www.sciencemag.org/cgi/content/full/science.1246886/DC1

Supplementary Materials for

Association of the Autoimmune Disease Scleroderma with an Immunologic Response to Cancer

Christine G. Joseph, Erika Darrah, Ami A. Shah, Andrew D. Skora, Livia A. Casciola-Rosen, Fredrick M. Wigley, Francesco Boin, Andrea Fava, Chris Thoburn, Isaac Kinde, Yuchen Jiao, Nickolas Papadopoulos, Kenneth W. Kinzler,* Bert Vogelstein,* Antony Rosen*

*Corresponding author. E-mail: kinzlke@jhmi.edu (K.W.K.); vogelbe@jhmi.edu (B.V.); arosen@jhmi.edu (A.R.)

Published 5 December 2013 on *Science* Express
DOI: 10.1126/science.1246886

This PDF file includes:

Supplementary Materials:

Materials and Methods
Brief clinical descriptions of patients with anti-RPC1 antibodies
Figures S1 to S5
Tables S1 to S6
Supplementary references (38-44)

Materials and Methods

Clinical methods

Consenting scleroderma patients with confirmed cancer diagnoses were recruited from the Johns Hopkins Scleroderma Center. Scleroderma patients met the American College of Rheumatology criteria for scleroderma (*38*). Existing cancer pathology specimens were obtained from prior surgical procedures performed as part of routine clinical care. The closest serum sample to cancer diagnosis was studied in all patients, and DNA and PBMC samples were obtained in consenting participants. The Johns Hopkins Institutional Review Board approved the acquisition of clinical data and all biological samples for this study.

Demographic and clinical data were abstracted from the Johns Hopkins Scleroderma Center database and careful medical record review. Cancer diagnosis dates and histology were determined by review of the initial diagnostic pathology report. The clinical onset of scleroderma was defined by the first scleroderma symptom, either Raynaud's or non-Raynaud's. The interval between scleroderma onset and cancer diagnosis was calculated for each subject (cancer date – scleroderma onset date). The scleroderma cutaneous subtype and modified Rodnan skin score were defined by established criteria (*39*, *40*). All sera were tested for autoantibodies against RPC1, TOP1, and CENPB as previously described (*3*). Demographic and clinical data were compared across autoantibody groups, and differences in continuous and dichotomous/categorical variables were assessed by the Kruskal-Wallis and Fisher's exact tests, respectively.

The clinical features and cancer types of the 16 patients evaluated in this study are listed in Table 1. Eight patients were positive for anti-RPC1 antibodies, five for anti-TOP1 antibodies and three for anti-CENPB antibodies. Enhanced nucleolar staining with the anti-RPC1 antibodies (*3*) was observed in the tumors of all eight patients. No subject was positive for more than one autoantibody. Clinical phenotypic characteristics were representative of those expected in each autoantibody group (e.g. severe diffuse disease in RPC1-positive patients), and patients with RPC1 autoantibodies had a shorter interval between scleroderma onset and cancer diagnosis (median of -0.1 years vs. 13.4 years for patients with TOP1 autoantibodies and 34.0 years for patients with CENPB autoantibodies; p=0.05). Seven of the 16 patients had a short interval (+/- 2 years) between scleroderma onset and cancer diagnosis, and 6 of these 7 patients (85.7%) were positive for anti-RNA polymerase III antibodies.

Preparation of Illumina genomic DNA libraries

Genomic DNA libraries were prepared following Illumina's (Illumina, San Diego, CA) suggested protocol with the following modifications. **(1)** 50 to 75 microliters (µl) of genomic DNA from tumor or normal cells in a total volume of 100 µl TE was fragmented in a Covaris sonicator (Covaris, Woburn, MA) to a size of 100 to 500 bp. DNA was purified with a Nucleospin Extract II kit (Cat # 740609, Macherey-Nagel, Germany) and eluted in 50 µl of elution buffer included in the kit. **(2)** 45 µl of purified, fragmented DNA was mixed with 40 µl of H_2O, 10 µl End repair reaction buffer and 5 µl of End Repair enzyme. All reagents used for this step and those described below were from New England Biolabs (NEB cat# E6040, Ipswich, MA) unless otherwise specified. The 100 µl end-repair mixture was incubated at 20^{o}C for 30 min, purified by a PCR purification kit (Cat # 28104, Qiagen) and eluted with 42 µl of elution buffer (EB). **(3)** To A-tail, all 42 µl of end-repaired DNA was mixed with 5 µl of 10 x dA-Tailing Reaction buffer and 3 µl of Klenow Fragment (3' to 5' exo-). The 50 µl mixture was incubated at 37^{o}C for 30 min before DNA was purified with a MinElute PCR purification kit (Cat # 28004, Qiagen). Purified DNA was eluted with 27 µl of 70^{o}C EB. **(4)** For adaptor ligation, 25 µl of A-tailed DNA was mixed with 10 µl of PE-adaptor (Illumina), 10 µl of 5x Ligation buffer and 5 µl of Quick T4 Ligase. The ligation mixture was incubated at room temperature (RT) or 20^{o}C for 15 min. **(5)** To purify adaptor- ligated DNA, 50 µl of ligation mixture from step (4) was mixed with 200 µl of NT buffer from NucleoSpin Extract II kit (cat# 636972, Clontech, Mountain View, CA) and loaded into a NucleoSpin column. The column was centrifuged at 14000 g in a desktop centrifuge for 1 min, washed once with 600 µl of wash buffer (NT3 from Clontech), and centrifuged again for 2 min to dry completely. DNA was eluted in 50 µl elution buffer included in the kit. **(6)** To obtain an amplified library, ten or twenty PCRs of 50 µl each were set up, each including 30 µl of H_2O, 2.5 µl dimethyl sulfoxide (DMSO), 10 µl of 5 x Phusion HF buffer, 1.0 µl of a dNTP mix containing 10 mM of each dNTP, 0.5 µl of Illumina PE primer #1, 0.5 µl of Illumina PE primer #2, 0.5 µl of Hot Start Phusion polymerase, and 2.5 or 5 µl of the DNA from step (5). The PCR program used was: 98^{o}C 1 minute; 10 to 16 cycles of 98^{o}C for 20 seconds, 65^{o}C for 30 seconds, 72 ^{o}C for 30 seconds; and 72 ^{o}C for 5 min. To purify the PCR product, 250 µl PCR mixture (from the ten PCR reactions) was mixed with 500 µl NT buffer from a NucleoSpin Extract II kit and purified as described in step (5). Library DNA was eluted with 70^{o}C elution buffer and the DNA concentration was estimated by absorption at 260 nm.

Target DNA enrichment

The targeted regions included all 53 exons of CENPB, *POLR3A*, TOP1. Capture probes were designed (*41*) to capture both the plus and the minus strand of the DNA and had a 33-base overlap and were custom-synthesized by Agilent Technologies en masse on a solid phase and used for capture, essentially as described (*42*). Approximately 3 µg of library DNA was used per capture. After washing, the captured libraries were ethanol-precipitated and redissolved in 20 µl of Tris-EDTA (TE) buffer. The DNA was then amplified in a PCR mix containing 51 µl of

distilled water (dH_2O), 20 μl of 5× Phusion buffer, 5 μl of dimethyl sulfoxide (DMSO), 2 μl of 10 mM dNTPs, 50 pmol of Illumina forward and reverse primers, and 1 μl of HotStart Phusion enzyme (New England Biolabs) with the following cycling program: 98°C for 30 s; 15 cycles of 98°C for 25 s, 65°C for 30 s, 72°C for 30 s; and 72°C for 5 min. The amplified PCR product was purified with a NucleoSpin column (Macherey Nagel Inc.) according to the manufacturer's suggested protocol, except that the NT buffer was not diluted and the DNA bound to the column was eluted in 45 μl of elution buffer. The captured libraries were quantified using an Agilent BioAnalyzer.

Somatic mutation identification and LOH analysis

Captured DNA libraries were sequenced with the Illumina GAIIx Genome Analyzer. Sequencing reads were analyzed and aligned to human genome hg18 with the Eland algorithm in CASAVA 1.6 software (Illumina). A mismatched base was identified as a mutation only when (i) it was identified by ten or more distinct pairs; (ii) the number of distinct tags containing a particular mismatched base was at least 2.5% of the total distinct tags; and (iii) it was not present in >0.5% of the tags in the matched normal sample. Mutations were confirmed by amplification of the relevant region with a single primer pair and evaluated as described in (*43*). LOH analysis was performed in a similar way, using the primer pairs described in table S3. A patient was considered "informative' for the SNP if DNA from the normal tissue of that patient was heterozygous for the SNP. A tumor was determined to have undergone LOH if >75% of the informative primer pairs in that patient had an allelic ratio less than the mean minus 2 standard deviations of those measured in control individuals without scleroderma. Note that this analysis can only assess allelic imbalance, i.e., a gain in one allele or a loss in the other allele, though it is often (including in the current study) interpreted as LOH, i.e., *loss* of an allele.

Peptide microarray

All experiments with the peptide microarray were performed at ProImmune Inc. (Oxford, UK). Peptides were synthesized as 15-mers with 10 overlapping amino acids from the previous peptide, spanning the entire RPC1 protein. Peptides were printed on glass slides with multiple arrays per slide separated with gaskets, allowing for multiple donor sera to be tested per slide. Donor serum was diluted 1:100, 1:500 and 1:1000 and incubated on the array. A fluorescent anti-Human IgG antibody was used as a secondary antibody, results were detected using a CCD camera and analysis was done using MS Excel. Peptides were determined to be potential binders if the normalized average signal intensity was greater than 4X the respective background negative control. Binding was considered positive if the signal intensity was 4X background for all three serum dilutions.

Autoantibody analysis

RPC1 antibodies were assayed by ELISA using a commercially available kit (Inova Diagnostics). CENP and TOP1 autoanitibody assays were performed as described in (*3*). To

define whether patient antibodies recognized patient-specific mutated forms of RPC1, full-length wild type human *POLR3A* cDNA was purchased from Origene, and site-directed mutagenesis was performed to generate the three different RPC1 mutants, each with a single point mutation: E1072Q, K1365N and I104T, corresponding to the tumor mutations detected in patients SCL-2, SCL-4 and SCL-42, respectively. All were sequence verified before use. ^{35}S-methionine-labeled products were generated from the wild type and mutant DNAs by IVTT reactions (Promega kit). Prior to use in immunoprecipitations, the radiolabeled proteins were electrophoresed on SDS-PAGE gels and visualized by fluorography. The radiolabeled signal generated by each of these products was similar (1 µl of E1072Q equivalent to: 1.1, 1.2 and 1.7 µl of I104T, wt and K1365N, respectively). Equivalent radioactive amounts ^{35}S-methionine-labeled wt and mutated RPC1 proteins were used in immunoprecipitations performed as described in (*44*) with sera from three cancer scleroderma patients. One µl of each serum was used to immunoprecipitate the wild-type form of *RPC1* as well as the specific *RPC1* mutation found in the tumor from that patient.

