He subsequently developed methotrexate intolerance with complaints of shaking, fever, and abdominal pain and was switched to etanercept. Persistent elevations in ESR prompted a switch to infliximab with initial dosing at 6 mg/kg and increased up to 10 mg/kg. His exam continued to be notable for persistently active bilateral arthritis of elbows, wrists, knees and ankles with an unsteady gait, and flexion contractures of the hands. He was also noted to have erythema of the upper eyelids, as well as a rash on the cheeks that was thought to be due to pseudoporphyria from naproxen use. Despite apparent improved range of motion on infliximab, he subsequently developed a right leg contracture with shiny firmness of the skin over his hands and feet as well as a scaly rash over his upper chest and hyperpigmentation of the neck. He continued to have diffuse arthritis as well as subcutaneous tissue loss without calcinosis. He was started on daily leflunomide, but switched to cyclosporine given worsening of his skin

findings including tautness of the skin over the dorsum of hands and forearms as well as over the dorsum of his feet and legs below the knees. He was also noted to have tautness of the skin around the lateral portion of his neck near his jaw as well as his upper chest and lower neck. Additional evaluations included an ophthalmologic exam negative for uveitis, negative chest CT, and normal cardiac echocardiogram. Light therapy was started with some improvement in skin lesions, but he continued to have severe limitation of wrists, no motion in the ankles or subtalar joints, severe flexion contractures of the fingers, and decreased range of motion of the wrists. Skin biopsy at an outside center was reported to be consistent with scleroderma.

He was started on treatment for localized scleroderma with methylprednisolone pulsing (3 consecutive days for 3 months) and combination therapy including leflunomide, cyclosporine and prednisone. Despite this regimen, he continued to develop large linear scleroderma lesions on the side of his face, along his neck and the back of his neck and extending from the distal arms into the fingers. He had severe contractures of the feet with the lesions extending up the thighs from the feet, sparing the knees. Leflunomide was discontinued due to elevated LFTs. Exam at that time demonstrated significant weight loss from the 40th percentile to the 10-25th percentile and height to be in the 3rd percentile of age appropriate normals. He continued to have worsening scleroderma of the face without ability to note any normal skin on the face and worsening knee contractions, and the diagnosis of pansclerotic morphea was made. The patient was continued on methylprednisolone pulses and cyclosporine, and minocycline was added. Labs were notable for worsening LFTs and aldolase values, preceding a presumed scleroderma renal crisis and right ankle cellulitis/bullous impetigo requiring IV antibiotics, and new diagnosis of restrictive lung disease. Given the refractory course and progressive systemic complications, autologous bone marrow transplant was proposed. For stem cell mobilization, he received 1 dose of 2 g/m^2 cyclophosphamide followed by G-CSF and stem cell collection. Shortly afterwards, he began to develop right lower leg ulcers with Rodnan score 3 diffusely (scleroderma skin scoring). He underwent a non-myeloablative preparatory regimen followed by stem cell rescue procedure, consisting of 50 mg/kg per day times 4 days of cyclophosphamide followed by 1.5 mg/kg of rapid anti-thymocyte globulin (ATG) with methylprednisolone followed by autologous stem cell infusion. After 3 months, his right leg ulcers persisted, however, the rest of his skin lesions appeared to have improved and he was noted to have new hair growth on the skin. Over the next several months, his Rodnan score improved from 3 to 1 or 2 in multiple quadrants of the body including the back, upper chest, lower abdomen, and upper arms, with improved laboratory values including normal CBCD, and LFTs. His IgG remained low at 473 (582-1441 mg/dL) and monthly replacement IVIG to keep IgG > 500 mg/dL was initiated. Unfortunately, 1 year post-transplant, he had relapse of his skin disease with worsening joint contractures, skin tightening, hair loss and loss of function. Immunomodulatory therapy was re-initiated with intravenous cyclophosphamide every 2 weeks and daily oral imatinib mesylate. The imatinib mesylate was stopped shortly afterwards due to persistent neutropenia and myelosuppression. He continued to have sclerodermatous skin with new breakdown in the chest region leading to initiation of bosentan. However, disease progression with resorption of the distal bones of his hands, worsening thrombocytopenia and bleeding from the skin led to discontinuation of bosentan. His case was further complicated by recurrent cellulitis requiring intravenous antibiotics and subsequent squamous cell carcinoma. Despite aggressive radiotherapy, he had poor healing after the biopsy and also developed several similar lesions on the dorsum of his right foot. Five years post-transplant, he underwent right above the knee amputation and bone marrow biopsy which showed mixed cellular marrow (75% cellularity) without malignancy, but his surgery was complicated by the development of an ulcer that required surgical revision, and his replacement immunoglobulin was increased to immunomodulatory dosing at 2 g/kg monthly. He continued to have persistent skin lesions with

frequent cellulitis and bleeding episodes, ultimately electing to have bilateral through-humerus amputations, complicated by post-operative liver failure, and right ventricular hypertension with subsequent development of splenomegaly, and portal hypertension. Nearly 9 years post-transplant, prompted by the genetic results, he was started on ruxolitinib. However, the patient then developed anxiety attacks and ruxolitinib was held. He subsequently underwent elective left lower extremity through joint amputation as well as chest wall biopsy. Ruxolitinib was held during this time but then restarted 2 weeks after surgery, with stabilization of his pulmonary hypertension. He continues on IVIG 2 g/kg for IgG levels < 1000 but has been able to avoid infusions for up to 2 months at a time over the last year. Most recent labs are notable for resolution of neutropenia with normal inflammatory markers, with persistent IgM and IgA deficiency.

The father of P1 and P2 has a history of oral ulcerations and less severe skin disease without a formal diagnosis, whereas the mother has no history of similar disease.

Family 2

The proband of Family 2 (STAT4 p.Ala650Asp) presented at 3 years of age with joint swelling of the ankles, knees, and elbows, and an inability to keep up with his peers. At the age of 7, he developed painful bilateral hand contractures and shortly after, developed a white patch on his left leg. He was given the diagnosis of juvenile rheumatoid arthritis, was hospitalized and treated with naproxen and methotrexate. He subsequently underwent biopsy consistent with scleroderma and he was treated with steroids for 6 months. By the following year, the fibrotic skin lesions became confluent and a diagnosis of generalized morphea was made. He was subsequently evaluated for possible eosinophilic fasciitis after a small ulcer on the right foot spread to the whole foot (except the sole) and progression of other skin ulcerations occurred. He had depigmented areas on the extremities and face with pruritus. In addition, he had increased contractions in the hands and feet and decreased use of the extremities. Lab studies showed an elevated ESR and profound peripheral eosinophilia. ANA and rheumatoid factor were negative. At that time diagnoses considered were scleroderma with possible overlap features of eosinophilic fasciitis. He was treated with infusions of methylprednisolone and then put on methotrexate. He responded to methotrexate without any major extension of the morphea. Six months after the drug was stopped, there was almost an "explosion" of skin sclerosis, preceded by inflammatory pruritus. He was treated with IM methotrexate and penicillamine for 8 months. Penicillamine was discontinued due to neutropenia. After 5 years of disease, he had generalized morphea in all stages of activity, involving the legs, buttocks, groin, shoulders and arms plus some deformity of the hands, ankles and feet with fixed flexion in the hands and wrist joints. He was treated with three pulses of methylprednisolone 500 mg for three consecutive weeks, as well as 10 mg of prednisone on a daily basis. Azathioprine and cyclosporine were added but he developed a severe infection, and he was treated with IV prostacyclin and hyperbaric oxygen. By report, the lesions did not get any worse after treatment with hyperbaric oxygen however when it was discontinued the ulcers worsened, and methotrexate was restarted. He underwent surgery to the hand to improve flexibility; upon surgical debridement of the foot ulcer, cultures grew *Serratia marcescens* and *Enterococcus faecalis*. After 8 years of disease, he had developed contractures of the knees, with maximum straightening of the right leg of about 150 degrees and maximum straightening of the left knee of less than 90 degrees, and contractures/subluxation of the ankles and elbows, and he was unable to close his hands. He had chronic foot ulcers on the right side, covering the back and distal lower legs with small ulcers on the upper legs and arms. Biopsy of the foot ulcer showed only granulation tissue, without evidence of neoplasia. His skin thickened and scarred causing extensive contractures, both axial and peripheral, limiting movement and leading to the patient being wheel chair bound. At initial evaluation, laboratory examinations were significant for ESR

of 32, aldolase 11, CPK 32. Imaging included a chest x-ray with mild interstitial changes. Pathology included a lymph node biopsy suggestive of CMV infection, with bronchoalveolar lavages demonstrating increased eosinophils in the vessels. His bone marrow biopsy was unremarkable. Over the course of disease, he had persistently elevated inflammatory markers including ESR and CRP with negative autoantibodies. Over the 15 years of disease, his immunosuppressive therapy regimen included oral and intravenous steroids with a maximum dose of 32 mg prednisolone, intermittent methotrexate, maximum dose of 25 mg subcutaneously, a trial of D-penicillamine, withdrawn because of severe leukopenia that resolved after using granulocyte and growth factors, azathioprine, cyclosporine and 17 pulses of intravenous gammaglobulin (1 g/kg/day for two days). Therapies were ultimately discontinued in the setting of open wounds and risks for infection. Ulcers were primarily treated by surgical debridement and local wound care. By his twenties, he was diagnosed with 50% hearing loss. He also lost vision suddenly in both eyes from cataracts. He had corrective cataract surgery with successful return of 20/20 vision in both eyes. At age 31, the patient died following an infection.

The proband's mother in Family 2 also has a presentation of joint deformities that began at age 20 years. She reported "electric pain" in her upper extremities with progressive hand joint swan-neck deformities over a 6-month period. The progression subsequently stabilized and was no longer painful, but she was left with marked residual deformities. She later developed bilateral cataracts requiring surgical correction at age 40, and bilateral hearing loss requiring hearing aids by the age of 50.

Family 3

The proband in Family 3 (STAT4 p.His623Tyr) developed a non-healing wound on his lower extremity at 9 months of age, followed by poor weight gain, loose stools, 'wasted appearance' of his lower extremities with firmness of the subcutis throughout the upper and lower extremities, with decreased range of motion and mild swelling of several small to large joints by age 12 months. Further history revealed recurrent tonsillitis, two pneumonias, with normal CBC and IgG but low IgA. Evaluation was extensive including cardiopulmonary, gastroenterology, genetics, and musculoskeletal studies. Significant findings in the musculoskeletal imaging pathology were found consistent with deep tissue inflammation and sclerosis of several levels from subcutis to muscle, consistent clinically with disabling pansclerotic morphea of childhood. An ultrasound of the lower extremities demonstrated obvious irregular nodular thickening of the subcutaneous soft tissues throughout bilaterally and arthritis of the knees. Further imaging (before treatment) included MRI hip/girdle and lower extremities (**Figure S1**) demonstrated subcutis edema with extensive fasciitis with associated adjacent muscle edema, which was directly supported histologically via full thickness skin biopsy of the left thigh (**Figure S1**) that showed marked sclerotic changes of the deep fascia and fibrous trabeculae of the subcutaneous fat with extension into the dermis with an associated moderate lymphoplasmacytic infiltrate of the periadnexal, perivascular and septal areas. Noted was extensive expanded fascial tissue with inflammation and sclerosis and associated neighboring skeletal muscle lymphocytic infiltrate (**Figure S1**).

He was treated (**Figure S4**) immediately with oral glucocorticoids (2mg/kg divided BID with taper over 9 months), in conjunction with methotrexate (1 mg/kg/week), UVA ~10 J/cm2 three times/week (4 months), and aggressive physical therapy and occupational therapy. Improvement in skin induration and joint range of motion (ROM) was noted within 6 months. Oral steroids were increased to potentially gain more benefit and tapered completely off by 24 months into therapy. Methotrexate continued. Repeat MRI of hip girdle/ lower extremities 18 months into treatment demonstrated full resolution of subcutis, fascia and muscle edema (and resolved inguinal and popliteal lymph node adenopathy). He had slow improvement with full

resolution of remaining skin thickness and improvement of ROM, but with remaining deficits mild in his fingers, elbows, hips and knees and more moderate in ankles and subtalar joints, over the next several years. He was weaned off methotrexate completely after 6 years of therapy, and within 6 months had a flare of tenosynovitis of the upper extremities, wrists and fingers, mostly PIPs, documented clinically and radiographically with MRI. There was no flare of the skin, subcutis or fascia. Methotrexate with oral prednisone taper was restarted and led to improvement within 6 months. Due to tolerance issues 2 years later, he was switched from methotrexate to adalimumab to treat polyarthritis, with his only remaining symptoms being morning stiffness and synovial hypertrophy on examination. He has experienced full arthritis response to adalimumab and he was been in clinical remission in skin and joint symptoms for the past 6 years. He is currently 17 years old and fully participates in track and football sports at high school and has no complaints. He does have remaining moderate joint contractures of the subtalar and ankle joints with associated radiographic narrowing of the anterior subtalar joint space and osteophyte formation at the talonavicular joint supporting early degenerative changes, but other joints have little or no remaining joint contractures. There have been no additional immune lab abnormalities, serious or chronic infections or other autoimmune diagnoses over the 15 years of follow up.

SUPPLEMENTARY METHODS

DNA isolation, library construction, and sequencing

Family 1. Rapid whole genome sequencing largely followed methods as previously outlined. Blood was drawn following consent for rapid whole-genome sequencing[1] of the family into Rady Children's Institute Genomic Biorepository NCT02917460. Return of results on this protocol was limited to pathogenic and likely pathogenic variants in response to the FDA oversight.[2] DNA was subsequently extracted using EZ1 DSP DNA Blood Kit and sequenced using 2 × 101 base pair run on a HiSeq 2500 System (Illumina) in rapid-run mode to a ~45-fold coverage. Rapid alignment and nucleotide variant calling were performed using the Dragen (Edico Genome) hardware and software. Single nucleotide variants were annotated and analyzed in Opal (Omicia). Initially, variants were filtered to retain those with allele frequencies of <1% in the Exome Variant Server, 1000 Genomes Samples, and Exome Aggregation Consortium database[3] (http://evs.gs.washington.edu/EVS/2016). A gene panel was built in Phenolyzer[4] using Human Phenotype Ontology (HPO).[5] Structural variants were identified with Manta[6] and CNVnator,[7] a combination that provided the highest sensitivity and precision on 21 samples with known structural variants. Structural variants were filtered to retain those affecting coding regions of known disease genes and with allele frequencies <2% in the RCIGM database. Variants were classified based on the guidelines established by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.[8] No variants reportable per protocol were identified. Analysis in QIAGEN Ingenuity Pathway Analysis revealed a rare variant in *STAT4* c.1904 C>T (p.Ala635Val) shared by both brothers and not inherited from their unaffected mother. These results were shared with the research team.

