ARTHRITIS & RHEUMATOLOGY
Vol. 66, No. 10, October 2014, pp 2706–2715
DOI 10.1002/art.38754
© 2014, American College of Rheumatology

Barcode-Enabled Sequencing of Plasmablast Antibody Repertoires in Rheumatoid Arthritis

Yann-Chong Tan,[1] Sarah Kongpachith,[1] Lisa K. Blum,[1] Chia-Hsin Ju,[1] Lauren J. Lahey,[1] Daniel R. Lu,[1] Xiaoyong Cai,[1] Catriona A. Wagner,[1] Tamsin M. Lindstrom,[2] Jeremy Sokolove,[1] and William H. Robinson[1]

Objective. **A hallmark of rheumatoid arthritis (RA) is the production of autoantibodies, including anti–citrullinated protein antibodies (ACPAs). Nevertheless, the specific targets of these autoantibodies remain incompletely defined. During an immune response, B cells specific for the inciting antigen(s) are activated and differentiate into plasmablasts, which are released into the blood. We undertook this study to sequence the plasmablast antibody repertoire to define the targets of the active immune response in RA.**

Methods. **We developed a novel DNA barcoding method to sequence the cognate heavy- and light-chain pairs of antibodies expressed by individual blood plasmablasts in RA. The method uses a universal 5′ adapter that enables full-length sequencing of the antibodies' variable regions and recombinant expression of the paired antibody chains. The sequence data sets were bioinformatically analyzed to generate phylogenetic trees that identify clonal families of antibodies sharing heavy- and light-chain VJ sequences. Representative antibodies were expressed, and their binding properties were characterized using anti–cyclic citrullinated peptide 2 (anti–CCP-2) enzyme-linked immunosorbent assay (ELISA) and antigen microarrays.**

Results. **We used our sequencing method to generate phylogenetic trees representing the antibody repertoires of peripheral blood plasmablasts from 4 individuals with anti-CCP+ RA, and recombinantly expressed 14 antibodies that were either "singleton" antibodies or representative of clonal antibody families. Anti–CCP-2 ELISA identified 4 ACPAs, and antigen microarray analysis identified ACPAs that differentially targeted epitopes on α-enolase, citrullinated fibrinogen, and citrullinated histone H2B.**

Conclusion. **Our data provide evidence that autoantibodies targeting α-enolase, citrullinated fibrinogen, and citrullinated histone H2B are produced by the ongoing activated B cell response in, and thus may contribute to the pathogenesis of, RA.**

Rheumatoid arthritis (RA) is a common autoimmune synovitis associated with the production of autoantibodies, including rheumatoid factor (RF) and anti–citrullinated protein antibodies (ACPAs) (1–3). ACPAs target proteins that have undergone citrullination (1,2,4), a posttranslational modification that converts peptidylarginine to peptidylcitrulline. Presently, such antibodies are detected in the clinic using the anti–cyclic citrullinated peptide (anti-CCP) assay (1,5).

Supported by the NIH (National Heart, Lung, and Blood Institute Proteomics Center grant N01-HV-00242, National Institute of Arthritis and Musculoskeletal and Skin Diseases grant R01-AR-063676, and grant U01-AI-101981 to Dr. Robinson) and by the Northern California Chapter of the Arthritis Foundation (Center of Excellence funding to Dr. Robinson). Dr. Tan is recipient of an A*STAR National Science Scholarship (PhD) Fellowship from the Agency for Science, Technology, and Research, Singapore.

[1]Yann-Chong Tan, PhD (current address: Atreca, Inc., San Carlos, California), Sarah Kongpachith, BS, Lisa K. Blum, PhD, Chia-Hsin Ju, PhD, Lauren J. Lahey, BS, Daniel R. Lu, BA, Xiaoyong Cai, BS, Catriona A. Wagner, BS, Jeremy Sokolove, MD, William H. Robinson, MD, PhD: Stanford University, Stanford, California, and VA Palo Alto Health Care System, Palo Alto, California. [2]Tamsin M. Lindstrom, PhD: Stanford University, Stanford, California.

Dr. Tan and Ms Kongpachith contributed equally to this work.

Dr. Tan owns stock or stock options in Atreca, Inc. Drs. Sokolove and Robinson have received consulting fees from Atreca, Inc. (more than $10,000) and own stock or stock options in Atreca. Drs. Tan, Sokolove, and Robinson are listed on patents (for a method to sequence antibody repertoires) owned by Stanford University and licensed to Atreca, Inc., for which they receive royalties.

Address correspondence to William H. Robinson, MD, PhD, Division of Immunology and Rheumatology, Stanford University School of Medicine, CCSR 4135, 269 Campus Drive, Stanford, CA 94305. E-mail: wrobins@stanford.edu.

Submitted for publication October 11, 2013; accepted in revised form June 17, 2014.

The anti-CCP assay uses as detector antigens a mixture of cyclized, citrulline-substituted peptides derived from filaggrin, a protein not expressed in joint tissue (4), and therefore does not identify the bona fide, in vivo targets of ACPAs. Thus, uncovering the specificity of the ACPAs that contribute to the pathogenesis of RA remains a critical challenge (1,4). To gain further insights into the specificity of the autoantibody response in RA, we developed and applied a DNA barcoding method to sequence the cognate heavy- and light-chain pairs of antibodies expressed by individual peripheral blood plasmablasts derived from individuals with anti-CCP autoantibody–positive RA (anti-CCP+ RA).