Cell culture, stimulation and flow cytometry

PBMCs were freshly isolated from whole blood by density-gradient centrifugation (Ficoll-Paque Plus, GE Healthcare), and were used fresh (SCL-42) or frozen (SCL4 and SCL2). For each patient, PBMCs from a donor expressing one matching HLA-DRB1 allele was selected and used as a control. Cells were resuspended to a concentration of 1.5 x 10^6 cells/150µl in RPMI medium supplemented with serum, 2mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin, and were plated onto 96-well flat bottom plates.1 µg/ml anti-human CD40 blocking antibody (G28.5, Biolegend) was added, and after 30 minutes, cells were stimulated for 18 hours as indicated: 4 µg/ml wild-type or mutant patient-matched RPC1 peptide, 4 µg/ml of peptidyl arginine deiminase 4 (PAD4) peptide as a negative control, or 4 µg/ml of a pool of class-II peptides from infectious agents antigens (CEFT) (Axxora) as a positive control. Peptide storage buffer was used for the unstimulated control. HLA restriction was assessed by stimulating cells in the presence of 1 µg/ml anti-HLA-DR blocking antibody (L243, Biolegend) or 1 µg/ml IgG-2a κ isotype control (MOPC-173, Biolegend).

Cells were washed with PBS after stimulation, stained with live/dead fixable blue dead cell stain (Molecular Probes), and then stained with BV510-conjugated CD3 antibody (OKT3, BioLegend), Pacific Blue-conjugated anti-CD4 antibody (RPA-T4, BD Pharmingen), APC-H7-conjugated anti-CD8 (SK1, BD), and PE-conjugated anti-CD154 antibody (TRAP1, BD Pharmingen). FACS analysis was performed on FACSAria flow cytometer-cell-sorter using FACSDiva (Becton Dickinson) and FlowJo software (Tree Star Inc. Ashland, OR, USA).

TCR spectratyping

The diversity of CDR3 regions for 22 TCR Vβ regions was assessed using the TCRExpress Quantitative Analysis Kit (Biomed Immunotech, Tampa FL). Briefly, RNA was isolated from SCL-42 PBMCs after culture for 6 days with wt or mutant peptides (Invitrogen, Carlsbad

CA), and cDNA was generated using random hexamers (Invitrogen) and You Prime First Strand Beads (GE, Buckinghamshire UK) following the manufacturers protocols. CDR3 regions were amplified from cDNA using two rounds of PCR with Vβ-family specific PCR primers as per the manufacturer's instructions. Fragment length analysis was performed by the Johns Hopkins DNA analysis facility. The distribution of CDR3 lengths for each Vβ-family was determined and expressed as "proportion of TCR". Peaks were considered to be antigen-driven when the observed proportion of a given fragment size differed by more than 10% between wt and mutant-stimulated cells.

TCR Sequencing

TCR libraries were prepared for sequencing using a Truseq sample preparation kit following the manufacturer's suggestions with the following modifications. Input DNA was prepared from the PCR products obtained from spectratyping analysis. Product from wells using primers specific to the CDR3 regions of Vβ3, Vβ5, Vβ7, Vβ12, Vβ16 and Vβ24 were purified using a Qiagen PCR purification kit. DNA was eluted in 30 ul of 65°C elution buffer. After A-tailing and igation to adaptors, the library was amplified in six reactions with a PCR mix containing 10 ul H20, 1.5 ul DMSO, 6 ul 5X Phusion buffer, 6 ul dNTPs, 3 ul each of Forward and Reverse Primers, 3 ul Phusion polymerase (2 U/ul) and 2 ul ligation reactions, and cycled using the following program, 98°C for 30 s; 14 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 30 s; and 72°C for 5 min. The resulting product was purified using Ampure beads, quantified with an Agilent Bioanalyzer, and sequenced using an Illumina instrument.

qPCR detection of specific TCRs

Custom Taqman assays for specific TCRs were developed using Primer express v2.0 software and synthesized by Applied Biosystems. For detection of the Vβ24 Jβ1.1 wt TCR, a FAM-labeled probe (ACTGAAGCTTTCTTTGGAC), forward primer (GCACCGGGACAGTGATGAA), and reverse primer (GGTCCTCTACAACTGTGAGTTTGGT) were synthesized. For detection of the Vβ24 Jβ1.5 mutant TCR, a FAM-labeled probe (ACAGTAAATCAGCCCCAGC), forward primer (TGTGTGCCACCAGCAGAGA), and reverse primer (AGTCGAGTCCCATCACCAAAA) were synthesized. cDNA from Day 0 SCL-42 PBMCs was prepared as described and was amplified by one round of PCR using Vβ24-family specific PCR primers (Biomed Immunotech, Tampa FL). Expression of the specific Vβ24 TCRs was determined in triplicate using standard ABI chemistry and reagents.

Clinical descriptions of anti-RPC1 positive patients

As detailed below, the anti-PRC1 positive scleroderma patients with cancer shared many features, including a short interval between the first clinical signs of scleroderma and cancer diagnosis, aggressive cutaneous disease, and a high risk of scleroderma renal crisis.

Patient SCL-1

Patient 1 palpated a breast mass in the summer of 2005 and was diagnosed with a breast invasive ductal carcinoma on 7/26/05. Around the time of her diagnosis, she developed hypertension, thrombocytopenia, seizures, and renal failure that progressed despite blood pressure control. She ultimately initiated peritoneal dialysis. Her breast cancer was treated with lumpectomy, radiation therapy, and doxorubicin. In October of 2005, she began to notice Raynaud's phenomenon, and in October of 2007, she developed skin thickening. When first seen in our Center on 2/4/08, she was noted to have extensive and severe scleroderma skin thickening with a modified Rodnan skin score (mRSS) of 41 (range is 0-51 with 51 representing most severe disease possible).

Patient SCL-2

Patient 2 was noted to have a mass on a chest radiograph, leading to a diagnosis of a small cell carcinoma of the lung on 3/22/06. She clearly had Raynaud's phenomenon by April 2006. She was treated with chemotherapy (completed 7/06) and radiation therapy with prophylactic brain irradiation (completed 9/06). By April of 2007, she began to notice worsening of her Raynaud's phenomenon and swelling of her hands followed by the onset of rapid, diffuse skin thickening. She was initially treated for her cutaneous disease with mycophenolate mofetil and was noted to have a mRSS of 47 on her visit to our Center on 9/6/07. She later required therapy with cyclophosphamide due to concern for interstitial lung disease, and by 12/13/11, her mRSS had decreased significantly to 5.

Patient SCL-4

Patient 4 had a known BRCA1 mutation and underwent a prophylactic oophorectomy based on her genetic risk. During this procedure, she was found to have stage III ovarian adenocarcinoma with papillary serous features (4/5/06). She completed 6 cycles of paclitaxel and cisplatin in 8/2006, and around this time developed Raynaud's phenomenon. Her chemotherapy course was complicated by the development of pericarditis with tamponade physiology requiring drainage and a pericardial window, and there was no evidence of an infected or malignant effusion. When first seen here on 1/10/08, she was noted to have significant skin disease (mRSS 21), numerous tendon friction rubs, a myopathy, and scleroderma renal crisis with a Cr of 1.9. In 2008-2009, she was treated with a number of immunosuppressive agents targeting her cutaneous, muscle, and joint disease including mycophenolate, methotrexate, azathioprine and hydroxychloroquine, and her skin disease was significantly improved by 8/08. In October 2011, she developed a small bowel obstruction with imaging findings consistent with serosal implants in the context of a

rising CA 125 level. She began weekly carboplatin, and her CA 125 level had normalized by April 2012.

Patient SCL-13

Patient 13 noticed bilateral hand and ankle swelling in May 2005 and Raynaud's phenomenon in August 2005. She was diagnosed with invasive ductal carcinoma of the breast on 8/24/05. She was treated with a mastectomy followed by doxorubicin and cyclophosphamide from 10-12/05. In 01/06, she developed worsening skin thickening in her hands, arthralgias and myalgias and began therapy with d-penicillamine and later methotrexate. She also initiated paclitaxel and trastuzumab for her cancer in 01/06. She was seen at our Center 4/16/07 and was noted to have progressive skin disease (mRSS 30); she was transitioned to mycophenolate for her cutaneous disease and gradually had an improvement in her skin disease (mRSS in 2012 was 2).

Patient SCL-35

Patient 35 was diagnosed with breast ductal carcinoma in situ on 8/16/04, treated with a mastectomy. In April 2006, she developed symptoms consistent with carpal tunnel syndrome, and by 8/06, she had lower extremity skin thickening and tendon friction rubs. Raynaud's phenomenon developed in 1/07. When first seen at our Center in 6/07, her mRSS was 48, and she was on letrozole for her malignancy. After therapy with cyclophosphamide and mycophenolate, her cutaneous disease significantly improved (mRSS 3 by 3/11).

Patient SCL-42

Patient 42 developed arthralgias and skin thickening in her fingers in 3/2007 that rapidly progressed to diffuse skin thickening. In 4/2007 she developed Raynaud's phenomenon also. By 10/07, she developed hypertension requiring ACE-inhibitor therapy, and her mRSS was 18. Despite therapy with mycophenolate, her cutaneous disease progressed, and by 3/08, her mRSS had increased to 37. In 4/08, methotrexate was added to her regimen with some improvement in flexibility and new hair growth; however her mRSS remained at 35 in 7/08. In 9/08, in the setting of increased cutaneous activity, she was diagnosed with a stage II, triple negative, invasive ductal cancer of the breast. She was treated with mastectomy (10/08) and chemotherapy with cyclophosphamide and docetaxel (12/08-2/09). Her cutaneous symptoms gradually improved in the setting of IVIG therapy, and by 12/11 her mRSS was 8.

Patient SCL-81

Patient 81 was diagnosed with an adenocarcinoma of the colon with positive lymph nodes in 5/05 requiring bowel resection and 5-FU and platinum chemotherapy. In 7/09, he developed diffuse cutaneous skin disease and by 11/09 had Raynaud's phenomenon. In 3/10, his mRSS was 46.

Patient SCL-82

Patient 82 developed Raynaud's phenomenon and hand swelling in 1/08 and was noted to have a mRSS of 14 in 8/08 while on methotrexate therapy. By 1/09, the patient was on combination mycophenolate and methotrexate therapy for a mRSS of 18. By 1/10, her skin disease was significant improved (mRSS 5), but she was diagnosed with a ductal carcinoma in situ (DCIS) of the breast on 6/28/10 and treated with a mastectomy.

Fig. S1. immunoprecipitations of wt and mutant RPC proteins by sera from scleroderma cancer patients. ^{35}S-methionine-labeled wt and mutated RPC1 proteins were generated by IVTT ("IVTT Input"). For each radiolabeled RPC1 protein, the amount used for the input gel samples was 1/20 of the amount used for immunoprecipitation. Immunoprecipitates were electrophoresed on SDS-polyacrylamide gels and visualized by fluorography. (**A**) Immunoprecipitations (performed in duplicate) with patient sera SCL-2, 4 and 42 ("IVTT IP"). The levels of anti-RPC1 antibodies in each of the sera (assayed by ELISA) is listed; values >80 units denote high levels of these antibodies. (**B**) Immunoprecipitations were performed with the indicated scleroderma patient and control sera (right panel, "IVTT IP").

Fig. S2. Peptide array. Peptides determined to be positive binders (see Methods) are shown in blue. Only peptides that bound to the sera of at least one patient are displayed, however all 276 peptides (table S4) spanning the entire wt RPC sequence, as well as peptides spanning the identified *POLR3A* mutations, were included on the array.

Fig. S3. The effect of HLA-DR blocking antibodies on activation of CD4+ T cells. PBMCs from patient SCL-42 were stimulated with patient-matched mutant and corresponding wild-type RPC1 peptides. The PAD4 peptide and the CEFT pool were used as negative and positive controls, respectively. CD4+ T cell responses to RPC1 peptides were reduced by the presence of HLA-DR blocking antibodies (1 ug/ml) but not by an isotype control antibody used at the same concentration.