The variant was confirmed by Sanger sequencing in peripheral blood mononuclear cells and fibroblasts. Subsequent to completion of functional studies the variant could be reclassified and reported clinically to the patients.

Family 2. Whole genome sequencing from genomic DNA was performed by HudsonAlpha Clinical Services Lab using the Illumina HiSeq X platform. Mapping of FASTQ files to the GRCh37 human reference was done using the Burrows-Wheeler Aligner[9] (v. 0.7.15) followed by processing with Picard tools (v.2.1.1, http://broadinstitute.github.io/picard) on the NIH HPC Biowulf cluster (http://hpc.nih.gov). A variant call file (VCF) was generated with GATK (v. 3.6)[10] and subsequently annotated using ANNOVAR.[11] Strict filtering was done by evaluating for variants absent from the ExAC database[3] with *in-silico* predictions of pathogenicity, ultimately identifying the *STAT4* c.1949C>A (p.Ala650Asp) variant in both the proband and mother. Sanger sequencing confirmed the *STAT4* variant was present in the proband and his mildly affected mother, but absent in his father.

Family 3. Exome sequencing of the proband and his unaffected parents was performed on DNA derived from peripheral blood. DNA extraction was performed using the Maxwell 16 Blood DNA Purification Kit. Illumina TruSeq DNA Sample Preparation Kit version 2 was used to prepare sequence libraries. Paired end sequencing was performed on the Illumina HiSeq2000 instrument at the National Institutes of Health (NIH) Intramural Sequencing Center (NISC). Sequence reads were aligned to the reference build hg19 with BWA. PCR duplicates were marked using Picard MarkDuplicates and the results were coordinate sorted using SAMtools. Base recalibration and realignment around microindels were performed using GATK which allowed for more accurate base quality scores. GATK's genotype workflow was used to identify SNVs and indels. Identified variants were annotated with the Variant Effect Predictor (VEP). Given the phenotypic severity, variants were filtered for ultrarare or novel (MAF < 0.0001) variants. GEMINI allowed for filtering based on de novo or inheritance patterns. Multiple variant

impact prediction tools were then used for further filtering for variants predicted to be pathogenic. Subsequent analysis revealed the *STAT4* c.1867C>T, p.His623Tyr variant. The variant was confirmed by Sanger Sequencing. The BigDye Terminator v1.1 Cycle Sequencing kit (Applied Biosystems) was used for sequencing coding exons of *STAT4*. The sequencing was performed on a Seq Studio Genetic Analyzer (Applied Biosystems). The sequencing data were reviewed using Sequencher (Gene Codes) and the chromatograms shown for P4 and Family 3 are derived from Sequencher.

Molecular dynamics methods

A full-length human STAT4 model (amino acids 1-748) predicted by AlphaFold[12,13] was obtained from UniProt (Uniprot ID: Q14765). For molecular dynamics (MD) simulations a dimeric model containing the SH2 domain (amino acids 572-679), the phosphorylated tail segment (P-tail segment, amino acids 680-706) and the transcriptional activation domain (TAD, amino acids 708-748) was built in Chimera[14] based on the STAT1 crystal structure (PDB ID: 1BF5,[15]) The obtained model was in silico phosphorylated at Y693 using Coot.[16] The final phosphorylated model was then minimized sequentially using Yasara [17] and Chimera to remove clashes. Individual MD simulations of wild-type STAT4 and STAT4 containing the single mutations were carried out using the Desmond simulation package (Schrödinger Release 2017-3). Proteins were prepared by Protein Preparation Wizard, and the optimized potentials for liquid simulation force field (OPLS_2005) parameters were used in restraint minimization and system building [18]. The system was set up for simulation using a predefined water model (TIP3P) as a solvent. The electrically neutral system for simulation was built with 0.15 M NaCl in 10 Å buffer. The NPT ensemble with 300°K, and a pressure of 1 bar was applied in the run. The simulation was performed for 50 ns, and the trajectory sampling was done at an interval of 5 ps. The short-range coulombic interactions were analyzed using a cutoff value of 9.0 Å using the short-range method. The smooth particle mesh Ewald method was used for handling long-range coulombic interactions. MD simulations supported a gain-of-function model, and are available upon request.

Cell lines

Human primary skin fibroblasts were cultured from skin biopsy samples as previously described.[19] Briefly, 1mm skin biopsy pieces were placed in a 6 well plated coated with 0.1% gelatin in DMEM with 20% FBS and 1% Antibiotic-Antimycotic (Gibco). Media was replaced every 2-3 days, and cell cultures were confluent after 2-3 weeks. After passage 2, cells were maintained in Gibco Medium 106 (ThermoFisher) supplemented with Gibco Low Serum Growth Supplement (ThermoFisher) and 1% Antibiotic-Antimycotic (Gibco). Cells were used between passages 4 and 9.

The IL-6 Leeporter™ Luciferase Reporter, an NIH3T3 derivative stably expressing a luciferase construct driven by human *IL6* promoter, was purchased from Abeomics, Inc. (San Diego, CA) and maintained in DMEM medium (w/ L-glutamine, 4.5g/L glucose, sodium pyruvate) supplemented with 10% heat-inactivated FBS and 1% pen/strep, plus 3 µg/ml of puromycin. The cell line was validated with dose-response curves by seeding 5 x 10^4 cells/well on a white solid-bottom 96-well plate and stimulating for 16 hours with LPS (0.1-10,000 ng/mL) or IL-6 (1-1000 ng/mL). Abeomics luciferase assay reagent (#17-1101) was added to each well, and luminescence measured within 1-5 minutes on an EnSpire Plate Reader (PerkinElmer Inc.). For all assays, cells were used between passages 4-9.

U3A cells[20,21] were acquired from Sigma Aldrich. Cells were maintained in DMEM medium (w/ L-glutamine, 4.5g/L glucose, sodium pyruvate) supplemented with 10% heat-inactivated FBS and 1% pen/strep.

Mammalian transfection and expression of recombinant proteins
The human *STAT4* ORF was purchased from Dharmacon (Lafayette, CO) and cloned into the Gateway pcDNA DEST40 (ThermoFisher Scientific), which contains a C-terminal V5-6x His tag, per the manufacturer's instructions. The STAT4 variants c.1949C>A mutation encoding STAT4 A650D, c.1867C>T mutation encoding STAT4 H623Y, and c.1904 C>T mutation, encoding A635V were introduced using a QuikChange protocol, per the manufacturer's instructions. The *STAT4* coding sequence in each of the mutated plasmids was verified by sequencing by Eton Bioscience, Inc. (San Diego, CA). Plasmid DNA (1.5µg) containing either wild-type or mutant *STAT4* were transfected into the Leeporter cell line using LipoFectamine 2000 per manufacturer's instructions (ThermoFisher Scientific). Cells were co-transfected with pCXLE-EGFP (Addgene) to verify similar transfection efficiencies prior to proceeding with downstream assays.

For flow cytometry studies, human STAT4 ORF was cloned into pDONR223. The c.1949C>A mutation encoding STAT4 A650D, c.1867C>T mutation encoding STAT4 H623Y, and c.1904 C>T mutation, encoding A635V were induced using QuikChange Lightning protocol, following manufacturer instructions. Mutations were verified by Sanger sequencing performed by Psomagen (Rockville, MD). U3A cells were stably transfected with *PiggyBac*[22] vector (Addgene, cat# 80479) containing STAT4 wild-type or patient variants. Transfected cells were then selected using puromycin (0.5 µg/ml). After puromycin selection, cells were stimulated with doxycycline (0.5 µg/ml) for a period of 24 hours to induce STAT4 expression before use in experiments.

Flow cytometric assessment of STAT4 phosphorylation
Intracellular STAT4 phosphorylation was measured by flow cytometry. For experiments in which cell lines were stimulated or inhibited, IFNα (10,000 IU/mL) and/or ruxolitinib (2.5 µM) were added at indicated time points. Cells were fixed in 4% paraformaldehyde (PFA) at room temperature for 10 minutes, then permeabilized in 100% methanol overnight at -20 degrees. Cells were then washed 3 times in PBS containing 0.5% BSA (FACS buffer), and then resuspended in FACS buffer. Staining was performed at room temperature, using Phospho-STAT4 (Tyr693) Monoclonal Antibody in PE (ThermoFisher, 12-9044-42) at a 1:200 dilution. For experiments in which a live dead stain was used, the stain was added before the fixation step at 37 for 10 mins. Analysis was performed using FlowJo and GraphPad software.

For primary human peripheral blood cells, ficolled PBMC were cultured at 1-2 x 10^6 cells/mL complete media (RPMI 1640, 2 mM glutamine, 10% FBS, penicillin and streptomycin) with phytohemagglutinin (PHA, 3 µg/mL, Sigma) for 40-72 hours, then washed and rested for 2 hours in a 37 degree C incubator. Cells were filtered and aliquoted to 0.3-0.5 x 10^6/tube per condition, resuspending in RPMI 1640 without FBS. After resting for 1 hour, cells were stimulated with human recombinant IL-12 (25 ng/mL, R&D Systems) at 37 degrees C for 20 minutes. Live/dead dye (ThermoFisher) was added per manufacturer's protocol and cells were fixed with 1.6% PFA followed by permeabilization with cold 100% methanol overnight. Cells were stained with antibodies to cell surface markers (CD4 L200, CD45RO UCHL1) and pSTAT4 (38/p-Stat4, BD Biosciences) for 30 minutes at room temperature in the dark, washed, and then evaluated by flow cytometry. Cells were analyzed on the FSC/SSC-gated blasted cells.

Reverse transcription and quantitative PCR
RNA was isolated from patient and healthy donor primary skin fibroblasts using Trizol reagent (Life Technologies) per manufacturer's instructions. cDNA was synthesized using Taqman Reverse Transcription reagents (Applied Biosystems). Relative gene expression for STAT4 was determined using the following primers: STAT4: 5'-CAGTGAAAGCCATCTCGGAGGA-3' and 5'-

TGTAGTCTCGCAGGATGTCAGC-3' with GAPDH 5'-ACATCGCTCAGACACCATG-3' and 5'-TGTAGTTGAGGTCAATGAAGGG-3' as reference gene. Quantitative PCR was performed using 100nM each forward and reverse primer and iQ SYBR Green supermix (Bio-Rad) with a Bio-Rad CFX96 Real-Time System, and reaction parameters as per Bio-Rad instructions Data was visualized with CFX Manager v3.0 software, and relative gene expression ws determined using the 2ΔΔct method.

Immunocytochemistry

Clinical immunohistochemistry was performed on skin biopsy samples with mouse anti-CD3 (clone# LN10) and mouse anti-SMA (clone# ASM-1) using standard protocols.

For immunofluorescence analysis of fibroblasts, primary skin fibroblasts were plated on glass coverslips (Corning) 24 hours prior to staining. Cells were washed with dPBS, and fixed in 4% PFA for 20 minutes. After washing, cells were permeabilized with 0.1% Triton X-100 for 20 minutes. After a second wash, cells were incubated in Alexa-Fluor 488-phalloidin (1:40 in 1% BSA for 30 minutes or phospho-STAT4 (PA5-105861, 1:200, for 1 hour) protected from light. After washing, cells were stained with 10 μM DRAQ5 (ThermoFisher) for 5 minutes, coverslips mounted to glass slides and sealed with ProLong Gold anti-fade mountant (Life Technologies). Slides were imaged using a Leica TCS-SPE confocal microscope with 10X objective.

For immunofluorescence analysis of transfected HEK 293T cells, 2.5 x 10^4 cells were plated on 35mm glass bottom dishes (MatTek, **#**P35G-1.5-14-C). The next day, cells were transiently transfected with plasmid DNA (20 ng) containing wild-type or variant *STAT4* tagged with GFP using LipoFectamine 2000 under manufacturer's instructions (ThermoFisher Scientific). Micrographs depicted in fig. 2d represent single frames after live imaging performed on a Zeiss LSM780 confocal system driven by the ZEN Black software (Zeiss). During imaging, cells were maintained at 37°C and 5% CO2 throughout the experiment using dedicated CO2 and temperature controllers (Zeiss) connected to a heated stage insert (Pecon). Time lapse was captured employing a Plan Apochromatic 40X/1.4NA oil immersion lens (Zeiss) maintained at 37°C by an objective heater (Bioptechs). Expression and localization of the STAT4-GFP protein was visualized using a 488nm argon ion laser with pinhole size set at 1 Airy Units (AU) and digital zoom set at 1.0 whereas laser power and detector gain were adjusted to avoid pixel saturation. Phase contrast images were acquired simultaneously. Cells were imaged continuously for 30 min and stimulated with INFα, which was added directly to the cells while on the microscope stage. Recordings (1,024 pixels wide) were exported as .czi files and then converted to .ims files to generate movies and single frame images in Imaris 9.9 (Bitplane). For each image intensity levels were linearly adjusted (when needed) in Imaris 9.9.

mRNA sequencing

U3A cells stably transfected with A650D, A635V, H623Y, wildtype, or phospho-dead Y693A variants were plated at 6 x 10^5 cells per well, and doxycycline was added to induce STAT4 protein expression. After 24 hours, cells were either left unstimulated, or stimulated with IFNα (10,000 IU/mL) for 4 hours. Total RNA was prepared using RNeasy Micro Kit (Qiagen, #74004). A fraction of total RNA (400 ng) was processed into mRNA-seq library using Quant-seq 3' mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen, #015.96) with the PCR Add-on Kit (Lexogen, #020.96) following manufacturer's protocol. The libraries were sequenced as 100 cycles (read 1) and index (8 bases) on V1.5 kit on NovaSeq 6000 (Illumina). Raw sequencing data were processed with bcl2fastq (v2.20.0.422) to generate FastQ files.