Fig. S4. Detection of wt and mutant-specific TCRs by qPCR. (**A**) Patient SCL-4 or SCL -42 PBMCs were cultured with patient-specific wt and mutant peptides for six days prior to cDNA isolation and amplification with Vβ24-based primers specific for the TCRs recognizing the wt or mutant forms of *POLR3A* found in patient SCL-42. As indicated. In the lower panel, the relative expression levels of the SCL-42 TCRs following stimulation with the indicated peptides were compared to those of GAPDH and displayed as $2^{-\Delta Ct}$ (lower panel). (**B**) Unstimulated patient SCL-42 PBMCs were used to generate cDNA which was then amplified with Vβ24-based primers specific for the TCRs recognizing the wt or mutant forms of *POLR3A* found in patient SCL-42. In the lower panel, the same cDNA was used to amplify the TCR-β constant region as a positive control.

Fig. S5. Mutant and wild type peptide-specific CD4+ T cells in patient SCL2. CD154 expression on CD4+ T cells was assayed by flow cytometry after stimulation (18 h) with patient-specific wild type or mutant RPC1 peptides, or a pool of peptides from infectious agent antigens (CEFT, positive control). Gate frequencies are expressed as percentage of CD4+ T cells.

Table S1. Demographic and clinical characteristics grouped by autoantibody status

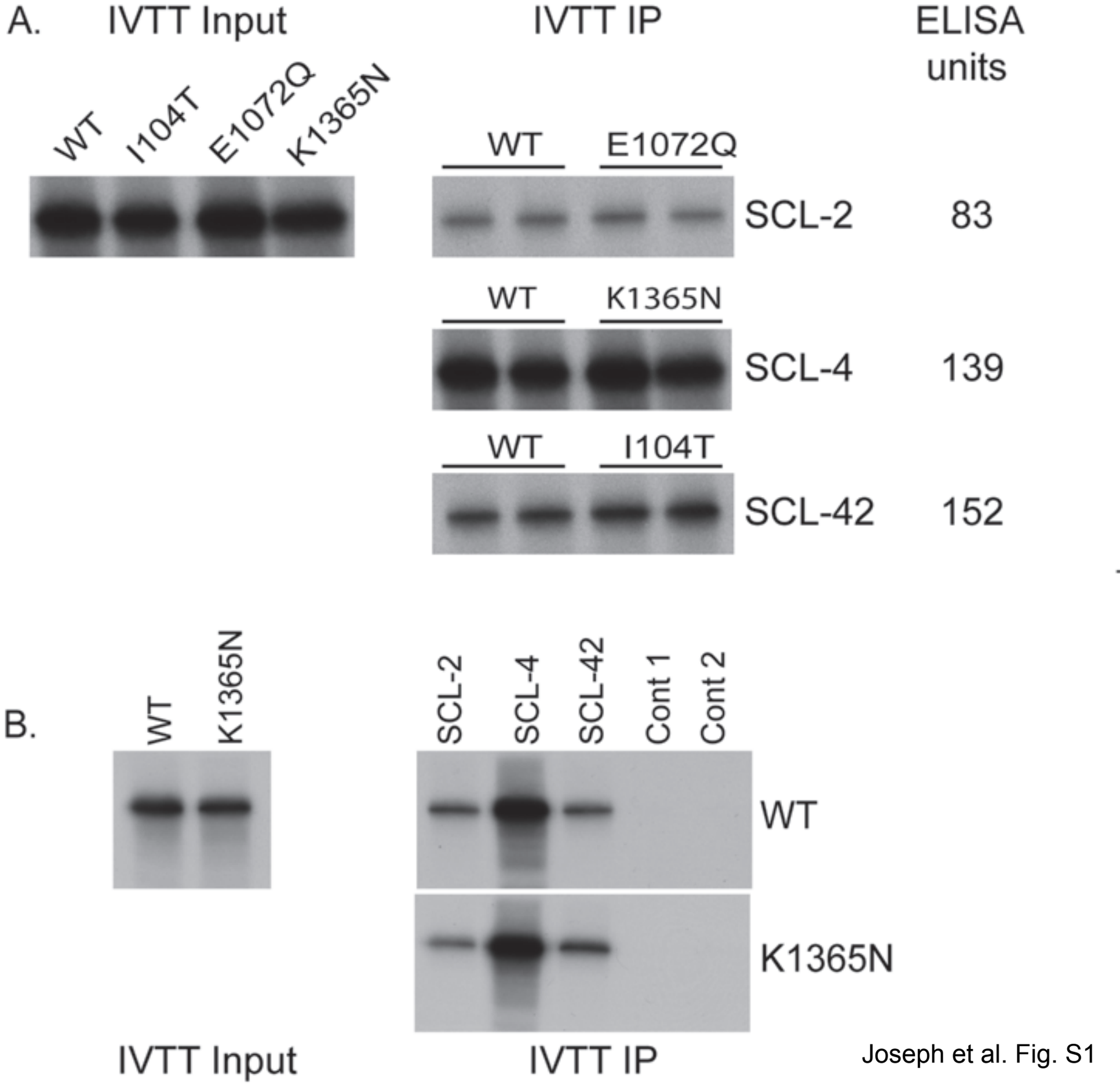
Table S2. Primers used for loss of heterozygosity analysis.

Table S3. Allelic ratios of SNP loci within and closely surrounding the TOP1 gene.

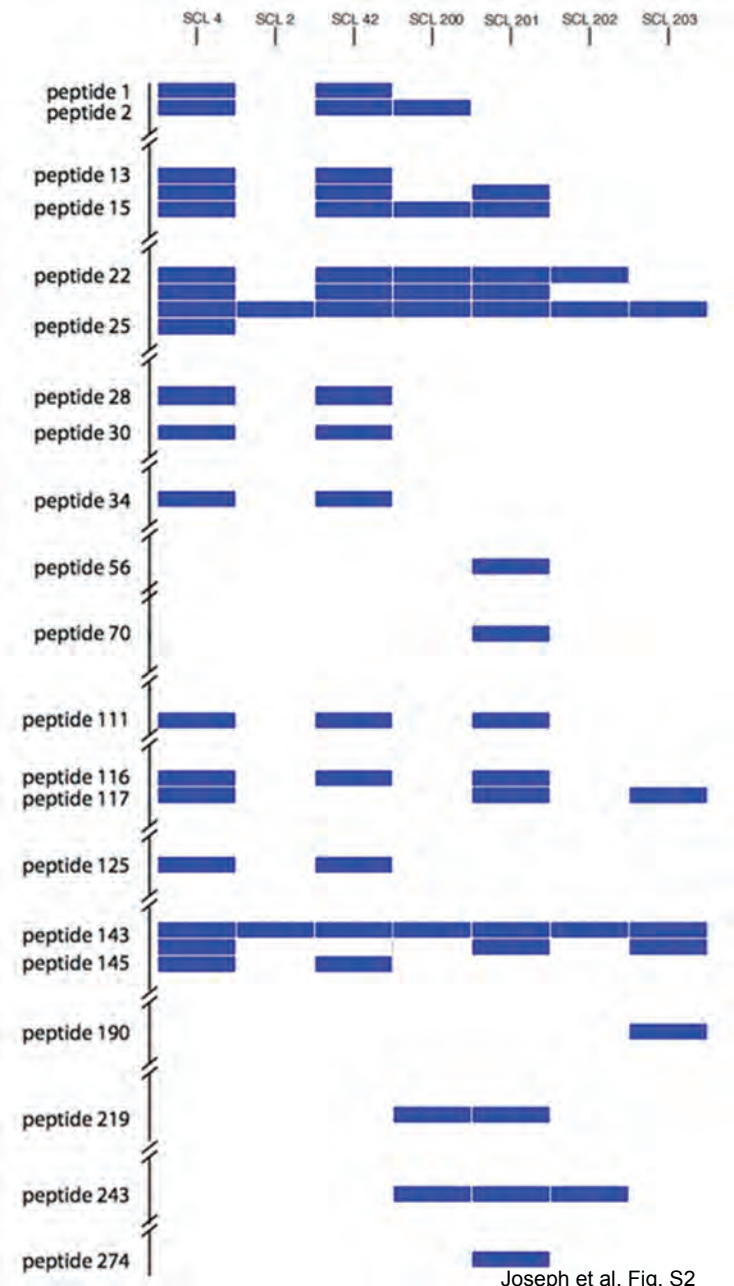
Table S4. Synthetic peptides assessed for antibody reactivity.

Table S5. MHC types of the patients with RPOL3A mutations.

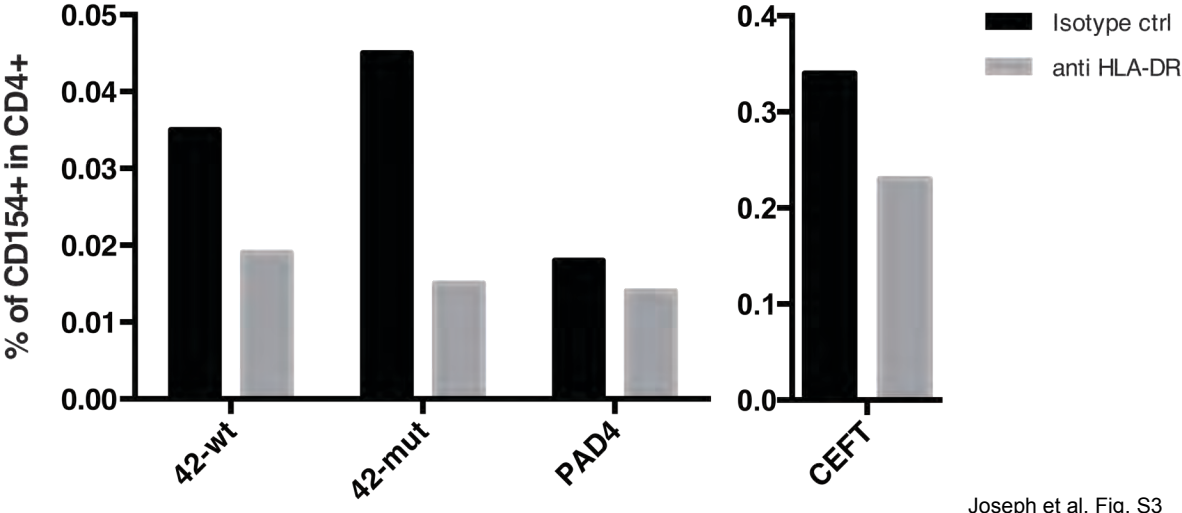
Table S6. Patient-specific MHC class I and class II peptides with highest predicted binding affinity



Joseph et al. Fig. S1

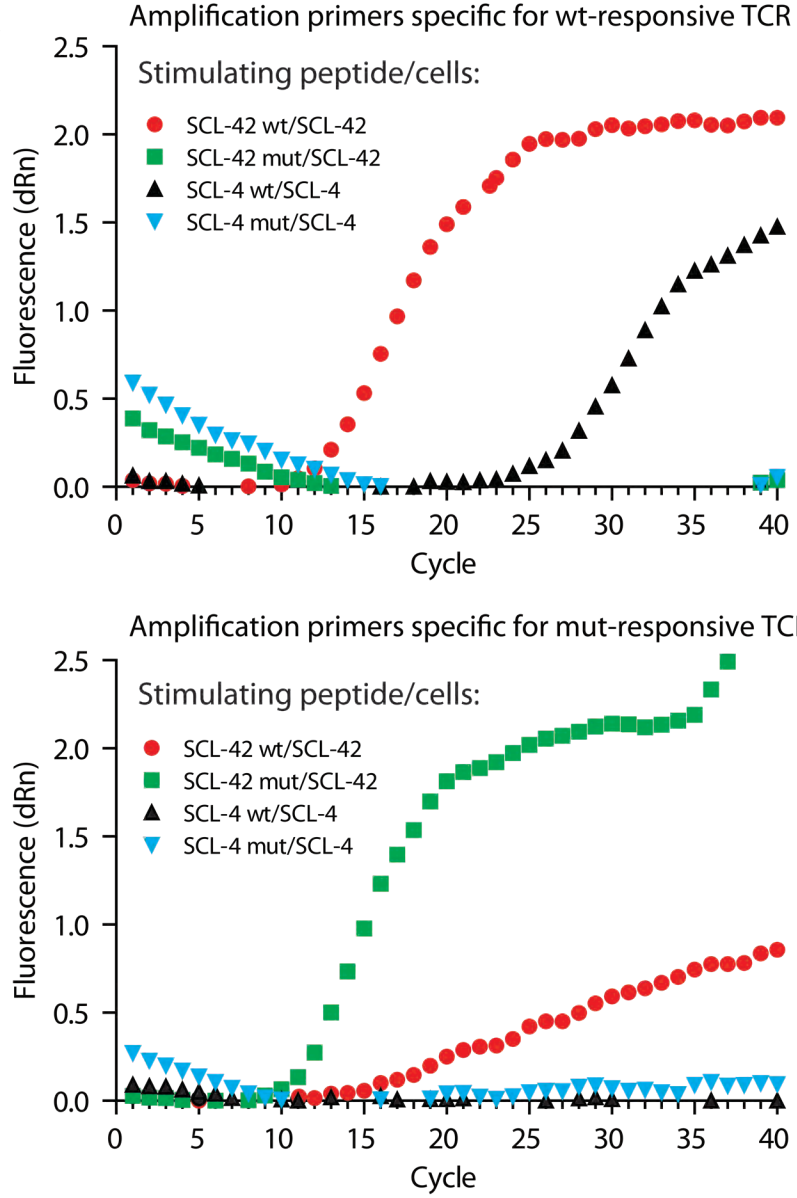


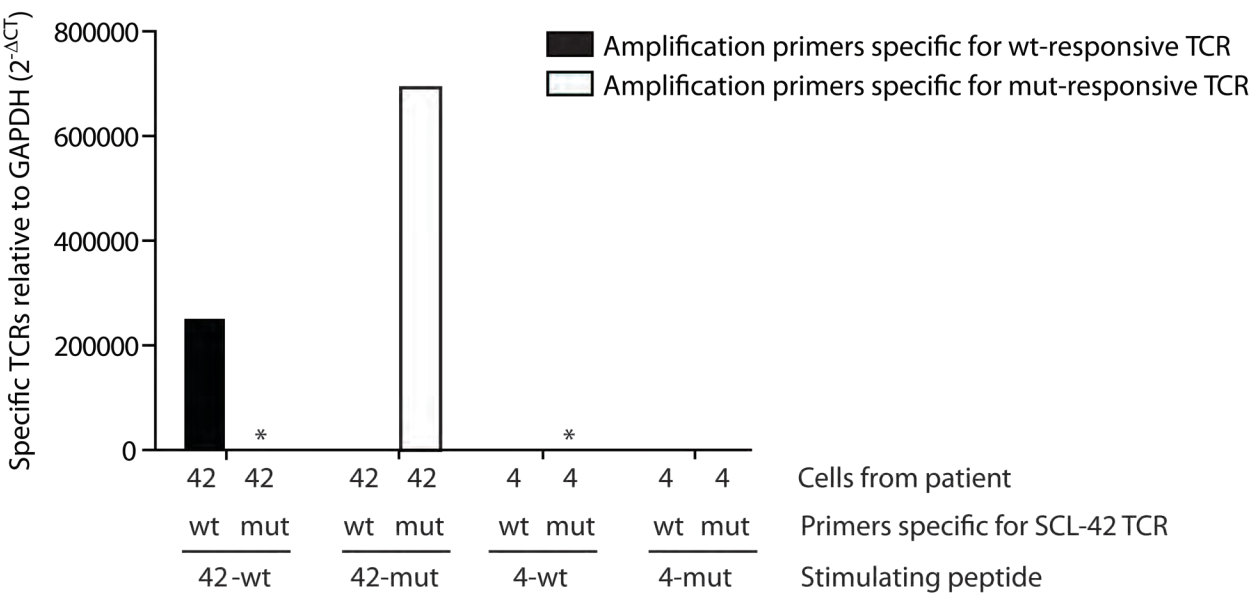
Joseph et al. Fig. S2



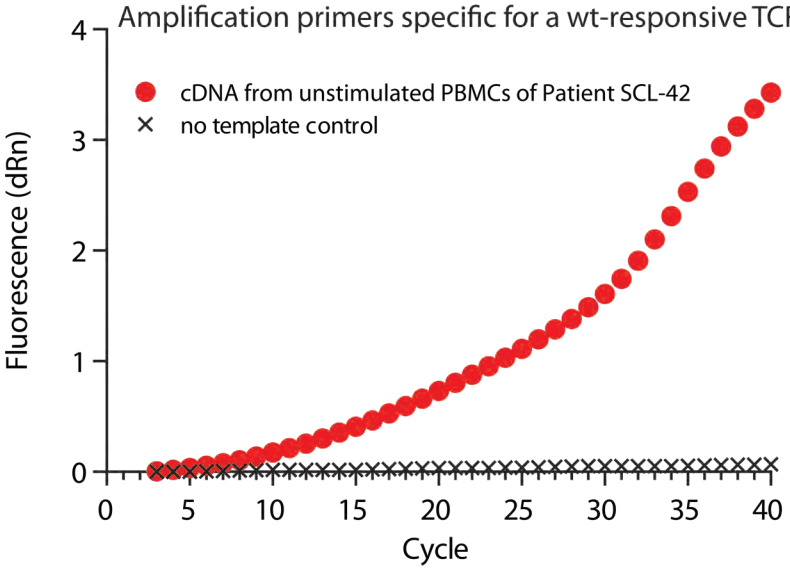
Joseph et al. Fig. S3

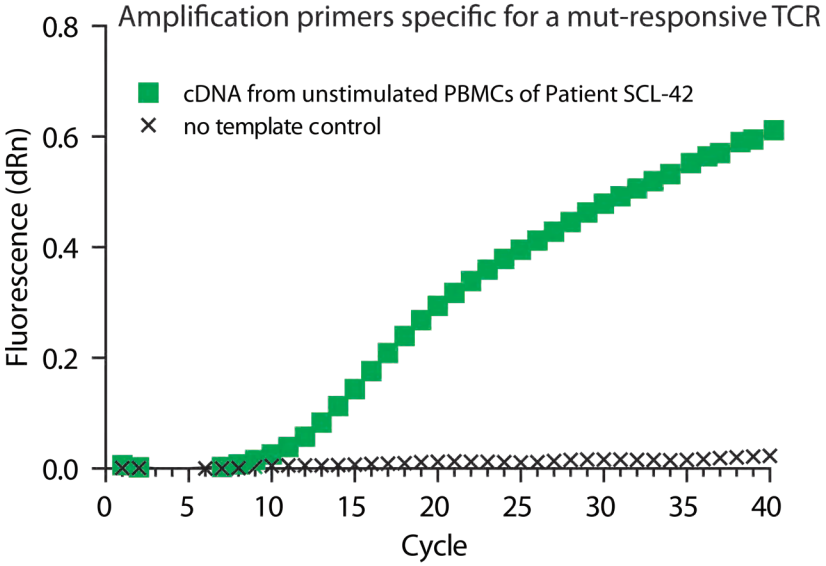
A

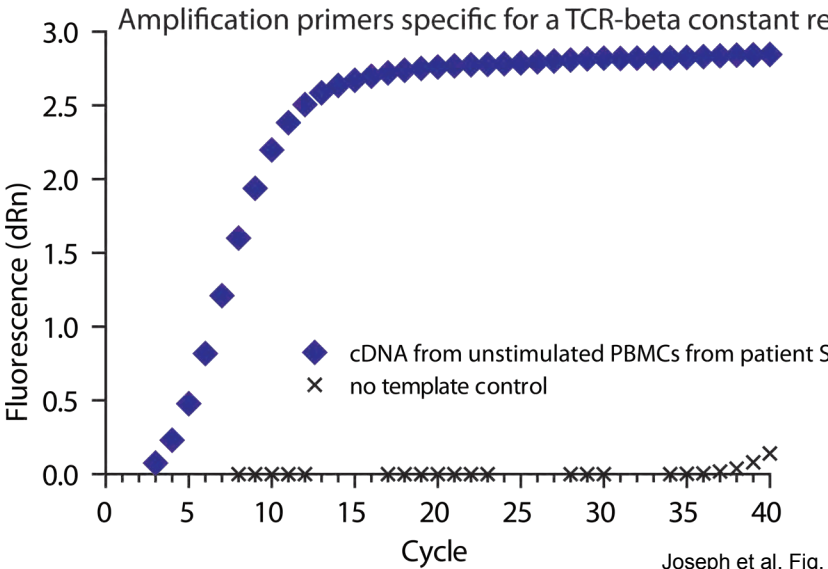

Amplification primers specific for wt-responsive TCR
Stimulating peptide/cells:
SCL-42 wt/SCL-42
SCL-42 mut/SCL-42
SCL-4 wt/SCL-4
SCL-4 mut/SCL-4
Fluorescence (dRn)
Cycle
Amplification primers specific for mut-responsive TCR
Stimulating peptide/cells:
SCL-42 wt/SCL-42
SCL-42 mut/SCL-42
SCL-4 wt/SCL-4
SCL-4 mut/SCL-4
Fluorescence (dRn)
Cycle


Amplification primers specific for wt-responsive TCR
Amplification primers specific for mut-responsive TCR
Specific TCRs relative to GAPDH (2-ΔCT)
42 42 42 42 4 4 4 4
wt mut wt mut wt mut wt mut
Cells from patient
Primers specific for SCL-42 TCR
42-wt
42-mut
4-wt
4-mut
Stimulating peptide

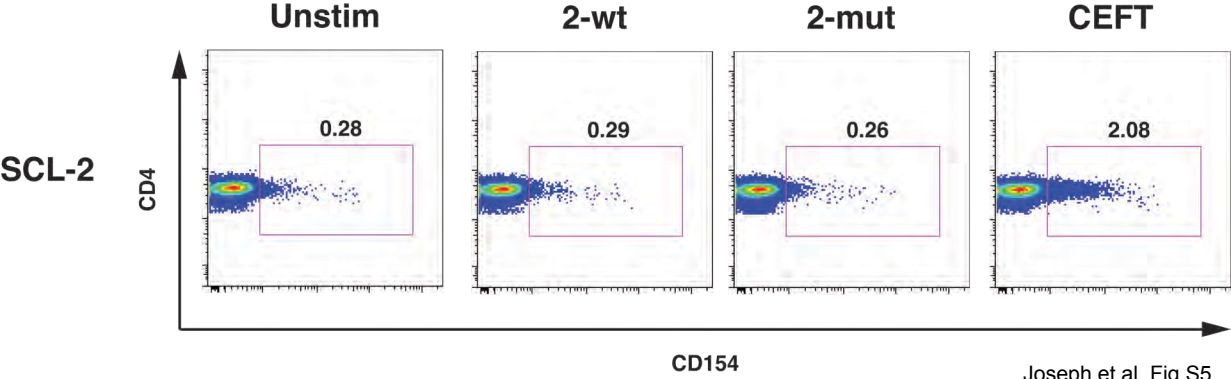
B







Joseph et al. Fig. S4B



Joseph et.al. Fig.S5

Table S1. Demographic and clinical characteristics grouped by autoantibody status