RNA-seq analysis
Sequence reads were trimmed with Cutadapt (v2.10), aligned to human genome (build hg38) using STAR (v2.5.4a, options: --outFilterType BySJout --outFilterMultimapNmax 200 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.6 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000), bam files were created with SAMtools (v1.10) and transcript abundance quantified using subread (v2.0.2). Differentially expressed genes were identified using the edgeR R package with the following criteria: log2 fold change > 0.5, FDR < 0.1, and logCPM > 0.3.

Enzyme-linked immunosorbent assay (ELISA) and antibody arrays
Secreted IL-6 was assessed using the human IL-6 Duo-Set kit (R and D Systems, Inc.) per manufacturer's protocol. Matrix production of pro-collagen I alpha-1 and fibronectin was evaluated using human Duo-Set ELISA kits (R and D Systems, Inc.) per manufacturer's instructions. ELISA absorbances were measured on an EnSpire Plate Reader.

Wound healing assay
Primary skin fibroblasts were plated at 2 x 10^5 cells per well on 0.1% gelatin coated 6-well plates in Gibco Medium 106 (ThermoFisher, #M106500) supplemented with Gibco Low Serum Growth Supplement (ThermoFisher, #S00310) and 1% Antibiotic-Antimycotic (Gibco). Cells were allowed to adhere and become confluent for 48 hours, and monolayers scratched manually with a p200 pipette tip. Cells were washed once to remove cellular debris. Prior to experiments, culture medium was replaced with fresh medium. In some experiments, cells were pre-treated for 24h with ruxolitinib (Cayman Chemical, #11609) or neutralizing anti-IL-6 monoclonal antibody (MQ2-13A5, ThermoFisher). Images were collected every 4-6 hours with a 4X objective using ToupView software (ToupTek Photonics, version x64). A minimum of 5 images per timepoint were collected for each assay, until wound closure. All cultures were used between passages 4 and 9, and maintained in a 37°C incubator with 5% CO_2.

Imaging analysis
Images, standardized at 1024 x 768 pixels, were imported into Image J. For each scratch, the analysis grid tool was overlaid and a line drawn horizontally on the grid, and measured, with 8 measurements per scratch. Measurements were exported to Microsoft Excel and normalized to the initial scratch width. A minimum of 3 independent scratches were performed.

Contraction assay
A collagen-based contraction assay was modified from a protocol by Cell BioLabs, Inc. Briefly, type I bovine collagen (3 mg/mL, Advanced Biomatrix, 5005), 5X DMEM (Gibco, 12100-061), and 0.5 M NaOH were mixed on ice. Primary skin fibroblasts were harvested by trypsin-EDTA treatment, and resuspended in Gibco Medium 106 (ThermoFisher, M106500) with low serum growth supplement (ThermoFisher, S00310) at 2 x 10^6 cells per mL. The collagen lattice was prepared by mixing 2 parts of cell suspension with 8 parts of ice-cold collagen solution. The cell-collagen lattice was added to a 24 well plate, at 500 μl per well and incubated for 1 hour in a 37°C incubator with 5% CO_2 to support collagen polymerization. After 1 hour, 1 mL of Gibco Medium 106 (ThermoFisher, M106500) with low serum growth supplement (ThermoFisher, S00310), was added to each well. In some experiments, cells were treated with TGF-β at 10 ng/mL (R&D Systems, 240-B-010). The cultures were incubated for 2 days, and then collagen gels released with a sterile syringe. The collagen gel size was measured hourly.

10X Genomics single cell RNA-seq

Sample processing and data sets

Peripheral blood samples from both patients underwent red blood cell lysis using the ACK buffer protocol (ThermoScientific), and PBMCs were resuspended in 1X PBS with 0.04% BSA. Libraries were prepared using Chromium Single Cell 3' Reagent Kits v3 (User Guide CG00183 RevA). 10X Genomics RNA sequencing was conducted at the IGM Genomics Center, University of California, San Diego, La Jolla, CA. Less than 1 hour passed from time of collection, to loading on the 10X Genomics Chip. Cells were at 97% viability using trypan blue exclusion. The healthy donor control dataset ("5k Peripheral blood mononuclear cells (PBMCs) from a healthy donor (v3 chemistry)") was obtained from the publicly available FASTQ files at https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k_pbmc_v3.

Single-cell RNA-seq gene expression quantitation

Alignment of exonic reads to transcripts, identification of single-cell-specific barcodes, extraction of unique molecular identifiers (UMI), and filtering of low RNA-content barcodes were used to quantitate UMI counts in each single cell. UMI counts were determined for scRNA-seq data from the PBMCs from both patients and the control (3 samples). Quantitation was performed with CellRanger's (v3.1.0) *count* function.[23] Default parameters were used, with the exception of increasing the expected number of cells for the healthy donor sample to 5,000. The GRCh38 GTF and reference FASTA files required for the analysis were built from the Ensembl release 93 files using CellRanger.

Single-cell normalization and dataset integration

Raw UMI counts for the three samples were normalized and integrated using Seurat (v4.1.1).[24] Unless otherwise stated, all analyses were performed using default parameters. Cells with fewer than 200 non-zero genes and genes found in fewer than 3 cells were excluded. To further limit the impact of low-quality cells,[25] each sample was manually assessed on standard quality control metrics (the per cell total number of counts, total number of unique genes, and fraction of counts mapping to mitochondrial genes)[26] and filtered using heuristic thresholds.[27,28] Cells from the control sample with total counts greater than or equal to 20,000, total counts less than or equal to 3,750, or percent of counts mapping to mitochondrial genes greater than or equal to 20% were excluded. Cells from Patient 1 with total counts greater than or equal to 10,000, total counts less than or equal to 1250, total unique genes detected less than or equal to 350, or percent of counts mapping to mitochondrial genes greater than or equal to 20% were excluded. Cells from Patient 2 with total counts greater than or equal to 15,000, total counts less than or equal to 1500, total unique genes detected less than or equal to 500, or percent of counts mapping to mitochondrial genes greater than or equal to 20% were excluded. Next, each sample was normalized for count depth, scaled by a factor of 1e6 to arrive at counts per million (CPM), and the CPM values with a pseudo-count of 1 added were natural log-transformed.

Finally, we accounted for batch effects using Seurat's *IntegrateData*. First, for each sample, the top 2000 variable features for each normalized sample were identified using Seurat's *FindVariableFeatures* function to help identify integration features. The samples were then scaled (normalized by the standard deviation for each gene) and centered (subtracted the average expression for each gene) prior to running principal components analysis (PCA) to 50 principal components. Finally, the reciprocal PCA method was used to identify the integration anchors.

Dimensionality reduction and clustering

The integrated dataset was analyzed by PCA to reduce the dimensionality to 47 principal components (PCs), at which point the additional variance explained by considering more

principal components was less than 0.1%. We ran the uniform manifold approximation and projection (UMAP) algorithm on the PCs to further reduce the dimensionality to two for visualization. We constructed a shared nearest neighbors (SNN) graph based on Euclidean distance in the PCA space with Seurat's *FindNeighbors* function. Next, we clustered cells using Louvain modularity optimization; the resolution parameter of Seurat's *FindClusters* function was set to 0.5. We confirmed that the integrated patient dataset did not possess batch-specific clusters; each cluster contained cells arising from multiple samples (**Figure S14a**).

Cell Type Identification

We identified cell types from the scaled, integrated expression matrix and the cluster annotations using ScType[29] (**Figure S14b**). Input markers for each cell type were selected from the "Immune system" tissue of ScType's database. Clusters were annotated as the cell type that received the highest ScType score.

To verify cell type annotations, a list of markers for each cluster was identified using differential expression (DE) analysis. DE tested for significantly up- or down-regulated genes in a given cluster relative to all other cells in the dataset using a Wilcoxon rank-sum test. Only genes present in at least 25% of cells in the cluster and that showed at least a log2-fold-change (LFC) of 0.5 were tested. Genes with a Bonferroni corrected p-value less than or equal to 0.05 and a LFC greater than 1.5 were retained for assessment. Uncertainty in the marker expression of three clusters (clusters 2, 6, and 9 labelled as Naïve CD4+ T-cells, Naïve B cells, and gamma-delta T-cells respectively) led us to further sub-cluster these data and manually annotate them based on DE markers. For each cluster, we subsetted the batch-corrected expression matrix to cells only in that cluster and re-ran the dimensionality reduction and clustering as previously described. The PCA dimensionality in these instances was 30 PCs. The resolution parameter of Seurat's *FindClusters* function was adjusted according to the number of mixed cell type populations we expected to observe in each cluster based on DE results. DE genes for each subcluster were identified using this same testing approach, and cell types were annotated based on significantly differentially expressed genes within these sub-clusters. Sub-clusters derived from cluster 2 were re-annotated as Effector or Memory CD4+ T-cells, those from cluster 6 were re-annotated as Plasma, Pre-, and Memory B cells, and those from cluster 9 were re-annotated as gamma-delta T-cells or Regulatory CD4+ T-cells.

Differential expression across samples and Ingenuity Pathway Analysis

Differential expression (DE) analysis between cell types across different samples was conducted using MAST (v1.20.0).[30] DE was performed on the log-normalized expression matrix. To account for technical variability and other nuisance factors, we introduced the cellular detection rate (CDR)--the fraction of genes expressed in a cell--as a covariate in the design matrix.

We tested for significant differential expression using a likelihood ratio test; the full model regressed expression on both the test condition and the CDR, and the reduced model regressed expression only on the CDR. We test for differential expression of genes in Patient 1 relative to Patient 2 (Condition 1) and genes in Patient 2 relative to the healthy donor (Condition 2). In Condition 1, we test each of NK cells, naïve CD4+ T-cells, effector CD4+ T-cells, naïve CD8+T-cells, and CD8+ NKT-like cells (Supplemental Tables S4-S9). In Condition 2, we test for CD4+ Effector T cells and CD8+ NKT-like cells to examine an exhaustive phenotype. For Condition 1, only genes present in at least 5% of both conditions and that showed at least a LFC of 0.5 were tested. Genes with a Benjamini-Hochberg false discovery rate (FDR) less than or equal to 0.1 were considered significantly differentially expressed. For Condition 2, only genes present in at least 10% of both conditions and that showed at least a LFC of 0.9 were

tested. Genes with a Benjamini-Hochberg FDR less than or equal to 0.01 were considered significantly differentially expressed.[29]

Finally, we ran Ingenuity Pathway Analysis (IPA) on the significantly differentially expressed genes for each cell type in Condition 1. Because CD8+ NKT cells had so many significant DE genes (1,416), within IPA we further filtered for genes with an FDR less than or equal to 0.01, yielding 641 “analysis ready” genes. Upstream regulators were identified by conducting a “Core Expression Analysis” on the LFC values. Default filters were used except for: 1) the Confidence filter included both “Experimentally Observed” and “High (predicted)” relationships and 2) the Species filter only included “Human” and “Uncategorized” molecules and relationships.

For cell-type specific analyses, IPA readouts on upstream regulators were sorted by "predicted activation state." Values flagged as “bias,” and molecule types “chemical,” “drug,” “complex” or “group” were removed. A total of 67 upstream regulators were enriched across the selected cell types.

URLs.
http://broadinstitute.github.io/picard
http://hpc.nih.gov
https://support.10xgenomics.com/single-cell-gene-expression/datasets/
https://support.10xgenomics.com/single-cell-gene-expression/software/release-notes/build#grch38_3.1.0

Code and Data availability
The authors declare that the data supporting the findings of this study are available within the paper. Single-cell raw UMI counts and fully processed data can be accessed on Zenodo (https://doi.org/10.5281/zenodo.7713993). All single-cell analyses and their respective codes (implemented in R and Jupyter Notebooks) are available in a public repository (https://github.com/hmbaghdassarian/STAT4_v2). Genome data for individual patients cannot be made publicly available for reasons of patient confidentiality. Researchers may apply for access to these data, pending approval of the individual Institutional Review Boards.