	RPC1 (N=8)	TOP1 (N=5)	CENPB (N=3)	p-value
Age at scleroderma onset, median (range), years	49.7 (42.4, 58.8)	42.4 (37.0, 65.3)	34.5 (27.3, 45.6)	0.12
Age at cancer diagnosis, median (range), years	50.9 (42.3, 54.6)	65.1 (52.1, 74.6)	64.2 (43.1, 68.6)	0.03
Interval between scleroderma onset and cancer diagnosis, median (range), years	-0.1 (-4.2, 2.5)	13.4 (0.4, 34.0)	34.0 (-2.5, 36.9)	0.05
Sex, no. (%) female	7 (87.5)	5 (100)	3 (100)	1.0
Race, no. (%)				
White	8 (100)	3 (60)	3 (100)	0.23
Black	0 (0)	1 (20)	0 (0)	
Asian	0 (0)	1 (20)	0 (0)	
Scleroderma classification, no. (%)				
Limited				
Diffuse	0 (0)	3 (60)	3 (100)	0.003
	8 (100)	2 (40)	0 (0)	
Ever smoker, no. (%)	2 (25)	2 (40)	1 (33.3)	1.0
Maximum mRSS, median (range)	38 (18, 48)	6 (4, 27)	4 (2, 10)	0.006
Renal crisis, no. (%)	3 (37.5)	0 (0)	0 (0)	0.24
Immunosuppressive therapy, ever use, no. (%)				
Methotrexate	4 (50)	1 (20)	0 (0)	0.39
Mycophenolate	6 (75)	0 (0)	0 (0)	0.008
IVIG	1 (12.5)	0 (0)	0 (0)	1.0
Cyclophosphamide	1 (12.5)	2 (40)	0 (0)	0.39
Cancer types	5 Breast 1 Lung 1 Ovary 1 Colon	4 Breast 1 Lung	1 Breast 1 Anal 1 Lymphoma	N/A

RP = Raynaud's phenomenon; mRSS = modified Rodnan skin score

Table S2. Primers used for loss of heterozygosity analysis

SNP Symbol	Related Gene	Coordinates	Distance (bp)*	Forward primer sequence**	Reverse primer sequence**
rs1054608	POLR3A	chr10:79213314	-193928	TGGACACACATCTTTCAGATTCA	CAAGATTTGGACTCAGCAGTTG
rs2579153	POLR3A	chr10:79218272	-188970	GCTACTTTTCTGGCCTGTGG	AGGGAGAACTAAAGGGGAACC
rs2165046	POLR3A	chr10:79222098	-185144	AGAATGCTGTGCTGTGGATG	GTGCTTCTGGGTCCTGGTT
rs10762763	POLR3A	chr10:79230809	-176433	TGAGGTCAGCTGAGTCTGTGG	GGATTTTCAGGGCTCCGAGT
rs2289311	POLR3A	chr10:79235661	-171581	GTGTCTGGCCCACCTTTG	ATTCTTGGTCCATCCTGTGG
rs10824579	POLR3A	chr10:79260691	-146551	CTCCTTTTCCAACTGGGATTC	CTGCCTCGTGAGGTTCAGA
rs1248888	POLR3A	chr10:79323400	-83842	TGTATCAATGGGCAGCAGTG	TATGTCATTCTGCCCCCAAG
rs2241547	POLR3A	chr10:79406970	-272	CCGATGGATGTATGCAGTGA	TTGTTTGCTTGCTTGAAACCT
rs3815891	POLR3A	chr10:79415795	0	TCTACTTGGGTGGGGGTGAT	CTCCAAGACGCGTTAGAACC
rs7094028	POLR3A	chr10:79419810	0	AACACAAGAAGCGAGGAGCTT	AAGATCAGCAGGCCAAAGAA
rs2818827	POLR3A	chr10:79424105	0	TGCTGTGGCTTTGTGTCTTC	GTGTTTGTTCTGGCCCACTC
rs2493568	POLR3A	chr10:79442860	0	TCGTCTTCTTAGAGGATGCAGTATT	TGCCTCACTATCACCGATCTC
rs4979801	POLR3A	chr10:79514072	54901	CCGTTGAGATTCATTCTCTCCT	TTGTTTAAAATTTCCCTGCTCCT
rs2253909	POLR3A	chr10:79546360	87189	TGGATTGATTAAACAGATGTTAAGGTT	CAGAAGGGCCGTTTACACTC
rs2253513	POLR3A	chr10:79549686	90515	AGCATGTCTGAGCCTCTTTTTC	CAGCAAAGCAGCAACAAAAG
rs2114907	POLR3A	chr10:79573735	114564	CACTCTCTCCACCACAAGCA	TTGGCCACCATCAGTAAGAAC
rs1249134	POLR3A	chr10:79615946	156775	AGAATGGGCCTCTTTGGACT	CAGGGAGCTGTTTTCAGGTG
rs1249126	POLR3A	chr10:79618728	159557	ACACTGTGGGGAGGGAAAA	AGCATTTCTGGCCTCCTGT
rs2434123	POLR3A	chr10:79654802	195631	GCCATCTAGTCTGCGAAAGG	GTGGGAGAAAGTCAGTTTGAATAAA
rs2865899	TOP1	chr20:38898952	-192170	CTCCATCGCCCTAGACTGAA	CTGCTGCAGGAGTCCACAC
rs6029375	TOP1	chr20:38899311	-191811	CCTCATCATCTGGCACCTCT	AGTGACCTCACACATTGAGGAGT
rs1005533	TOP1	chr20:38920524	-170598	TCATTTGTGTGGGTTTGTGTG	CTCCTTATGCCTCCCCTGA
rs1109393	TOP1	chr20:38927262	-163860	GGGAGCTGACCCAATCCAG	CCTGAAGGTTGCCCATAAAA
rs6029395	TOP1	chr20:38929002	-162120	CCCAGGTCTTCCAAACACAG	ACTGGAAGCTGCCACTTGAG
rs6129694	TOP1	chr20:38931100	-160022	CCACATTTCAGAGCCCAGAG	GCCAAGGCTACACATGGTTT
rs4142346	TOP1	chr20:38947279	-143843	TTTTAACTACTCATACATCCCCATGTC	CCCTCTTCCATCCCGTATTT
rs6129700	TOP1	chr20:38954995	-136127	CCAAGCCAATTTTCCAAAGA	GCATCCTCAGGCTGTTTGAC
rs6124288	TOP1	chr20:38955031	-136091	GGGATGTGCACTGAAACTGAT	TCAACTCAGCTCAAGGATTGC
rs2207309	TOP1	chr20:39011156	-79966	TGGGGGACAAACTTTTAGGG	GGCTAGGAGTTGCCTTCAATC
rs6072263	TOP1	chr20:39138960	0	CTCTTCACCCCTGACTCACTG	ACACAAAGGGATACACACACACA
rs6129731	TOP1	chr20:39094101	0	TCTGGAAGATGTGGTGTTGGT	GACATGTCTGGGCATAATTAAACA
rs6029549	TOP1	chr20:39188109	2758	GTCCAAGAACAGGCCAGATG	GGAGAAAAGCCAAGTCAGCA
rs926345	TOP1	chr20:39205361	20010	TGCCTCTGCTGTCACTGTTC	CCTCACTCCCTTCCTCCAAC
rs3795128	TOP1	chr20:39207577	22226	TCTGGGAAGCAAGTGTGATG	AACCCATGGCCCTACTTGAG
rs6093446	TOP1	chr20:39214346	28995	GGAGAGGGTGATTGGTGAGA	GGGAAAAACAGTAAAAGCAGCA
rs2076148	TOP1	chr20:39222269	36918	CAGGCTTCTCTCTCCAGGTG	CCGGGCAGCTACTTAGAGG
rs6124323	TOP1	chr20:39227903	42552	TTCCTCGTGCGAGAGAGTG	CTCTGGGTGATGTGATGTCTGT
rs2235367	TOP1	chr20:39263536	78185	GAGGCTGGGAACTCAGTGAC	TTTTGTTCCTCAGTTTCTCTGGA
rs6029610	TOP1	chr20:39348159	162808	TGTTGAGAATGCACTTTCTTGAAT	TAAGTGGCTAATTTTGGGTAATGG
rs4812493	TOP1	chr20:39382799	197448	ATTGGAATGGGCTACACCTG	CCCACCCAAAGCCTAAAGAT
rs4810312	TOP1	chr20:39402121	216770	GGTCTTTGGTCCAGGAAAGC	GGACAGGTGAGCCCAGGT
rs6029636	TOP1	chr20:39402927	217576	GGTCTTTCCAGGGAGTGGAG	CTCAGTCACTTGGCGCTTC
rs12625565	TOP1	chr20:39419499	234148	TGGAGCTGCCTTGGAGTAAC	CAGCCCAGGCACTCACAG
rs2072966	TOP1	chr20:39436751	251400	GATTGGAGCTCGACTCAGAAAT	CAGGCTGAGGACTTGAGGAG

* Distance between the SNP and the indicated gene is listed as "0" if the SNP was within the coding region of the gene. Otherwise, the distance listed represents that between the SNP and the ATG of the coding sequence (negative numbers) or the SNP and the stop codon (numbers >0).

** 5'-cacacaggaaacagctatgaccatg was added to 5'-end of Forward PCR primers and 5'-cgacgtaaaacgacggccagtNNNNNNNNNNNNNNwas added to the 5'-end of Reverse PCR primers to enable the second ampification step, as described in (40).

Table S3. Allelic ratios of SNP loci within and closely surrounding the *TOP1* gene.