Figure S1. Additional clinical images and evaluations of patients. Hypopigmented "tank top" sign on the back **(a, P3)**, and loss of entire subcutis of legs with joint contractures of ankle, subtalar and toe **(b, P4).** MRI of pelvic girdle **(c, P4)** and extremities (femur **(d)**, knee **(e), P3**) reveals diffuse fasciitis, adjacent myositis, and subcutaneous calcifications. Full thickness skin biopsy **(f,g, P4)** shows thickened fascia with sclerosis and pockets of lymphoplasmacytic inflammation throughout **(f)** with adjacent muscle demonstrating lymphocytic infiltrate throughout the muscle bundles **(g).**

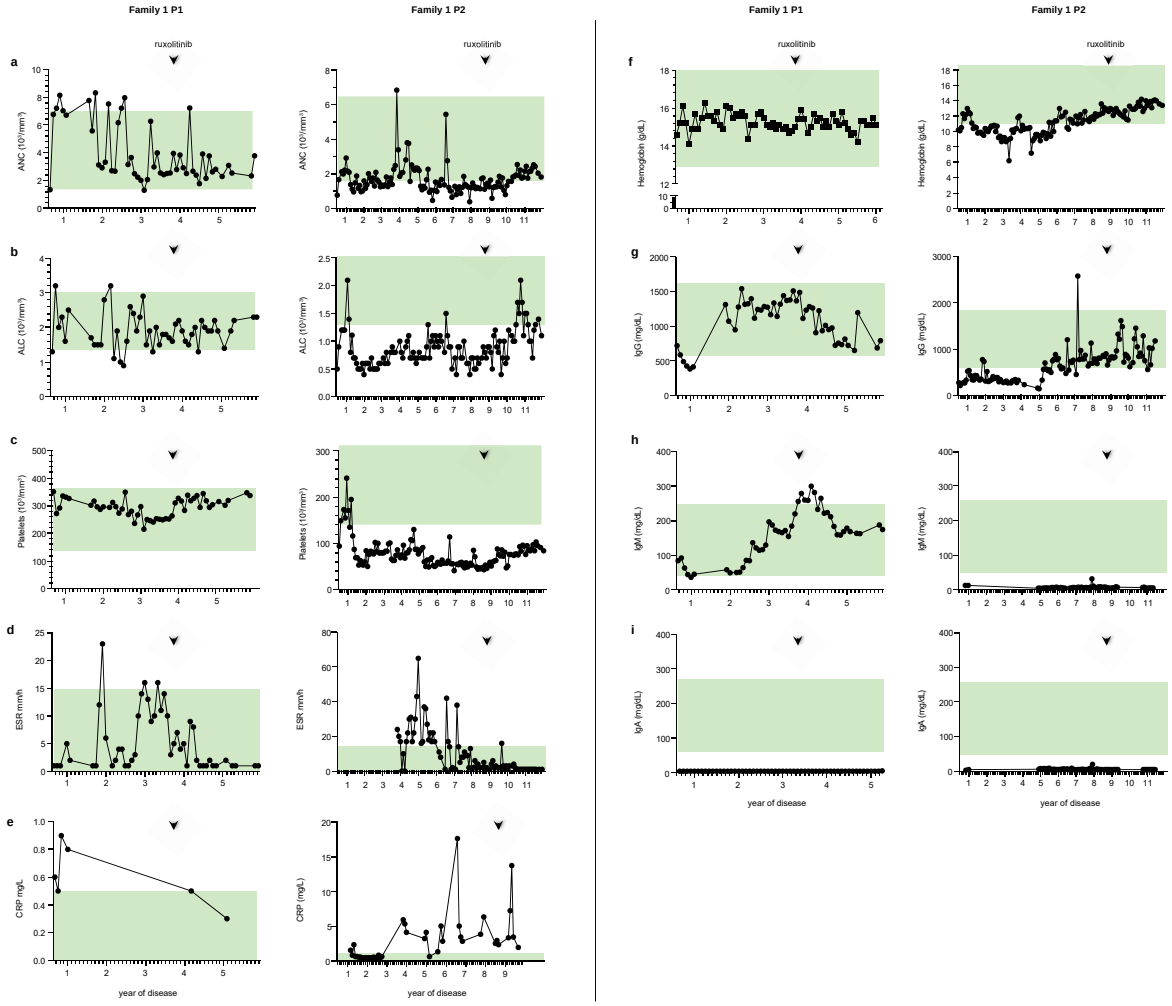


Figure S2. Summary of clinical labs for patients of Family 1. (a) Absolute neutrophil count (ANC), **(b)** absolute lymphocyte count (ALC), **(c)** platelet count, **(d)** erythrocyte sedimentation rate (ESR), **(e)** C-reactive protein levels (CRP), **(f)** hemoglobin, **(g)** IgG, **(h)** IgM, and **(i)** IgA levels over the course of disease. Normal ranges are shown in green; arrowheads indicate initiation of ruxolitinib therapy.

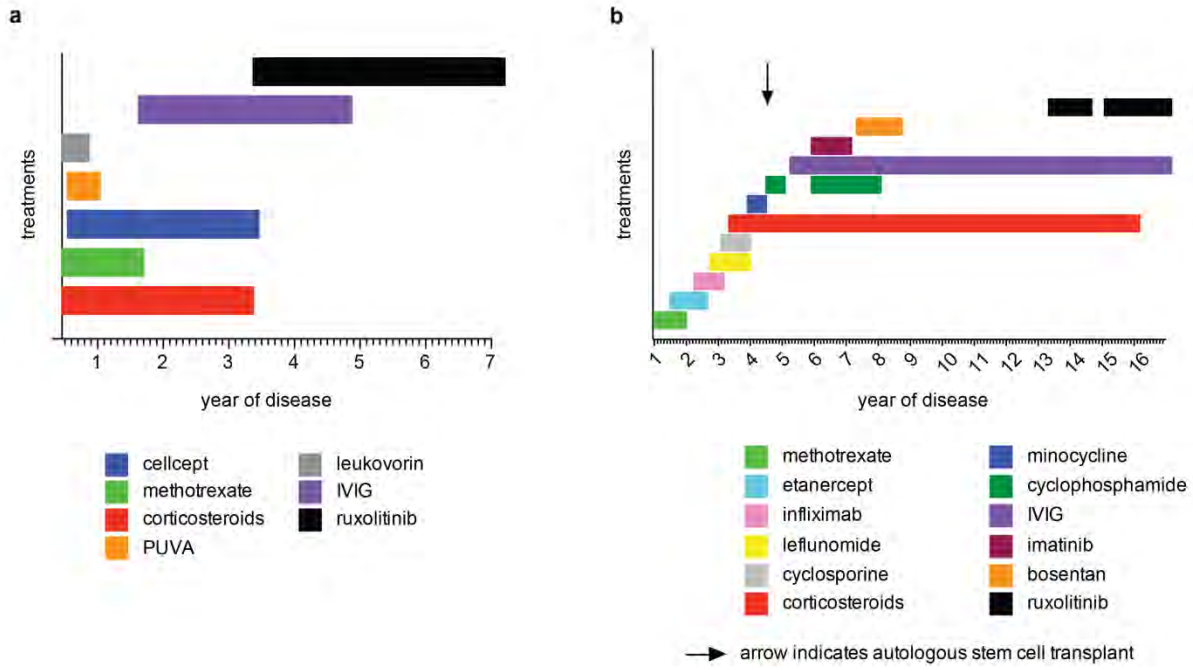


Figure S3. Treatment summary for Family 1. Timeline for immunomodulatory treatments for Patient 1 **(a)** and Patient 2 **(b)**.

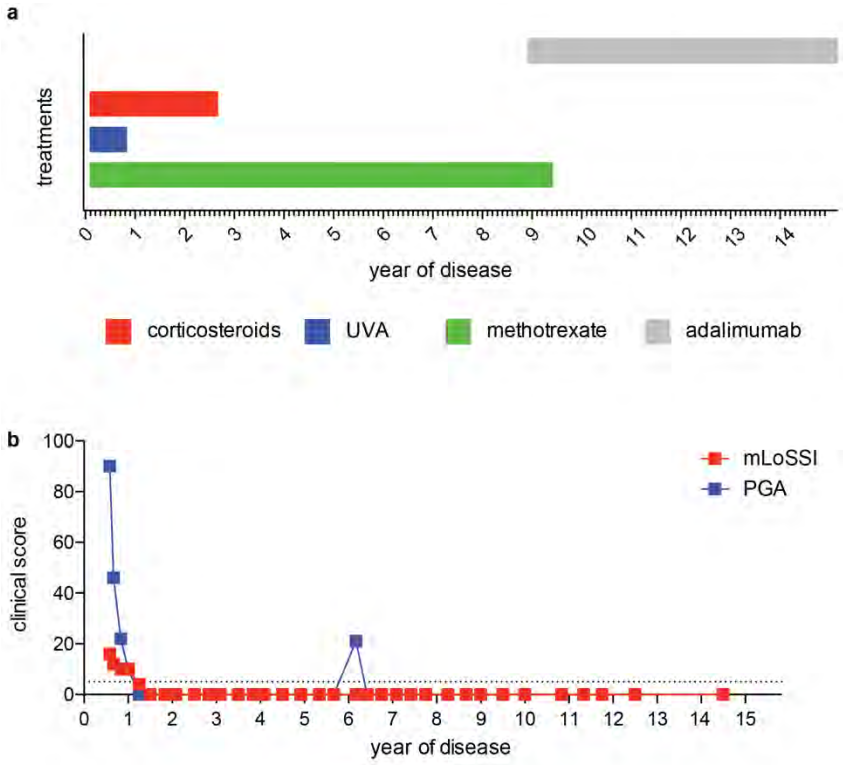


Figure S4. Treatment summary for Patient 4, Family 3. Timeline for immunomodulatory treatments for the family 3 proband **(a;** *UVA, ultraviolet A***)** and clinical scoring by the modified LS Skin Severity Index (mLoSSI, red) and Physician Global Assessment (PGA, blue) of Disease Activity. Dotted line represents absence of disease **(b)**.

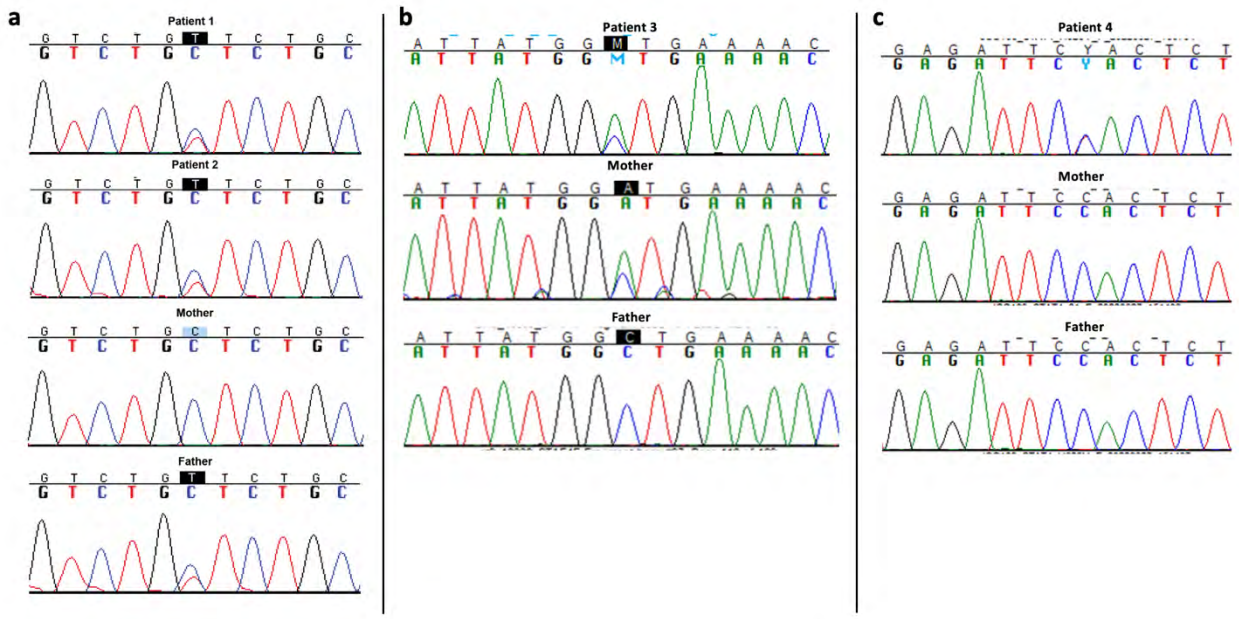
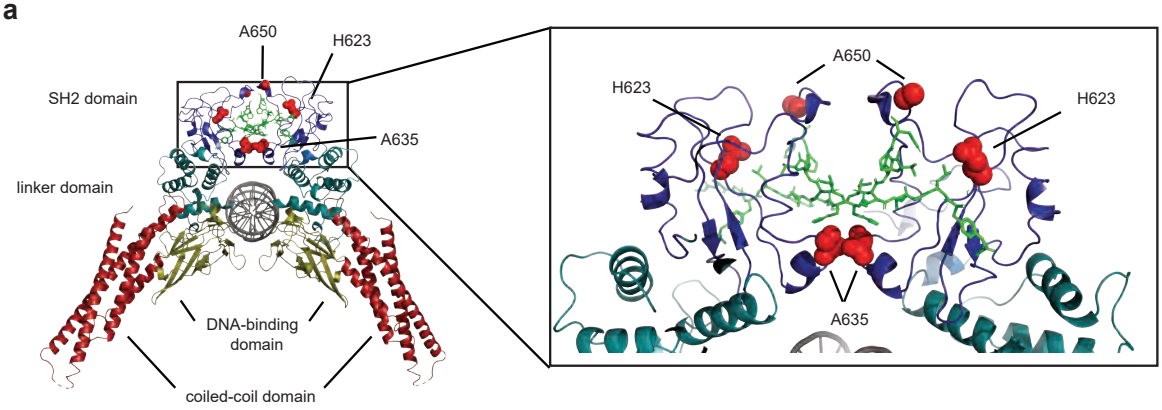
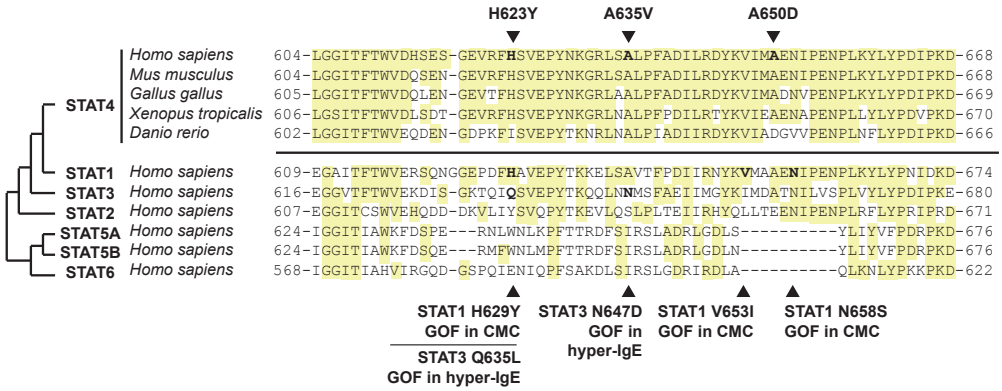


Figure S5. Sanger sequencing. (a) Chromatograms show confirmation of heterozygous *STAT4* mutation, c.1904 C>T, encoding A635V. **(b)** Confirmation of heterozygous *STAT4* mutation, c.1949 C>A, encoding A650D in Family 2. **(c)** Confirmation of *de novo* heterozygous *STAT4* mutation, c.1867 C>T, encoding H623Y in Family 3.

a



b



c

	Allele Frequencies						
Variant	**gnomADg**	**gnomAD_exomes**	**ExAC**	**1000Gp3**	**TopMed**	**AllofUs**	**deCAF**
STAT4:p.A650D	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STAT4:p.A635V	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STAT4:p.H623Y	0.0	0.0	0.0	0.0	0.0	0.0	0.0

d

Variant	**CADD _PHRED**	**BayesDel _addAF**	**VEST4**	**REVEL**	**FATHMM pred (score)**	**GERP++ _NR**	**SIFT pred (score)**	**PolyPhen2 pred (score)**
STAT4:p.A650D	22.2	0.04026	0.169	0.367	D (-3.87)	5.38	tolerated (0.41)	benign (0.001)
STAT4:p.A635V	31	0.335585	0.49	0.718	D (-4.03)	5.38	deleterious (0.03)	probably damaging (0.995)
STAT4:p.H623Y	26.6	0.176164	0.451	0.623	D (-3.99)	5.38	deleterious (0.01)	possibly damaging (0.752)
Deleterious cutoff	≥20	≥0.0692655	≥0.5	≥0.65	from output	≥0.047	from output	from output

Figure S6. Structure, conservation and predicted impact of amino acid substitutions caused by STAT4 patient mutations. (a) The structure of STAT4 modeled on STAT1 (PDB 1BF5; linker domain green, DNA binding domain yellow, coiled-coil domains red, and the A635, H623 and A650 residues (red spheres) proximal to the SH2 domain (cyan)). The box shows a

magnification of the phosphotyrosine binding pocket. **(b)** Sequences for *Homo sapiens* STAT1 (NP_009330.1), STAT2 (NP_005410.1), STAT3 (NP_644805.1), STAT4 (NP_003142.1), STAT5A (NP_003143.2), STAT5B (NP_036580.2), and STAT6 (NP_003144.3) were aligned with STAT4 sequences from *Mus musculus* (mouse, NP_035617.1), *Gallus gallus* (chicken, NP_001254484.2), *Xenopus tropicalis* (western clawed frog, XP_031749081.1), and *Danio rerio* (zebrafish, XP_005167937). Triangles at top indicate the positions of amino acids changed by patient mutations. Highlighted residues are identical to the *H. sapiens* STAT4 sequence. Numbers indicate amino acid positions. Alignments were performed with Clustal Omega.[31] Cladogram on left was derived from previous phylogenetic analyses.[32] Amino acid substitutions associated with a STAT1 gain-of-function and chronic mucocutaneous candidiasis (CMC) and STAT3 gain-of-function and hyper-IgE syndrome are indicated by triangles at bottom.[33-35] The amino acids at STAT1 H629 and STAT3 Q635 are homologous to STAT4 H623 and the STAT3 N647 amino acid is homologous to STAT4 A635. **(c)** The allele frequencies of the patient mutations in databases of single nucleotide variants reveal that these mutations are not present in the general population. **(d)** Eight programs were used to make *in silico* predictions for the effect of patient amino acid substitutions on STAT4 function (BayesDel,[36] CADD_phred,[37] FATHMM[38] GERP++,[39] PolyPhen2,[40] REVEL,[41] SIFT,[42] VEST[43]). Predictions indicating a potential or probable deleterious effect are highlighted in red. Cutoffs used for predicting deleterious effects for each program are reported at bottom.

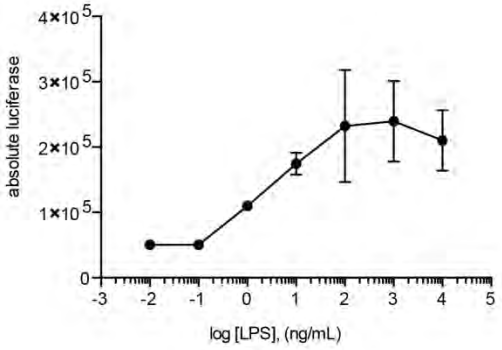


Figure S7. Induction of IL-6 promoter activity by LPS in IL-6 Leeporter™ cells. IL-6 Leeporter™ cells were seeded at 5 x 10^4 cells/well into a white solid-bottom 96-well microplate. Cells were stimulated with various concentrations of LPS for 16 hours, and luciferase activity measured. Shown as average ± SEM (n=2).

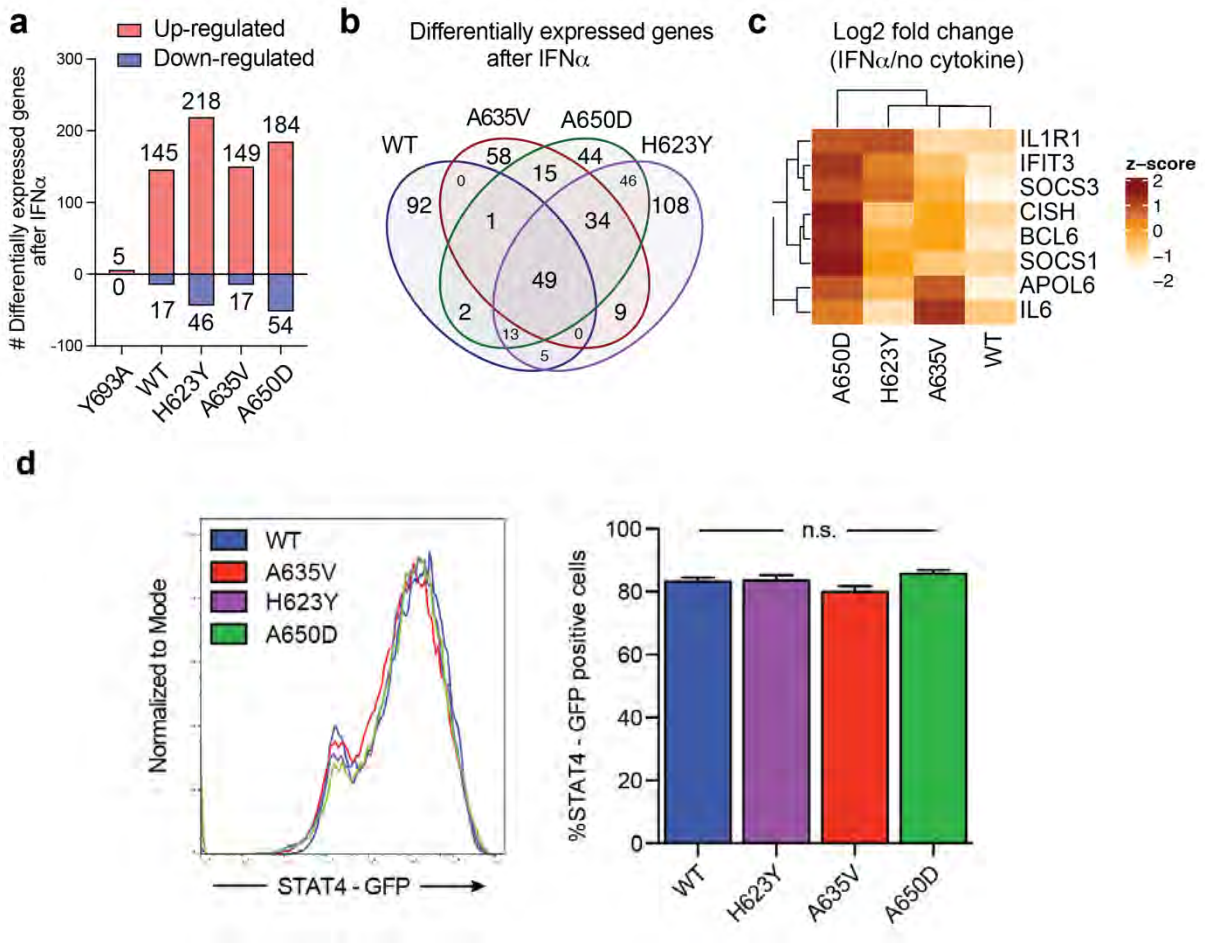


Figure S8. STAT4 variants induce a pro-inflammatory state.
(a) Number of differentially expressed genes measured by RNAseq between unstimulated and IFNα-stimulated U3A cells stably transfected with the indicated STAT4 variants (WT, H623Y, A635V, A650D, or the phospho-dead variant Y693A). Genes differentially expressed after IFNα stimulation were identified using the following criteria: log2 fold change > 0.5, FDR < 0.1, and logCPM > 0.3. (**b)** Venn Diagram showing the number of differentially expressed genes similar and different between each condition. **(c)** Log2 fold change expression of top induced genes with IFNα stimulation. **(d)** Baseline expression of total STAT4 does not vary among unstimulated U3A cells stably transfected with the indicated STAT4 variants (H623Y, A635V, A650D) compared to wildtype (WT) (n=3 per cell line; n.s., not significant).

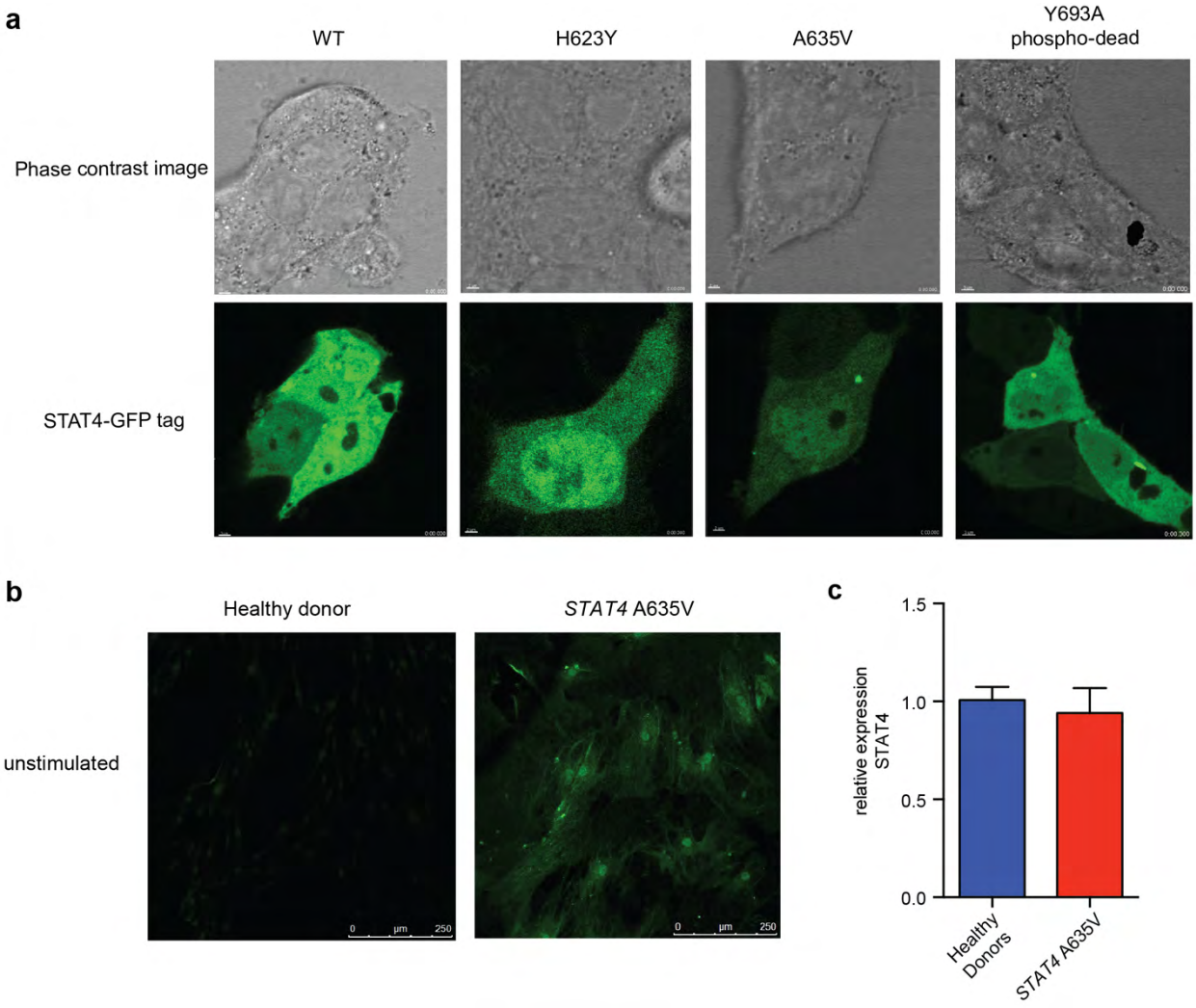


Figure S9. Nuclear localization of variant STAT4. (a) HEK 293T cells transiently transfected with plasmids containing wildtype H623Y, A635V or phospho-dead Y693A STAT4 tagged with GFP. Unstimulated cells transfected with H623Y or A635V variants show a greater accumulation of STAT4 in the nucleus compared to wildtype and Y693A cells in single fluorescence channel images. **(b)** Single channel immunofluorescence imaging of primary skin fibroblasts from patient have prominent pSTAT4 (green) which is not observed in healthy donor fibroblasts at baseline. **(c)** Relative expression of STAT4 mRNA by qPCR in healthy donor and patient derived skin fibroblasts.