		Patients with RPC-1 antibodies								Patients without RPC-1 antibodies							
Chr 20 position	**SNP ID**	**SCL-1**	**SCL-2**	**SCL-4**	**SCL-13**	**SCL-35**	**SCL-42**	**SCL-81**	**SCL-82**	**SCL-5**	**SCL-8**	**SCL-11**	**SCL-12**	**SCL-19**	**SCL-24**	**SCL-32**	**SCL-85**
38,898,952	rs2865899	NI	NI	103%	NI	NI	NI	98%	NI	NI	NI	NI	103%	NI	NI	NI	NI
38,899,311	rs6029375	NI	NI	103%	NI	102%	103%	96%	103%	NI	NI	NI	NI	98%	104%	NI	NI
38,920,524	rs1005533	NI	NI	102%	NI	101%	102%	93%	99%	NI	NI	NI	103%	96%	104%	NI	NI
38,931,100	rs6129694	NI	NI	101%	95%	NI	NI	**87%**	NI	NI	NI	NI	NI	104%	NI	NI	NI
38,947,279	rs4142346	NI	NI	101%	NI	102%	103%	NI	96%	103%	NI	NI	NI	NI	NI	NI	NI
38,954,995	rs6129700	NI	NI	99%	NI	103%	101%	NI	NI	NI	NI	NI	NI	NI	103%	NI	NI
38,955,031	rs6124288	NI	NI	101%	NI	103%	NI	NI	101%	NI	NI	NI	NI	102%	102%	NI	NI
39,011,156	rs2207309	NI	NI	104%	104%	NI	NI	NI	NI	93%	101%	99%	NI	NI	103%	NI	NI
39,094,101	rs6129731	NI	NI	NI	NI	NI	NI	95%	NI	NI	NI	NI	101%	NI	NI	NI	104%
39,138,960	rs6072263	**81%**	NI	NI	104%	NI	102%	94%	NI	NI	NI	NI	NI	104%	NI	NI	104%
39,188,109	rs6029549	**83%**	NI	NI	102%	NI	NI	98%	NI	NI	NI	NI	NI	98%	NI	NI	98%
39,205,361	rs926345	100%	NI	NI	104%	NI	100%	95%	NI	98%	NI	NI	NI	NI	NI	NI	98%
39,207,577	rs3795128	**86%**	NI	NI	104%	NI	104%	94%	NI	103%	NI	NI	NI	95%	NI	NI	102%
39,214,346	rs6093446	NI	NI	NI	NI	NI	NI	NI	NI	**83%**	NI	103%	96%	NI	NI	NI	98%
39,222,269	rs2076148	103%	NI	NI	NI	NI	101%	**87%**	NI	92%	NI	NI	NI	NI	NI	NI	NI
39,227,903	rs6124323	**88%**	NI	NI	98%	NI	102%	**86%**	NI	92%	NI	NI	NI	NI	NI	NI	101%
39,263,536	rs2235367	NI	NI	NI	NI	NI	NI	100%	NI	96%	NI	NI	NI	NI	NI	NI	NI
39,348,159	rs6029610	NI	NI	NI	NI	NI	101%	100%	NI	NI	NI	NI	NI	NI	NI	NI	100%
39,382,799	rs4812493	NI	NI	NI	94%	NI	101%	NI	NI	NI	NI	NI	NI	NI	NI	NI	93%
39,402,927	rs6029636	NI	102%	NI	95%	101%	102%	103%	NI	NI	NI	101%	NI	NI	NI	NI	99%
39,436,751	rs2072966	**88%**	97%	NI	NI	97%	NI	NI	NI	NI	NI	93%	104%	NI	NI	NI	NI
Average allelic ratio		**90%**	**99%**	**102%**	**100%**	**101%**	**102%**	**95%**	**100%**	**95%**	**101%**	**99%**	**101%**	**99%**	**103%**	**NI**	**100%**
Standard deviation (%)		**9%**	**3%**	**2%**	**4%**	**2%**	**1%**	**5%**	**3%**	**6%**	**0%**	**4%**	**3%**	**4%**	**1%**	**NI**	**3%**

*Entries represent the allelic ratios of the indicated SNPs. The values in DNA from normal individuals was 100% +/- 4.40%. Allelic ratios less than 2 standard deviations (SD) from the mean (i.e., <91.2%) are highlighted in red. A tumor was considered to exhibit LOH if more than 3/4 of the informative SNPs exhibited alleic ratios <91.2%. NI: Non-informative, i.e., the SNP was not heterozygous in the normal cells of the patient (see Supplemental material, Materials and Methods).

Table S4. Synthetic peptides assessed for antibody reactivity		
Peptide ID	**RPC1 amino acid numbers**	**Sequence**
1	1-15	MVKEQFRETDVAKKI
2	6-20	FRETDVAKKISHICF
3	11-25	VAKKISHICFGMKSP
4	16-30	SHICFGMKSPEEMRQ
5	21-35	GMKSPEEMRQQAHIQ
6	26-40	EEMRQQAHIQVVSKN
7	31-45	QAHIQVVSKNLYSQD
8	36-50	VVSKNLYSQDNQHAP
9	41-55	LYSQDNQHAPLLYGV
10	46-60	NQHAPLLYGVLDHRM
11	51-65	LLYGVLDHRMGTSEK
12	56-70	LDHRMGTSEKDRPCE
13	61-75	GTSEKDRPCETCGKN
14	66-80	DRPCETCGKNLADCL
15	71-85	TCGKNLADCLGHYGY
16	76-90	LADCLGHYGYIDLEL
17	81-95	GHYGYIDLELPCFHV
18	86-100	IDLELPCFHVGYFRA
19	91-105	PCFHVGYFRAVIGIL
20	96-110	GYFRAVIGILQMICK
21	101-115	VIGILQMICKTCCHI
22	106-120	QMICKTCCHIMLSQE
23	111-125	TCCHIMLSQEEKKQF
24	116-130	MLSQEEKKQFLDYLK
25	121-135	EKKQFLDYLKRPGLT
26	126-140	LDYLKRPGLTYLQKR
27	131-145	RPGLTYLQKRGLKKK
28	136-150	YLQKRGLKKKISDKC
29	141-155	GLKKKISDKCRKKNI
30	146-160	ISDKCRKKNICHHCG
31	151-165	RKKNICHHCGAFNGT
32	156-170	CHHCGAFNGTVKKCG
33	161-175	AFNGTVKKCGLLKII
34	166-180	VKKCGLLKIIHEKYK
35	171-185	LLKIIHEKYKTNKKV
36	176-190	HEKYKTNKKVVDPIV
37	181-195	TNKKVVDPIVSNFLQ
38	186-200	VDPIVSNFLQSFETA
39	191-205	SNFLQSFETAIEHNK
40	196-210	SFETAIEHNKEVEPL
41	201-215	IEHNKEVEPLLGRAQ
42	206-220	EVEPLLGRAQENLNP
43	211-225	LGRAQENLNPLVVLN
44	216-230	ENLNPLVVLNLFKRI
45	221-235	LVVLNLFKRIPAEDV
46	226-240	LFKRIPAEDVPLLLM
47	231-245	PAEDVPLLLMNPEAG
48	236-250	PLLLMNPEAGKPSDL
49	241-255	NPEAGKPSDLILTRL

50	246-260	KPSDLILTRLLVPPL
51	251-265	ILTRLLVPPLCIRPS
52	256-270	LVPPLCIRPSVVSDL
53	261-275	CIRPSVVSDLKSGTN
54	266-280	VVSDLKSGTNEDDLT
55	271-285	KSGTNEDDLTMKLTE
56	276-290	EDDLTMKLTEIIFLN
57	281-295	MKLTEIIFLNDVIKK
58	286-300	IIFLNDVIKKHRISG
59	291-305	DVIKKHRISGAKTQM
60	296-310	HRISGAKTQMIMEDW
61	301-315	AKTQMIMEDWDFLQL
62	306-320	IMEDWDFLQLQCALY
63	311-325	DFLQLQCALYINSEL
64	316-330	QCALYINSELSGIPL
65	321-335	INSELSGIPLNMAPK
66	326-340	SGIPLNMAPKKWTRG
67	331-345	NMAPKKWTRGFVQRL
68	336-350	KWTRGFVQRLKGKQG
69	341-355	FVQRLKGKQGRFRGN
70	346-360	KGKQGRFRGNLSGKR
71	351-365	RFRGNLSGKRVDFSG
72	356-370	LSGKRVDFSGRTVIS
73	361-375	VDFSGRTVISPDPNL
74	366-380	RTVISPDPNLRIDEV
75	371-385	PDPNLRIDEVAVPVH
76	376-390	RIDEVAVPVHVAKIL
77	381-395	AVPVHVAKILTFPEK
78	386-400	VAKILTFPEKVNKAN
79	391-405	TFPEKVNKANINFLR
80	396-410	VNKANINFLRKLVQN
81	401-415	INFLRKLVQNGPEVH
82	406-420	KLVQNGPEVHPGANF
83	411-425	GPEVHPGANFIQQRH
84	416-430	PGANFIQQRHTQMKR
85	421-435	IQQRHTQMKRFLKYG
86	426-440	TQMKRFLKYGNREKM
87	431-445	FLKYGNREKMAQELK
88	436-450	NREKMAQELKYGDIV
89	441-455	AQELKYGDIVERHLI
90	446-460	YGDIVERHLIDGDVV
91	451-465	ERHLIDGDVVLFNRQ
92	456-470	DGDVVLFNRQPSLHK
93	461-475	LFNRQPSLHKLSIMA
94	466-480	PSLHKLSIMAHLARV
95	471-485	LSIMAHLARVKPHRT
96	476-490	HLARVKPHRTFRFNE
97	481-495	KPHRTFRFNECVCTP
98	486-500	FRFNECVCTPYNADF
99	491-505	CVCTPYNADFDGDEM
100	496-510	YNADFDGDEMNLHLP
101	501-515	DGDEMNLHLPQTEEA

102	506-520	NLHLPQTEEAKAEAL
103	511-525	QTEEAKAEALVLMGT
104	516-530	KAEALVLMGTKANLV
105	521-535	VLMGTKANLVTPRNG
106	526-540	KANLVTPRNGEPLIA
107	531-545	TPRNGEPLIAAIQDF
108	536-550	EPLIAAIQDFLTGAY
109	541-555	AIQDFLTGAYLLTLK
110	546-560	LTGAYLLTLKDTFFD
111	551-565	LLTLKDTFFDRAKAC
112	556-570	DTFFDRAKACQIIAS
113	561-575	RAKACQIIASILVGK
114	566-580	QIIASILVGKDEKIK
115	571-585	ILVGKDEKIKVRLPP
116	576-590	DEKIKVRLPPPTILK
117	581-595	VRLPPPTILKPVTLW
118	586-600	PTILKPVTLWTGKQI
119	591-605	PVTLWTGKQIFSVIL
120	596-610	TGKQIFSVILRPSDD
121	601-615	FSVILRPSDDNPVRA
122	606-620	RPSDDNPVRANLRTK
123	611-625	NPVRANLRTKGKQYC
124	616-630	NLRTKGKQYCGKGED
125	621-635	GKQYCGKGEDLCAND
126	626-640	GKGEDLCANDSYVTI
127	631-645	LCANDSYVTIQNSEL
128	636-650	SYVTIQNSELMSGSM
129	641-655	QNSELMSGSMDKGTL
130	646-660	MSGSMDKGTLGSGSK
131	651-665	DKGTLGSGSKNNIFY
132	656-670	GSGSKNNIFYILLRD
133	661-675	NNIFYILLRDWGQNE
134	666-680	ILLRDWGQNEAADAM
135	671-685	WGQNEAADAMSRLAR
136	676-690	AADAMSRLARLAPVY
137	681-695	SRLARLAPVYLSNRG
138	686-700	LAPVYLSNRGFSIGI
139	691-705	LSNRGFSIGIGDVTP
140	696-710	FSIGIGDVTPGQGLL
141	701-715	GDVTPGQGLLKAKYE
142	706-720	GQGLLKAKYELLNAG
143	711-725	KAKYELLNAGYKKCD
144	716-730	LLNAGYKKCDEYIEA
145	721-735	YKKCDEYIEALNTGK
146	726-740	EYIEALNTGKLQQQP
147	731-745	LNTGKLQQQPGCTAE
148	736-750	LQQQPGCTAEETLEA
149	741-755	GCTAEETLEALILKE
150	746-760	ETLEALILKELSVIR
151	751-765	LILKELSVIRDHAGS
152	756-770	LSVIRDHAGSACLRE
153	761-775	DHAGSACLRELDKSN

154	766-780	ACLRELDKSNSPLTM
155	771-785	LDKSNSPLTMALCGS
156	776-790	SPLTMALCGSKGSFI
157	781-795	ALCGSKGSFINISQM
158	786-800	KGSFINISQMIACVG
159	791-805	NISQMIACVGQQAIS
160	796-810	IACVGQQAISGSRVP
161	801-815	QQAISGSRVPDGFEN
162	806-820	GSRVPDGFENRSLPH
163	811-825	DGFENRSLPHFEKHS
164	816-830	RSLPHFEKHSKLPAA
165	821-835	FEKHSKLPAAKGFVA
166	826-840	KLPAAKGFVANSFYS
167	831-845	KGFVANSFYSGLTPT
168	836-850	NSFYSGLTPTEFFFH
169	841-855	GLTPTEFFFHTMAGR
170	846-860	EFFFHTMAGREGLVD
171	851-865	TMAGREGLVDTAVKT
172	856-870	EGLVDTAVKTAETGY
173	861-875	TAVKTAETGYMQRRL
174	866-880	AETGYMQRRLVKSLE
175	871-885	MQRRLVKSLEDLCSQ
176	876-890	VKSLEDLCSQYDLTV
177	881-895	DLCSQYDLTVRSSTG
178	886-900	YDLTVRSSTGDIIQF
179	891-905	RSSTGDIIQFIYGGD
180	896-910	DIIQFIYGGDGLDPA
181	901-915	IYGGDGLDPAAMEGK
182	906-920	GLDPAAMEGKDEPLE
183	911-925	AMEGKDEPLEFKRVL
184	916-930	DEPLEFKRVLDNIKA
185	921-935	FKRVLDNIKAVFPCP
186	926-940	DNIKAVFPCPSEPAL
187	931-945	VFPCPSEPALSKNEL
188	936-950	SEPALSKNELILTTE
189	941-955	SKNELILTTESIMKK
190	946-960	ILTTESIMKKSEFLC
191	951-965	SIMKKSEFLCKYMRA
192	956-970	SEFLCKYMRAQMEPG
193	961-975	KYMRAQMEPGSAVGA
194	966-980	QMEPGSAVGALCAQS
195	971-985	SAVGALCAQSIGEPG
196	976-990	LCAQSIGEPGTQMTL
197	981-995	IGEPGTQMTLKTFHF
198	986-1000	TQMTLKTFHFAGVAS
199	991-1005	KTFHFAGVASMNITL
200	996-1010	AGVASMNITLGVPRI
201	1001-1015	MNITLGVPRIKEIIN
202	1006-1020	GVPRIKEIINASKAI
203	1011-1025	KEIINASKAICQDSF
204	1016-1130	ASKAICQDSFLQEIK
205	1021-1035	CQDSFLQEIKKFIKG

206	1026-1040	LQEIKKFIKGVSEKI
207	1031-1045	KFIKGVSEKIKKTRD
208	1036-1050	VSEKIKKTRDKYGIN
209	1041-1055	KKTRDKYGINDNGTT
210	1046-1060	KYGINDNGTTEPRVL
211	1051-1065	DNGTTEPRVLYQLDR
212	1056-1070	EPRVLYQLDRITPTQ
213	1061-1075	YQLDRITPTQVEKFL
214	1066-1080	ITPTQVEKFLETCRD
215	1071-1085	VEKFLETCRDSTPII
216	1076-1090	ETCRDSTPIITAQLD
217	1081-1095	STPIITAQLDKDDDA
218	1086-1100	TAQLDKDDDADYARL
219	1091-1105	KDDDADYARLVKGRI
220	1096-1110	DYARLVKGRIEKTLL
221	1101-1115	VKGRIEKTLLGEISE
222	1106-1120	EKTLLGEISEYIEEV
223	1111-1125	GEISEYIEEVFLPDD
224	1116-1130	YIEEVFLPDDCFILV
225	1121-1135	FLPDDCFILVKLSLE
226	1126-1140	CFILVKLSLERIRLL
227	1131-1145	KLSLERIRLLRLEVN
228	1136-1150	RIRLLRLEVNAETVR
229	1141-1155	RLEVNAETVRYSICT
230	1146-1160	AETVRYSICTSKLRV
231	1151-1165	YSICTSKLRVKPGDV
232	1156-1170	SKLRVKPGDVAVHGE
233	1161-1175	KPGDVAVHGEAVVCV
234	1166-1180	AVHGEAVVCVTPREN
235	1171-1185	AVVCVTPRENSKSSM
236	1176-1190	TPRENSKSSMYYVLQ
237	1181-1195	SKSSMYYVLQFLKED
238	1186-1200	YYVLQFLKEDLPKVV
239	1191-1205	FLKEDLPKVVVQGIP
240	1196-1210	LPKVVVQGIPEVSRA
241	1201-1215	VQGIPEVSRAVIHID
242	1206-1220	EVSRAVIHIDEQSGK
243	1211-1225	VIHIDEQSGKEKYKL
244	1216-1230	EQSGKEKYKLLVEGD
245	1221-1235	EKYKLLVEGDNLRAV
246	1226-1240	LVEGDNLRAVMATHG
247	1231-1245	NLRAVMATHGVKGTR
248	1236-1250	MATHGVKGTRTTSNN
249	1241-1255	VKGTRTTSNNTYEVE
250	1246-1260	TTSNNTYEVEKTLGI
251	1251-1265	TYEVEKTLGIEAART
252	1256-1270	KTLGIEAARTTIINE
253	1261-1275	EAARTTIINEIQYTM
254	1266-1280	TIINEIQYTMVNHGM
255	1271-1285	IQYTMVNHGMSIDRR
256	1276-1290	VNHGMSIDRRHVMLL
257	1281-1295	SIDRRHVMLLSDLMT

258	1286-1300	HVMLLSDLMTYKGEV
259	1291-1305	SDLMTYKGEVLGITR
260	1296-1310	YKGEVLGITRFGLAK
261	1301-1315	LGITRFGLAKMKESV
262	1306-1320	FGLAKMKESVLMLAS
263	1311-1325	MKESVLMLASFEKTA
264	1316-1330	LMLASFEKTADHLFD
265	1321-1335	FEKTADHLFDAAYFG
266	1326-1340	DHLFDAAYFGQKDSV
267	1331-1345	AAYFGQKDSVCGVSE
268	1336-1350	QKDSVCGVSECIIMG
269	1341-1355	CGVSECIIMGIPMNI
270	1346-1360	CIIMGIPMNIGTGLF
271	1351-1365	IPMNIGTGLFKLLHK
272	1356-1370	GTGLFKLLHKADRDP
273	1361-1375	KLLHKADRDPNPPKR
274	1366-1380	ADRDPNPPKRPLIFD
275	1371-1385	NPPKRPLIFDTNEFH
276	1376-1390	PLIFDTNEFHIPLVT
I104T mutant #1	91-105	PCFHVGYFRAVIGTL
I104T mutant #2	96-110	GYFRAVIGTLQMICK
I104T mutant #3	101-115	VIGTLQMICKTCCHI
E1072Q mutant #1	1061-1075	YQLDRITPTQVQKFL
E1072Q mutant #2	1066-1080	ITPTQVQKFLETCRD
E1072Q mutant #3	1071-1085	VQKFLETCRDSTPII
K1365N #1	1351-1365	IPMNIGTGLFKLLHN
K1365N #2	1356-1370	GTGLFKLLHNADRDP
K1365N #3	1361-1375	KLLHNADRDPNPPKR

Table S5. MHC types of the patients with *RPOL3A* mutations

Patient #	MHC
SCL-2	A 1/3
	B 7/8
	C 7/7
	DRB1 3/15
	DQB1 2/6
SCL-4	A 2/26
	B 35/56
	C 1/4
	DRB1 1/11
	DQB1 3/5
SCL-42	A 1/29
	B 37/44
	C 6/16
	DRB1 7/10
	DQB1 2/5

Table S6: Patient-specific MHC class I and class II peptides with highest predicted binding affinity

Patient	Mutation	HLA Allele	wt peptide	IC50	mut peptide	IC50	Prediction Method
SCL-2	E1072Q	A*0101	TLGVPRIK**E**I	33377	TLGVPRIK**Q**I	32553	ANN
		A*0301	RIK**E**IINASK	**78**	RIK**Q**IINASK	**58**	ANN
		B*0702	VPRIK**E**IINAS	1407	VPRIK**Q**IINAS	953	CombLib
		B*0801	VPRIK**E**II	759	VPRIK**Q**II	447	CombLib
		C*0701	ITLGVPRIK**E**I	1539	ITLGVPRIK**Q**I	2221	ANN
		C*0702	SMNITLGVPRIK**E**I	11367	SMNITLGVPRIK**Q**I	11587	ANN
		DR*0301	GVPRIK**E**IINASKAIST	8235	GVPRIK**Q**IINASKAIST	6854	SMM
		DR*1501	GVPRIK**E**IINASKAIST	591	GVPRIK**Q**IINASKAIST	416	SMM
SCL-4	K1365N	A*0201	GLFKLLH**K**A	**46**	GLFKLLH**N**A	**27**	ANN
		A*2601	GTGLFKLLH**K**	31605	GTGLFKLLH**N**	33271	ANN
		B*3501	FKLLH**K**AD	15618	FKLLH**N**AD	4544	SMM
		B*5601	TGLFKLLH**K**A	24218	TGLFKLLH**N**A	21971	NetMHCpan
		C*0102	**K**ADRDPNPPKRPL	13637	**N**ADRDPNPPKRPL	21533	NetMHCpan
		C*0401	KLLH**K**ADRDPNPPK	7973	KLLH**N**ADRDPNPPK	7794	ANN
		DR*0101	IGTGLFKLLH**K**ADRDPN	**69**	IGTGLFKLLH**N**ADRDPN	**4**	CombLib
		DR*1101	IGTGLFKLLH**K**ADRDPN	**26**	IGTGLFKLLH**N**ADRDPN	**113**	SMM
SCL-42	I104T	A*0101	VIG**I**LQMI	21734	VIG**T**LQMI	21971	NetMHCpan
		A*2902	YFRAVIG**I**LQM	1158	YFRAVIG**T**LQM	743	NetMHCpan
		B*3701	FRAVIG**I**L	3417	FRAVIG**T**L	3237	NetMHCpan
		B*4403	AVIG**I**LQM	22452	AVIG**T**LQM	21501	NetMHCpan
		C*0602	FRAVIG**I**LQMI	**53**	FRAVIG**T**LQMI	**40**	ANN
		C*1601	YFRAVIG**I**LQM	822	YFRAVIG**T**LQM	458	NetMHCpan
		DR*0701	FHVGYFRAVIG**I**LQMI	**0**	FHVGYFRAVIG**T**LQMI	**0.5**	CombLib
		DR*1001	FHVGYFRAVIG**I**LQMI	**54**	FHVGYFRAVIG**T**LQMI	**21**	NetMHCpan

Moderate to high affinity peptides are highlighted in red

ANN= artificial neural network

CombLib = scoring matrices derived from combinatorial peptide libraries

SMM = stabilized matrix method

References and Notes

1. A. Gabrielli, E. V. Avvedimento, T. Krieg, Scleroderma. *N. Engl. J. Med.* **360**, 1989–2003 (2009). Medline doi:10.1056/NEJMra0806188