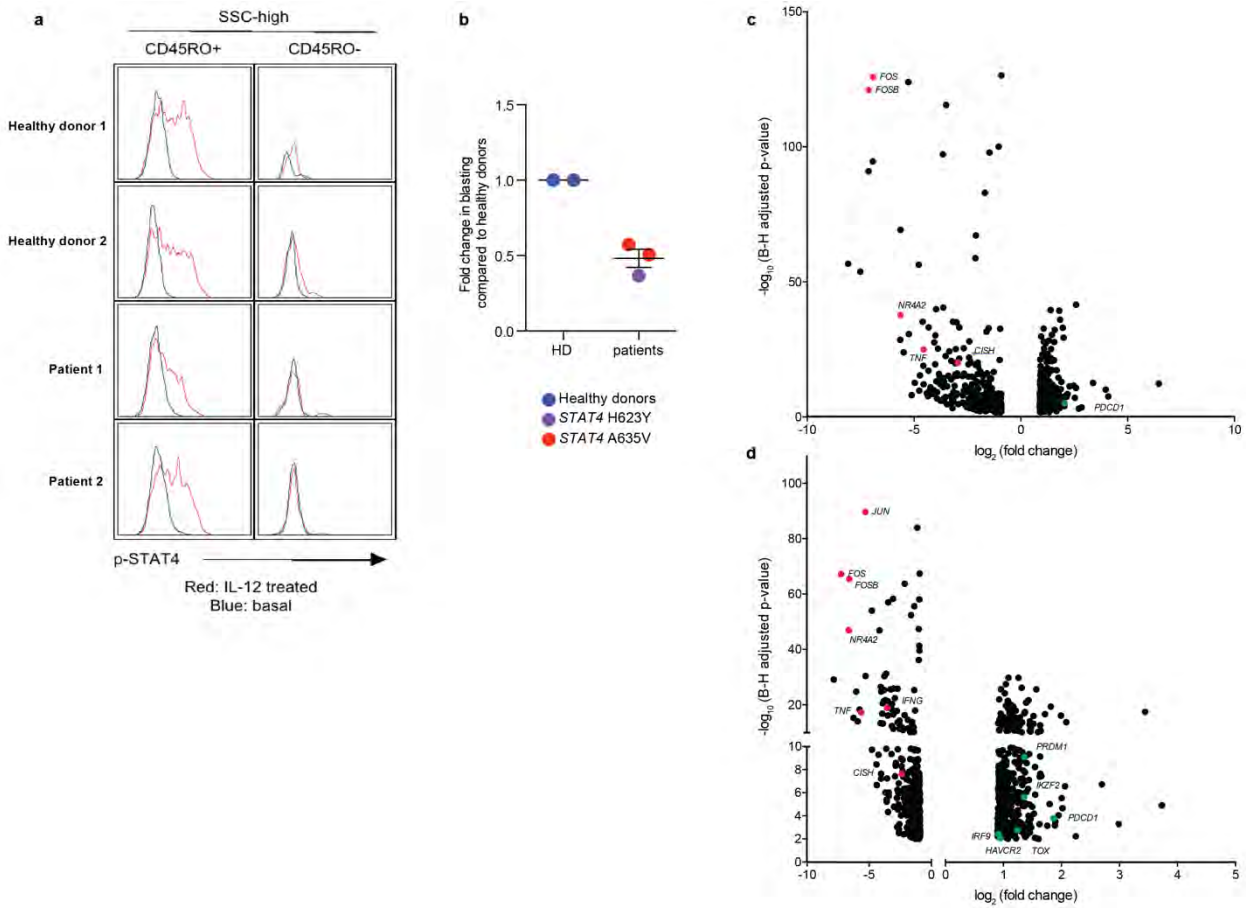


Figure S10. pSTAT4 activation and effects on T cells in peripheral blood. **(a)** Flow cytometry of CD3+CD4+ CD45RO+ or – cells from 2 healthy donors compared to Family 1 P1 and P2. pSTAT4 measured in response to PHA and IL-12 stimulation. **(b)** PHA-induced blasting was reduced in peripheral blood T cells isolated from patients carrying either the H623Y or A635V variants, compared to healthy donor T cells. **(c,d)** Differential expression analysis of CD4 **(c)** and CD8 T cells **(d)** from untreated patient compared to control, demonstrating upregulation of genes associated with T cell exhaustion shown in green and downregulation of *JUN, FOS, FOSB, NR4A2, CISH, TNF* and *IFNG*, shown in red. Non-significant genes with low fold changes are not shown. *STAT4* was not differentially expressed among T cell subsets between patient and control samples.

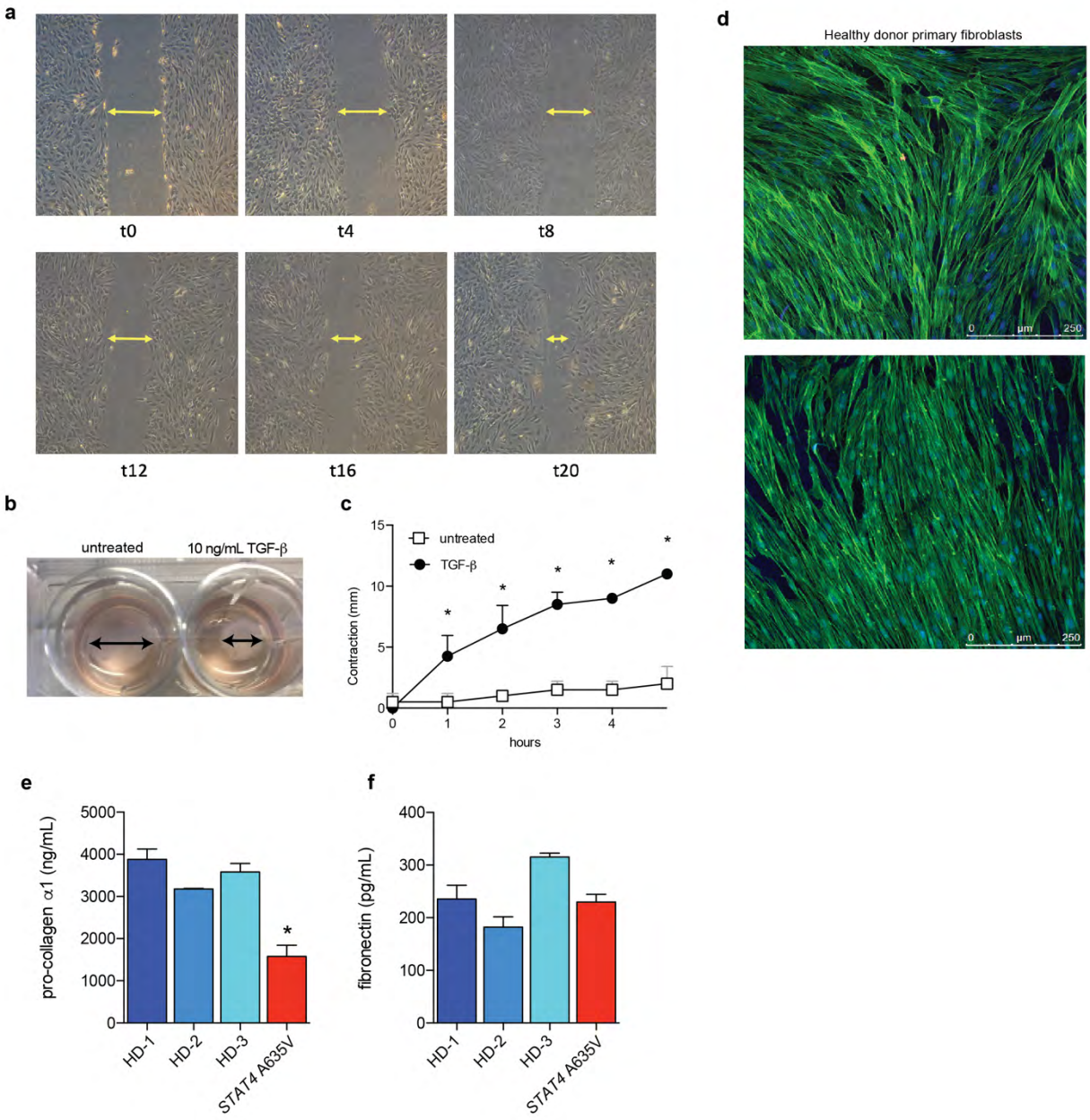


Figure S11. Evaluation of fibroblast function in vitro. (a) Representative wound closure ("scratch") assay. Confluent monolayers are scratched with a 200-µl pipette tip and diameter of scratch measured every 4-6 hours until closure. **(b,c)** TGF-β induces contraction in healthy donor control skin fibroblasts in a collagen matrix. Primary skin fibroblasts were embedded in a type I bovine collagen matrix, and incubated for 2 days with or without 10ng/mL TGF-β. After release, collagen gel diameter (arrow) was measured hourly. Representative collagen gel disks with contraction in the presence of TGF-β, compared to the untreated disk is shown in **(b)**. (*, P values between 0.01 and 0.05; n=4). **(d)** F-actin immunocytochemistry of 2 additional healthy donors. **(e,f)** Secretion of pro-collagen $\alpha 1$ **(e)** and fibronectin **(f)** at baseline in cell cultures from patient or healthy donors (n=3).

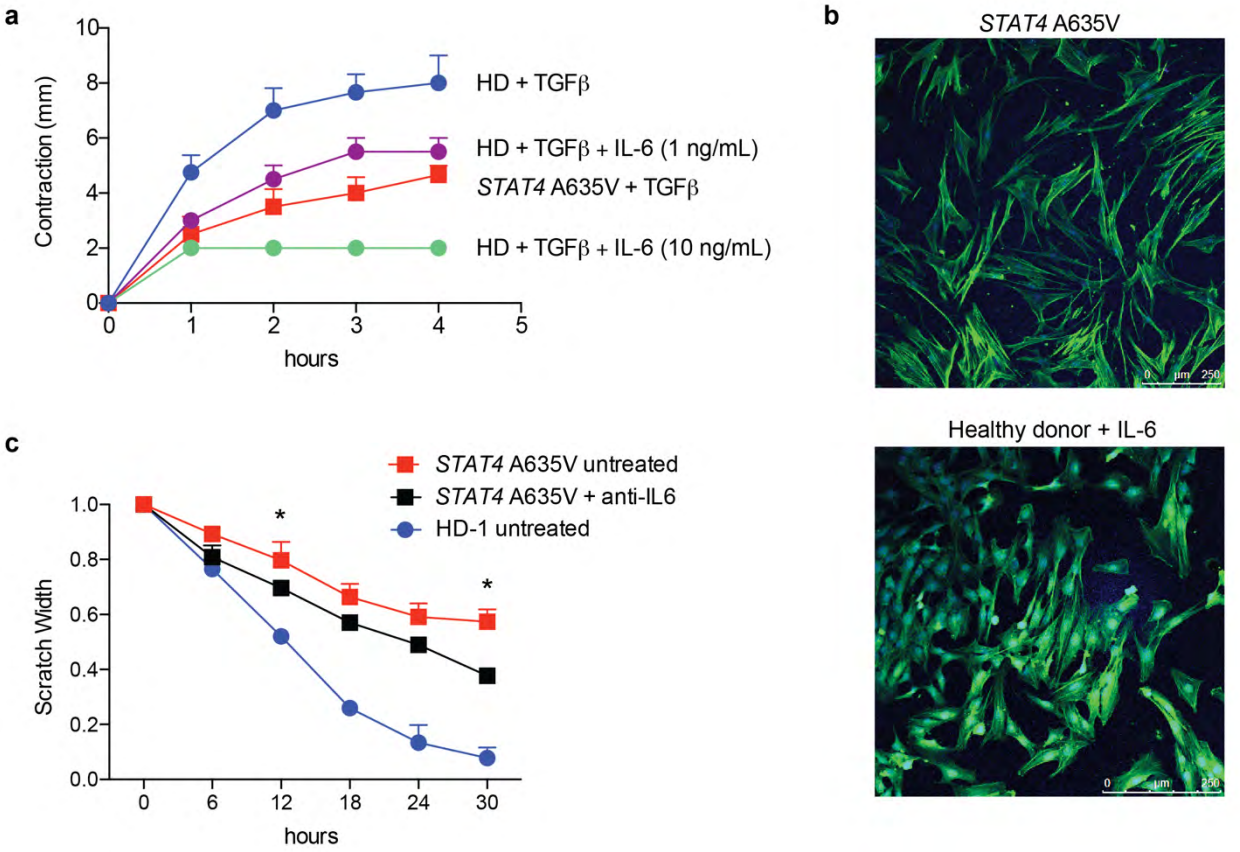


Figure S12. Role of IL-6 in fibroblast inflammation. (a) TGF-β -induced contraction of collagen matrix by healthy donor derived fibroblasts is reduced by IL-6 in a dose-dependent fashion (n=4). **(b)** Treatment of healthy donor fibroblasts with IL-6 (10 ng/mL) leads to enlarged cells (10X, images representative of 2 independent experiments). **(c)** Pre-treatment with anti-IL-6 leads to improved fibroblast migration in wound healing assays. (n=3 experiments, 6 scratches each *, P value between 0.01 and 0.05; by 2-way ANOVA).

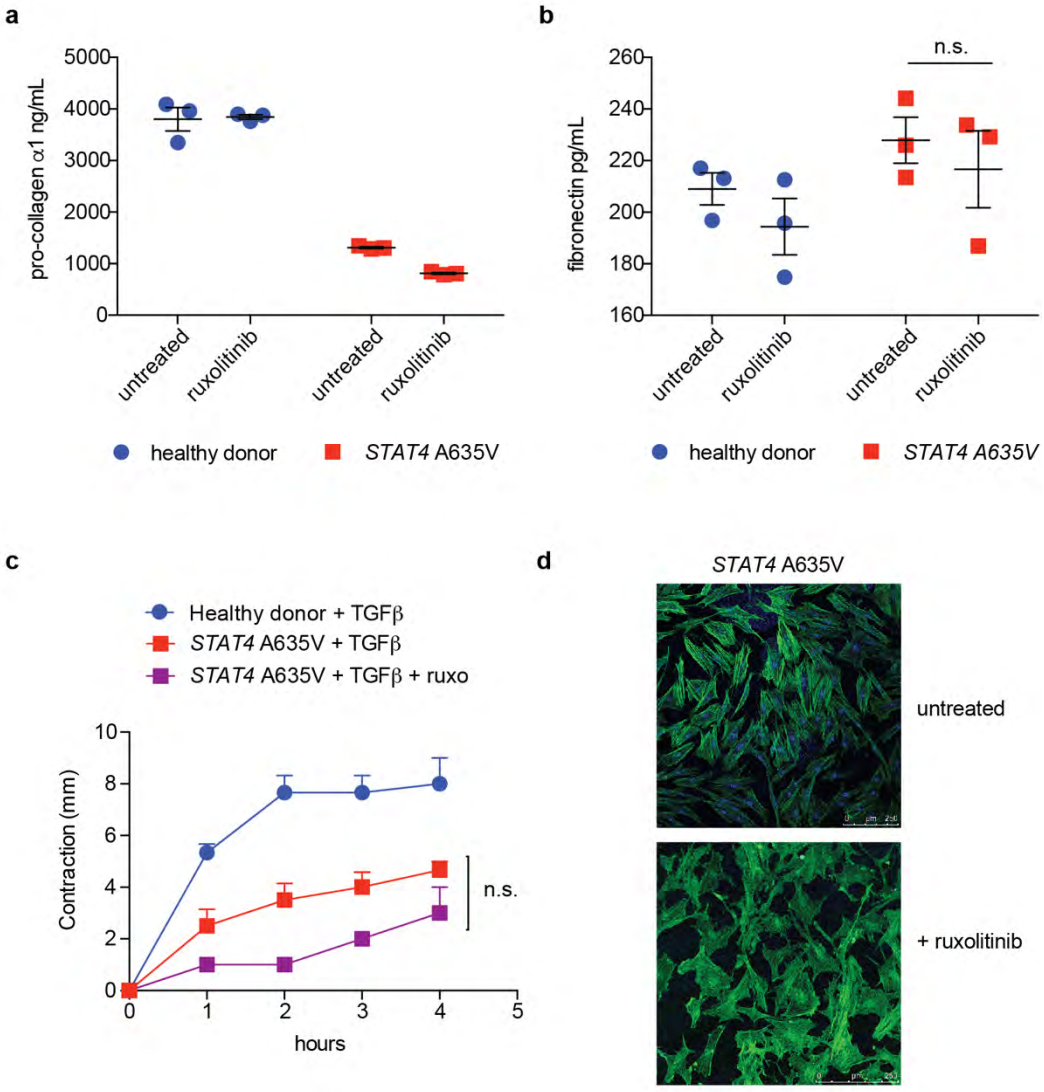


Figure S13. Ruxolitinib treatment of fibroblasts in vitro. (a,b) ELISA of primary skin fibroblast supernatants shows similar secretion of pro-collagen α1 **(a)** and fibronectin **(b)** at baseline, and when treated with ruxolitinib in both patient and normal donor samples (n = 3). **(c)** Collagen contraction induced by TGF-β remains impaired despite treatment with ruxolitinib, compared to healthy donor fibroblasts (n=3). **(d)** F-actin immunocytochemistry shows disorganized distribution and enhanced stress fibers in patient primary skin fibroblasts compared to healthy donor that is not dramatically improved with ruxolitinib treatment. (10X, images representative of 2 independent experiments)

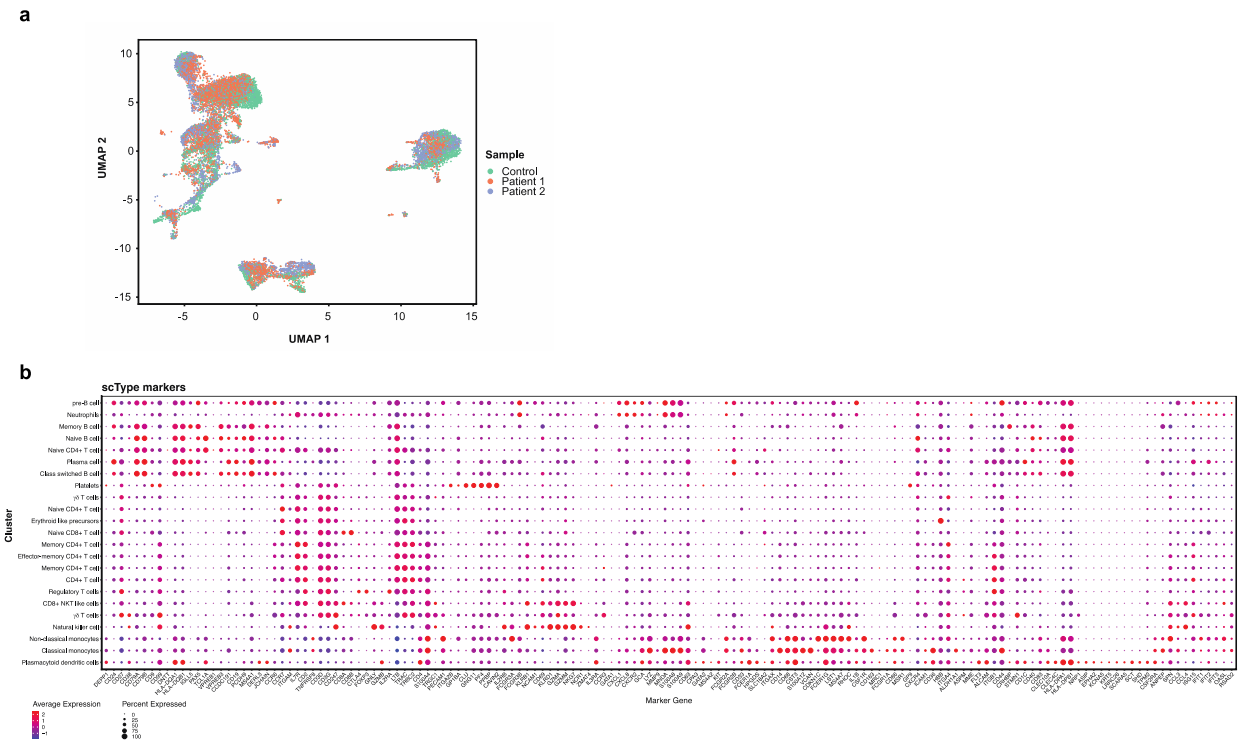


Figure S14. Cell type identification and integration of the scRNA-seq datasets. (a) Cells plotted in UMAP space after running the dimensionality reduction on 47 PCs for the integrated patient dataset, colored by sample. **(b)** Dotplot of canonical PBMC markers (x-axis). Clusters (y-axis, left-hand side) were labeled as cell types based on expression. Markers visualized are taken from those expected to be upregulated in immune tissue according to ScType's database. Dot size reflects the fraction of cells in the cluster expressing the marker, and color reflects the average expression of the marker. Expression values are taken from the integrated, scaled expression matrix.

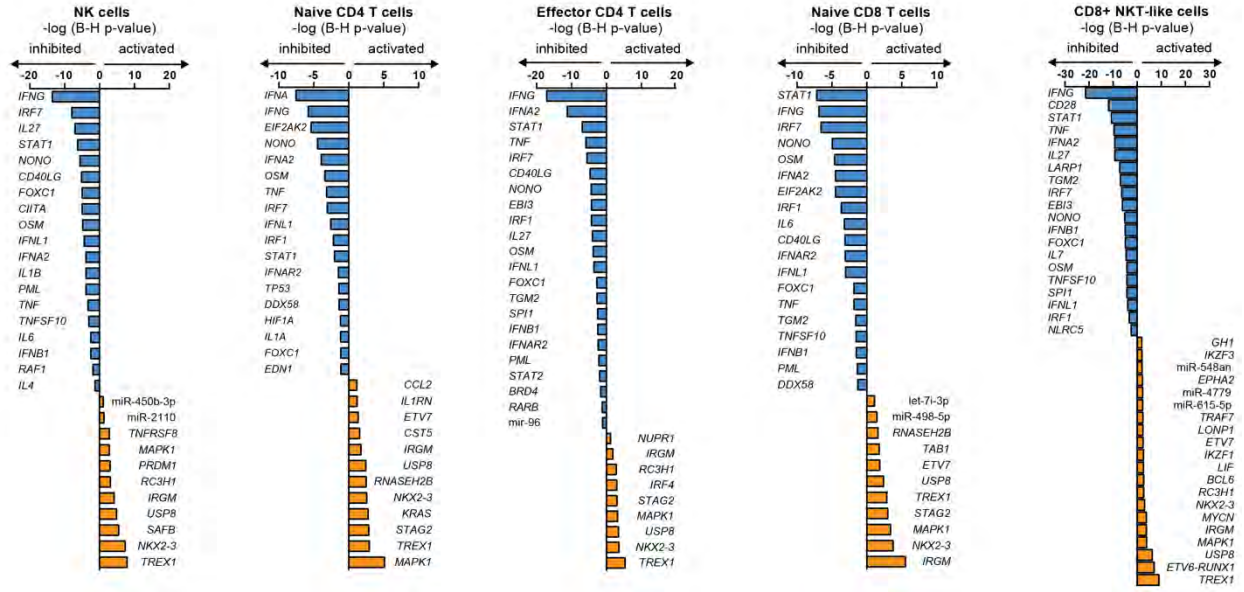


Figure S15. Upstream regulators identified by Ingenuity Pathway Analysis (IPA). IPA analysis of NK cell, CD4+ T cell and CD8+ T cell clusters. Differentially expressed genes used as input to IPA were identified by comparing expression levels in Patient 1 relative to Patient 2 in each cell type. Upstream regulators that are activated are shown in orange, and those that are inhibited are shown in blue.

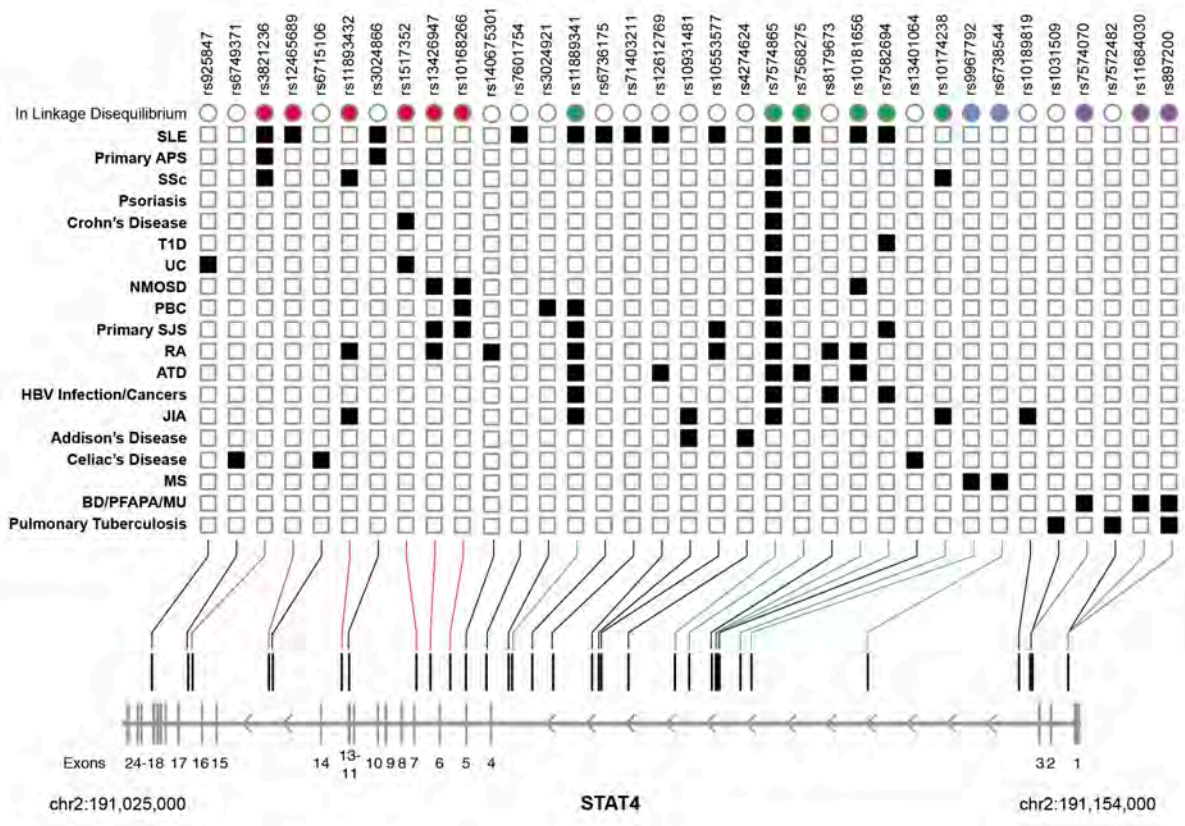


Figure S16. SNPs in *STAT4* associated with immune disease by GWAS studies. Top. Filled in squares indicate an association between the SNP (columns) and the disease (rows). Colored circles (red, green, blue, purple) indicate groups of SNPs that are in linkage disequilibrium with each other (r^2>0.8). **Bottom.** Positions of the SNPs within the *STAT4* gene are shown. Coordinates are from the GRCh38 build of the human genome. APS: Antiphospholipid Syndrome; ATD: Autoimmune Thyroid Disease; BD: Behçet disease; JIA: Juvenile Idiopathic Arthritis; HBV: Hepatitis B Virus; MS: Multiple Sclerosis; MU: Mouth Ulcers; NMOSD: Neuromyelitis Optica Spectrum Disorder; PBC: Primary Biliary Cirrhosis or Cholangitis; PFAPA: Periodic fever, aphthous stomatitis, pharyngitis, and cervical adenitis syndrome; RA: Rheumatoid Arthritis; SjS: Sjögren's Syndrome; SLE: Systemic Lupus Erythematous; SSc: Systemic Sclerosis; T1D: Type I Diabetes; UC: Ulcerative Colitis.[44-88]