2. M. L. Harris, A. Rosen, Autoimmunity in scleroderma: The origin, pathogenetic role, and clinical significance of autoantibodies. *Curr. Opin. Rheumatol.* **15**, 778–784 (2003). Medline doi:10.1097/00002281-200311000-00016

3. A. A. Shah, A. Rosen, L. Hummers, F. Wigley, L. Casciola-Rosen, Close temporal relationship between onset of cancer and scleroderma in patients with RNA polymerase I/III antibodies. *Arthritis Rheum.* **62**, 2787–2795 (2010). Medline doi:10.1002/art.27549

4. A. A. Shah, A. Rosen, Cancer and systemic sclerosis: Novel insights into pathogenesis and clinical implications. *Curr. Opin. Rheumatol.* **23**, 530–535 (2011). Medline doi:10.1097/BOR.0b013e32834a5081

5. Materials and methods are available as supplementary materials on *Science* Online.

6. A. G. Knudson, Hereditary cancer: Two hits revisited. *J. Cancer Res. Clin. Oncol.* **122**, 135–140 (1996). Medline doi:10.1007/BF01366952

7. Y. Kim, J. Ponomarenko, Z. Zhu, D. Tamang, P. Wang, J. Greenbaum, C. Lundegaard, A. Sette, O. Lund, P. E. Bourne, M. Nielsen, B. Peters, Immune epitope database analysis resource. *Nucleic Acids Res.* **40** (W1), W525–W530 (2012). Medline doi:10.1093/nar/gks438

8. P. Wang, J. Sidney, C. Dow, B. Mothé, A. Sette, B. Peters, A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLOS Comput. Biol.* **4**, e1000048 (2008). Medline doi:10.1371/journal.pcbi.1000048

9. P. Wang, J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, B. Peters, Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* **11**, 568 (2010). Medline doi:10.1186/1471-2105-11-568

10. Z. C. Ding, G. Zhou, Cytotoxic chemotherapy and CD4+ effector T cells: An emerging alliance for durable antitumor effects. *Clin. Dev. Immunol.* **2012**, 890178 (2012). Medline doi:10.1155/2012/890178

11. I. Mellman, G. Coukos, G. Dranoff, Cancer immunotherapy comes of age. *Nature* **480**, 480–489 (2011). Medline doi:10.1038/nature10673

12. P. K. Chattopadhyay, J. Yu, M. Roederer, A live-cell assay to detect antigen-specific CD4+ T cells with diverse cytokine profiles. *Nat. Med.* **11**, 1113–1117 (2005). Medline doi:10.1038/nm1293

13. M. Frentsch, O. Arbach, D. Kirchhoff, B. Moewes, M. Worm, M. Rothe, A. Scheffold, A. Thiel, Direct access to CD4+ T cells specific for defined antigens according to CD154 expression. *Nat. Med.* **11**, 1118–1124 (2005). Medline doi:10.1038/nm1292

14. G. T. Nepom, J. H. Buckner, E. J. Novak, S. Reichstetter, H. Reijonen, J. Gebe, R. Wang, E. Swanson, W. W. Kwok, HLA class II tetramers: Tools for direct analysis of antigen-specific CD4+ T cells. *Arthritis Rheum.* **46**, 5–12 (2002). Medline doi:10.1002/1529-0131(200201)46:1<5::AID-ART10063>3.0.CO;2-S

15. K. J. Jackson, M. J. Kidd, Y. Wang, A. M. Collins, The Shape of the lymphocyte receptor repertoire: Lessons from the B cell receptor. *Front. Immunol.* **4**, 263 (2013).

16. M. G. McHeyzer-Williams, M. M. Davis, Antigen-specific development of primary and memory T cells in vivo. *Science* **268**, 106–111 (1995). Medline doi:10.1126/science.7535476

17. R. Buchbinder, A. Forbes, S. Hall, X. Dennett, G. Giles, Incidence of malignant disease in biopsy-proven inflammatory myopathy. A population-based cohort study. *Ann. Intern. Med.* **134**, 1087–1095 (2001). Medline doi:10.7326/0003-4819-134-12-200106190-00008

18. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: Mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39** (Database), D945–D950 (2011). Medline doi:10.1093/nar/gkq929

19. H. A. Doyle, M. J. Mamula, Posttranslational modifications of self-antigens. *Ann. N. Y. Acad. Sci.* **1050**, 1–9 (2005). Medline doi:10.1196/annals.1313.001

20. M. J. Mamula, R. J. Gee, J. I. Elliott, A. Sette, S. Southwood, P. J. Jones, P. R. Blier, Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.* **274**, 22321–22327 (1999). Medline doi:10.1074/jbc.274.32.22321

21. R. H. Lin, M. J. Mamula, J. A. Hardin, C. A. Janeway Jr., Induction of autoreactive B cells allows priming of autoreactive T cells. *J. Exp. Med.* **173**, 1433–1439 (1991). Medline doi:10.1084/jem.173.6.1433

22. M. J. Mamula, S. Fatenejad, J. Craft, B cells process and present lupus autoantigens that initiate autoimmune T cell responses. *J. Immunol.* **152**, 1453–1461 (1994). Medline

23. M. L. Albert, R. B. Darnell, Paraneoplastic neurological degenerations: Keys to tumour immunity. *Nat. Rev. Cancer* **4**, 36–44 (2004). Medline doi:10.1038/nrc1255

24. C. Gaudin, F. Kremer, E. Angevin, V. Scott, F. Triebel, A hsp70-2 mutation recognized by CTL on a human renal cell carcinoma. *J. Immunol.* **162**, 1730–1738 (1999). Medline

25. R. F. Wang, X. Wang, A. C. Atwood, S. L. Topalian, S. A. Rosenberg, Cloning genes encoding MHC class II-restricted antigens: Mutated CDC27 as a tumor antigen. *Science* **284**, 1351–1354 (1999). Medline doi:10.1126/science.284.5418.1351

26. T. Wölfel, M. Hauer, J. Schneider, M. Serrano, C. Wölfel, E. Klehmann-Hieb, E. De Plaen, T. Hankeln, K. H. Meyer zum Büschenfelde, D. Beach, A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science* **269**, 1281–1284 (1995). Medline doi:10.1126/science.7652577

27. M. E. Engelhorn, J. A. Guevara-Patiño, G. Noffz, A. T. Hooper, O. Lou, J. S. Gold, B. J. Kappel, A. N. Houghton, Autoimmunity and tumor immunity induced by immune responses to mutations in self. *Nat. Med.* **12**, 198–206 (2006). Medline doi:10.1038/nm1363

28. L. Casciola-Rosen, K. Nagaraju, P. Plotz, K. Wang, S. Levine, E. Gabrielson, A. Corse, A. Rosen, Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. *J. Exp. Med.* **201**, 591–601 (2005). Medline doi:10.1084/jem.20041367

29. R. D. Schreiber, L. J. Old, M. J. Smyth, Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion. *Science* **331**, 1565–1570 (2011). Medline doi:10.1126/science.1203486

30. S. Bernatsky, J. F. Boivin, L. Joseph, R. Rajan, A. Zoma, S. Manzi, E. Ginzler, M. Urowitz, D. Gladman, P. R. Fortin, M. Petri, S. Edworthy, S. Barr, C. Gordon, S. C. Bae, J. Sibley, D. Isenberg, A. Rahman, C. Aranow, M. A. Dooley, K. Steinsson, O. Nived, G. Sturfelt, G. Alarcón, J. L. Senécal, M. Zummer, J. Hanly, S. Ensworth, J. Pope, H. El-Gabalawy, T. McCarthy, Y. St Pierre, R. Ramsey-Goldman, A. Clarke, An international cohort study of cancer in systemic lupus erythematosus. *Arthritis Rheum.* **52**, 1481–1490 (2005). Medline doi:10.1002/art.21029

31. E. Tatsis, E. Reinhold-Keller, K. Steindorf, A. C. Feller, W. L. Gross, Wegener's granulomatosis associated with renal cell carcinoma. *Arthritis Rheum.* **42**, 751–756 (1999). Medline doi:10.1002/1529-0131(199904)42:4<751::AID-ANR19>3.0.CO;2-D

32. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz Jr., K. W. Kinzler, Cancer genome landscapes. *Science* **339**, 1546–1558 (2013). Medline doi:10.1126/science.1235122

33. M. Burnet, Cancer—A Biological Approach. *BMJ* **1**, 841–847 (1957). Medline doi:10.1136/bmj.1.5023.841

34. F. M. Burnet, The concept of immunological surveillance. *Prog. Exp. Tumor Res.* **13**, 1–27 (1970). Medline

35. S. A. Quezada, K. S. Peggs, Exploiting CTLA-4, PD-1 and PD-L1 to reactivate the host immune response against cancer. *Br. J. Cancer* **108**, 1560–1565 (2013). Medline doi:10.1038/bjc.2013.117

36. M. DuPage, C. Mazumdar, L. M. Schmidt, A. F. Cheung, T. Jacks, Expression of tumour-specific antigens underlies cancer immunoediting. *Nature* **482**, 405–409 (2012). Medline doi:10.1038/nature10803

37. H. Matsushita, M. D. Vesely, D. C. Koboldt, C. G. Rickert, R. Uppaluri, V. J. Magrini, C. D. Arthur, J. M. White, Y. S. Chen, L. K. Shea, J. Hundal, M. C. Wendl, R. Demeter, T. Wylie, J. P. Allison, M. J. Smyth, L. J. Old, E. R. Mardis, R. D. Schreiber, Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012). Medline doi:10.1038/nature10755

38. A. T. Masi, Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee, Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum.* **23**, 581–590 (1980). Medline doi:10.1002/art.1780230510

39. P. J. Clements *et al*., Skin thickness score in systemic sclerosis: An assessment of interobserver variability in 3 independent studies. *J. Rheumatol.* **20**, 1892–1896 (1993). Medline

40. E. C. LeRoy, C. Black, R. Fleischmajer, S. Jablonska, T. Krieg, T. A. Medsger Jr., N. Rowell, F. Wollheim, Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J. Rheumatol.* **15**, 202–205 (1988). Medline

41. J. Wu, H. Matthaei, A. Maitra, M. Dal Molin, L. D. Wood, J. R. Eshleman, M. Goggins, M. I. Canto, R. D. Schulick, B. H. Edil, C. L. Wolfgang, A. P. Klein, L. A. Diaz Jr., P. J. Allen, C. M. Schmidt, K. W. Kinzler, N. Papadopoulos, R. H. Hruban, B. Vogelstein, Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development. *Sci. Transl. Med.* **3**, 92ra66 (2011). Medline doi:10.1126/scitranslmed.3002543

42. D. S. Herman, G. K. Hovingh, O. Iartchouk, H. L. Rehm, R. Kucherlapati, J. G. Seidman, C. E. Seidman, Filter-based hybridization capture of subgenomes enables resequencing and copy-number detection. *Nat. Methods* **6**, 507–510 (2009). Medline doi:10.1038/nmeth.1343

43. I. Kinde, J. Wu, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9530–9535 (2011). Medline doi:10.1073/pnas.1105422108

44. L. A. Casciola-Rosen, A. F. Pluta, P. H. Plotz, A. E. Cox, S. Morris, F. M. Wigley, M. Petri, A. C. Gelber, A. Rosen, The DNA mismatch repair enzyme PMS1 is a myositis-specific autoantigen. *Arthritis Rheum.* **44**, 389–396 (2001). Medline doi:10.1002/1529-0131(200102)44:2<389::AID-ANR58>3.0.CO;2-R