Table S1: Clinical phenotype of patients with disabling pansclerotic morphea

Characteristic	Subjects (n = 4)	Reference Ranges
Demographics		
Male sex – no. (%)	4 (100)	
Median age at onset (range, yr)	3 (0.75 - 5 yrs)	
Key Features		
Skin sclerosis – no. (%)	4 (100)	
Skin/mucosal ulceration – no. (%)	4 (100)	
Muscular atrophy – no. (%)	3 (75)	
Joint contractures – no. (%)	3 (75)	
Squamous cell carcinoma – no. (%)	1 (25)	
Recurrent infections – no. (%)	2 (50)	
Laboratory findings*		
Median ANC (range) – x 10^3/ul	1.75 (0.76 - 2.79)	1000 – 6500
Median ALC (range) – x 10^3/ul	1.26 (0.5 - 2.82)	1200 – 3000
Median Platelets (range) – x 10^3/ul	278 (94 - 351)	135 – 380
Median IgG (range) – mg/dL	304 (272 - 721)	345 – 1236
Median IgA (range) – mg/dL	6.5 (<5 – 8)	14 – 154
Median IgM (range) – mg/dL	65 (13 – 84)	41 – 200
Median CRP (range) – mg/L	11 (<0.5 – 46.2)	< 0.5
Median ESR (range) – mm/h	10 (1 – 24)	<15
Histologic findings*		
Dermal thickening – no. (%)	4 (100)	
Hyalinization – no. (%)	4 (100)	
Inflammatory infiltrate – no. (%)	4 (100)	

**at initial evaluation. Range indicates observed patient values.*

Supplemental References

1. Farnaes L, Hildreth A, Sweeney NM, et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. NPJ Genom Med 2018;3:10.
2. Milko LV, Chen F, Chan K, et al. FDA oversight of NSIGHT genomic research: the need for an integrated systems approach to regulation. NPJ Genom Med 2019;4:32.
3. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 2017;45:D840-D5.
4. Yang H, Robinson PN, Wang K. Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. Nat Methods 2015;12:841-3.
5. Kohler S, Vasilevsky NA, Engelstad M, et al. The Human Phenotype Ontology in 2017. Nucleic Acids Res 2017;45:D865-D76.
6. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 2016;32:1220-2.
7. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res 2011;21:974-84.
8. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.
9. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010;26:589-95.
10. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297-303.
11. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
12. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596:583-9.
13. Varadi M, Anyango S, Deshpande M, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 2022;50:D439-D44.
14. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605-12.
15. Chen X, Vinkemeier U, Zhao Y, Jeruzalmi D, Darnell JE, Jr., Kuriyan J. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. Cell 1998;93:827-39.
16. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66:486-501.
17. Land H, Humble MS. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. Methods Mol Biol 2018;1685:43-67.
18. Banks JL, Beard HS, Cao Y, et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). J Comput Chem 2005;26:1752-80.
19. Vangipuram M, Ting D, Kim S, Diaz R, Schule B. Skin punch biopsy explant culture for derivation of primary human fibroblasts. J Vis Exp 2013:e3779.
20. Muller M, Laxton C, Briscoe J, et al. Complementation of a mutant cell line: central role of the 91 kDa polypeptide of ISGF3 in the interferon-alpha and -gamma signal transduction pathways. EMBO J 1993;12:4221-8.
21. Torpey N, Maher SE, Bothwell AL, Pober JS. Interferon alpha but not interleukin 12 activates STAT4 signaling in human vascular endothelial cells. J Biol Chem 2004;279:26789-96.

22. Kim SI, Oceguera-Yanez F, Sakurai C, Nakagawa M, Yamanaka S, Woltjen K. Inducible Transgene Expression in Human iPS Cells Using Versatile All-in-One piggyBac Transposons. Methods Mol Biol 2016;1357:111-31.
23. Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun 2017;8:14049.
24. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184:3573-87 e29.
25. Ilicic T, Kim JK, Kolodziejczyk AA, et al. Classification of low quality cells from single-cell RNA-seq data. Genome Biol 2016;17:29.
26. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol 2019;15:e8746.
27. Buttner M, Ostner J, Muller CL, Theis FJ, Schubert B. scCODA is a Bayesian model for compositional single-cell data analysis. Nat Commun 2021;12:6876.
28. Mathys H, Davila-Velderrain J, Peng Z, et al. Single-cell transcriptomic analysis of Alzheimer's disease. Nature 2019;570:332-7.
29. Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. Nat Commun 2022;13:1246.
30. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 2015;16:278.
31. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011;7:539.
32. Gorissen M, de Vrieze E, Flik G, Huising MO. STAT genes display differential evolutionary rates that correlate with their roles in the endocrine and immune system. J Endocrinol 2011;209:175-84.
33. Meesilpavikkai K, Dik WA, Schrijver B, et al. A Novel Heterozygous Mutation in the STAT1 SH2 Domain Causes Chronic Mucocutaneous Candidiasis, Atypically Diverse Infections, Autoimmunity, and Impaired Cytokine Regulation. Front Immunol 2017;8:274.
34. Ovadia A, Sharfe N, Hawkins C, Laughlin S, Roifman CM. Two different STAT1 gain-of-function mutations lead to diverse IFN-gamma-mediated gene expression. NPJ Genom Med 2018;3:23.
35. Sobh A, Chou J, Schneider L, Geha RS, Massaad MJ. Chronic mucocutaneous candidiasis associated with an SH2 domain gain-of-function mutation that enhances STAT1 phosphorylation. J Allergy Clin Immunol 2016;138:297-9.
36. Feng BJ. PERCH: A Unified Framework for Disease Gene Prioritization. Hum Mutat 2017;38:243-51.
37. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 2019;47:D886-D94.
38. Shihab HA, Gough J, Cooper DN, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 2013;34:57-65.
39. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol 2010;6:e1001025.
40. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.
41. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;99:877-85.

42. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863-74.
43. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics 2013;14 Suppl 3:S3.
44. Abelson AK, Delgado-Vega AM, Kozyrev SV, et al. STAT4 associates with systemic lupus erythematosus through two independent effects that correlate with gene expression and act additively with IRF5 to increase risk. Ann Rheum Dis 2009;68:1746-53.
45. Aiba Y, Yamazaki K, Nishida N, et al. Disease susceptibility genes shared by primary biliary cirrhosis and Crohn's disease in the Japanese population. J Hum Genet 2015;60:525-31.
46. Bi C, Li B, Cheng Z, Hu Y, Fang Z, Zhai A. Association study of STAT4 polymorphisms and type 1 diabetes in Northeastern Chinese Han population. Tissue Antigens 2013;81:137-40.
47. Clark A, Gerlach F, Tong H, et al. A trivial role of STAT4 variant in chronic hepatitis B induced hepatocellular carcinoma. Infect Genet Evol 2013;18:257-61.
48. Diaz-Gallo LM, Palomino-Morales RJ, Gomez-Garcia M, et al. STAT4 gene influences genetic predisposition to ulcerative colitis but not Crohn's disease in the Spanish population: a replication study. Hum Immunol 2010;71:515-9.
49. Dieude P, Guedj M, Wipff J, et al. STAT4 is a genetic risk factor for systemic sclerosis having additive effects with IRF5 on disease susceptibility and related pulmonary fibrosis. Arthritis Rheum 2009;60:2472-9.
50. Fan ZD, Wang FF, Huang H, et al. STAT4 rs7574865 G/T and PTPN22 rs2488457 G/C polymorphisms influence the risk of developing juvenile idiopathic arthritis in Han Chinese patients. PLoS One 2015;10:e0117389.
51. Glas J, Seiderer J, Nagy M, et al. Evidence for STAT4 as a common autoimmune gene: rs7574865 is associated with colonic Crohn's disease and early disease onset. PLoS One 2010;5:e10373.
52. Han JW, Zheng HF, Cui Y, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 2009;41:1234-7.
53. Hou S, Yang Z, Du L, et al. Identification of a susceptibility locus in STAT4 for Behcet's disease in Han Chinese in a genome-wide association study. Arthritis Rheum 2012;64:4104-13.
54. Huang X, Wang Z, Jia N, et al. Association between STAT4 polymorphisms and the risk of juvenile idiopathic arthritis in Han Chinese populations. Clin Exp Rheumatol 2019;37:333-7.
55. Jiang DK, Sun J, Cao G, et al. Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus-related hepatocellular carcinoma. Nat Genet 2013;45:72-5.
56. Kim LH, Cheong HS, Namgoong S, et al. Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci of STAT4 and HLA-DQ in a Korean population. Infect Genet Evol 2015;33:72-6.
57. Kobayashi S, Ikari K, Kaneko H, et al. Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. Arthritis Rheum 2008;58:1940-6.
58. Korman BD, Alba MI, Le JM, et al. Variant form of STAT4 is associated with primary Sjogren's syndrome. Genes Immun 2008;9:267-70.
59. Lee HS, Park H, Yang S, Kim D, Park Y. STAT4 polymorphism is associated with early-onset type 1 diabetes, but not with late-onset type 1 diabetes. Ann N Y Acad Sci 2008;1150:93-8.
60. Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 2007;13:455-60.
61. Lessard CJ, Li H, Adrianto I, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome. Nat Genet 2013;45:1284-92.

62. Li Y, Zhang K, Chen H, et al. A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjogren's syndrome at 7q11.23. Nat Genet 2013;45:1361-5.
63. Lu Y, Zhu Y, Peng J, Wang X, Wang F, Sun Z. STAT4 genetic polymorphisms association with spontaneous clearance of hepatitis B virus infection. Immunol Res 2015;62:146-52.
64. Manthiram K, Preite S, Dedeoglu F, et al. Common genetic susceptibility loci link PFAPA syndrome, Behcet's disease, and recurrent aphthous stomatitis. Proc Natl Acad Sci U S A 2020;117:14405-11.
65. Mirkazemi S, Akbarian M, Jamshidi AR, et al. Association of STAT4 rs7574865 with susceptibility to systemic lupus erythematosus in Iranian population. Inflammation 2013;36:1548-52.
66. Mitchell AL, Macarthur KD, Gan EH, et al. Association of autoimmune Addison's disease with alleles of STAT4 and GATA3 in European cohorts. PLoS One 2014;9:e88991.
67. Moon CM, Cheon JH, Kim SW, et al. Association of signal transducer and activator of transcription 4 genetic variants with extra-intestinal manifestations in inflammatory bowel disease. Life Sci 2010;86:661-7.
68. Nordmark G, Kristjansdottir G, Theander E, et al. Association of EBF1, FAM167A(C8orf13)-BLK and TNFSF4 gene variants with primary Sjogren's syndrome. Genes Immun 2011;12:100-9.
69. Orozco G, Alizadeh BZ, Delgado-Vega AM, et al. Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. Arthritis Rheum 2008;58:1974-80.
70. Palomino-Morales RJ, Rojas-Villarraga A, Gonzalez CI, Ramirez G, Anaya JM, Martin J. STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. Genes Immun 2008;9:379-82.
71. Piotrowski P, Lianeri M, Wudarski M, Olesinska M, Jagodzinski PP. Contribution of STAT4 gene single-nucleotide polymorphism to systemic lupus erythematosus in the Polish population. Mol Biol Rep 2012;39:8861-6.
72. Prahalad S, Hansen S, Whiting A, et al. Variants in TNFAIP3, STAT4, and C12orf30 loci associated with multiple autoimmune diseases are also associated with juvenile idiopathic arthritis. Arthritis Rheum 2009;60:2124-30.
73. Remmers EF, Plenge RM, Lee AT, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357:977-86.
74. Rueda B, Broen J, Simeon C, et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet 2009;18:2071-7.
75. Sabri A, Grant AV, Cosker K, et al. Association study of genes controlling IL-12-dependent IFN-gamma immunity: STAT4 alleles increase risk of pulmonary tuberculosis in Morocco. J Infect Dis 2014;210:611-8.
76. Saevarsdottir S, Stefansdottir L, Sulem P, et al. Multiomics analysis of rheumatoid arthritis yields sequence variants that have large effects on risk of the seropositive subset. Ann Rheum Dis 2022;81:1085-95.
77. Shi Z, Zhang Q, Chen H, et al. STAT4 Polymorphisms are Associated with Neuromyelitis Optica Spectrum Disorders. Neuromolecular Med 2017;19:493-500.
78. Sigurdsson S, Nordmark G, Garnier S, et al. A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5. Hum Mol Genet 2008;17:2868-76.
79. Stock CJW, De Lauretis A, Visca D, et al. Defining genetic risk factors for scleroderma-associated interstitial lung disease : IRF5 and STAT4 gene variants are associated with scleroderma while STAT4 is protective against scleroderma-associated interstitial lung disease. Clin Rheumatol 2020;39:1173-9.
80. Tsuchiya N, Kawasaki A, Hasegawa M, et al. Association of STAT4 polymorphism with systemic sclerosis in a Japanese population. Ann Rheum Dis 2009;68:1375-6.

81. Xu L, Dai WQ, Wang F, et al. Association of STAT4 gene rs7574865G > T polymorphism with ulcerative colitis risk: evidence from 1532 cases and 3786 controls. Arch Med Sci 2014;10:419-24.
82. Yan N, Meng S, Zhou J, et al. Association between STAT4 gene polymorphisms and autoimmune thyroid diseases in a Chinese population. Int J Mol Sci 2014;15:12280-93.
83. Yi J, Fang X, Wan Y, Wei J, Huang J. STAT4 polymorphisms and diabetes risk: a meta-analysis with 18931 patients and 23833 controls. Int J Clin Exp Med 2015;8:3566-72.
84. Yi L, Wang JC, Guo XJ, et al. STAT4 is a genetic risk factor for systemic sclerosis in a Chinese population. Int J Immunopathol Pharmacol 2013;26:473-8.
85. Yin H, Borghi MO, Delgado-Vega AM, Tincani A, Meroni PL, Alarcon-Riquelme ME. Association of STAT4 and BLK, but not BANK1 or IRF5, with primary antiphospholipid syndrome. Arthritis Rheum 2009;60:2468-71.
86. Zervou MI, Goulielmos GN, Castro-Giner F, Tosca AD, Krueger-Krasagakis S. STAT4 gene polymorphism is associated with psoriasis in the genetically homogeneous population of Crete, Greece. Hum Immunol 2009;70:738-41.
87. Zervou MI, Mamoulakis D, Panierakis C, Boumpas DT, Goulielmos GN. STAT4: a risk factor for type 1 diabetes? Hum Immunol 2008;69:647-50.
88. Zhao X, Jiang K, Liang B, Huang X. STAT4 gene polymorphism and risk of chronic hepatitis B-induced hepatocellular carcinoma. Cell Biochem Biophys 2015;71:353-7.