Antibodies are comprised of heavy and light chains, each containing an antigen-binding domain that is generated by the recombination, junctional diversification, and somatic hypermutation of variable (V), joining (J), and/or diversity (D) gene segments. Several methods exist for the profiling and isolation of native human antibodies, including reverse transcriptase–polymerase chain reaction (RT-PCR) of single B cells (6–12). However, RT-PCR of single B cells is laborious, requiring Sanger sequencing of each B cell, followed by production and screening of a large number of antibodies (7–10).

Two recently developed methods have begun to address the issue of heavy- and light-chain pairing on a larger scale. One method involves the deposition of single B cells in high-density microwell plates followed by the sequencing of the third complementarity-determining region (CDR3) of their antibody genes (13). The other involves mass spectrometric analysis of circulating antibodies against specific antigens followed by combinatorial expression and screening of possible heavy- and light-chain pairs (14). Although useful tools, these methods have shortcomings: they use V-gene–specific primers that fail to amplify all immunoglobulin sequences (especially mutated 5′-end sequences that have arisen from extensive somatic hypermutation that may confer interesting biologic properties), they cannot distinguish between sequencing errors and closely related sequences that have arisen through somatic hypermutation, they sequence only the CDR3 regions and thus cannot accurately identify clonal families of antibodies that share other heavy- and light-chain variable region sequences, they cannot accurately determine the size of clonal antibody families, and they require PCR cloning and Sanger sequencing to deliver complete V-region sequences.

To overcome these shortcomings, we developed a novel approach that combines high-throughput sequencing with DNA barcode–enabled pairing of cognate heavy- and light-chain antibody sequences expressed by individual B cells (15). We focus our analysis on the antibodies expressed by peripheral blood plasmablasts; these antibody-producing cells arise from both the naive and the memory B cells activated in an immune response (11,16–18), and their antibody repertoires therefore provide a comprehensive snapshot of the ongoing antibody response. By bioinformatically analyzing the resulting sequence data sets, we can generate phylogenetic trees of the antibody responses and rationally select key antibodies for cloning, expression, and characterization of their binding and functional properties.

To demonstrate the power of our DNA barcoding method and to further study the specificities of the autoantibody response in RA, we applied our DNA barcoding method to characterize the autoantibody response of peripheral blood plasmablasts derived from individuals with anti-CCP+ RA. Phylogenetic trees representing the plasmablast antibody repertoires revealed the production of affinity-matured clonal families of antibodies in anti-CCP+ RA, and these trees were used to guide selection of representative antibodies for recombinant expression. We demonstrate that recombinant antibodies derived from RA plasmablasts bind to CCP in the anti–CCP-2 enzyme-linked immunosorbent assay (ELISA) and to putative, citrullinated autoantigens—including α-enolase, citrullinated fibrinogen, and citrullinated histone H2B–contained on RA antigen microarrays.

MATERIALS AND METHODS

Human samples. Samples were collected after obtaining informed consent and according to human subject protocols approved by the Investigational Review Board at Stanford University. For RA samples, we collected blood from individuals who met at least 4 of the 7 American College of Rheumatology 1987 classification criteria for RA (19).

Assessment of ACPA production by bulk cultured plasmablasts. Peripheral blood mononuclear cells (PBMCs) were stained, and IgG+ plasmablasts were bulk sorted by gating for CD19+CD20−CD27+CD38++IgA−IgM− cells. Plasmablasts were then cultured in complete RPMI medium for 7 days in 96-well plates at a density of 150,000 cells/well. We used the supernatants at a 1:1 dilution in RA antigen microarray assays.

Single-cell sorting of plasmablasts. PBMCs were stained, and IgG+ plasmablasts were single-cell sorted into 96-well plates based on gating for CD19+CD20−CD27+CD38++IgA−IgM− cells (11).

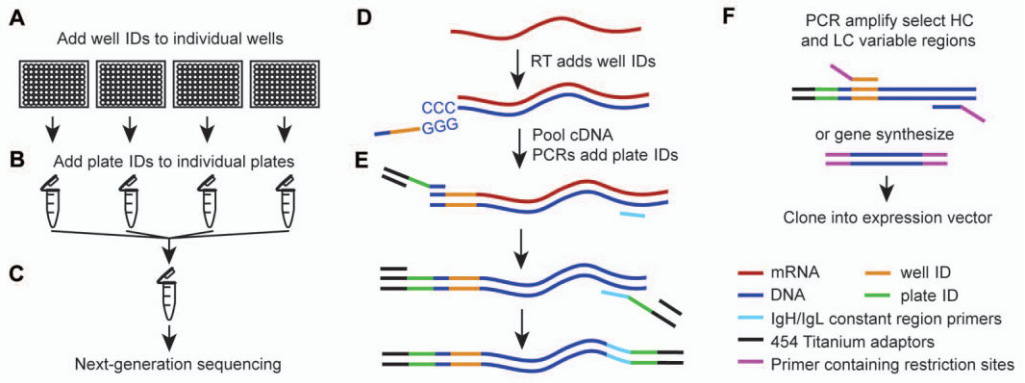

Figure 1. Schematic representation of a DNA barcode–enabled method for high-throughput sequencing and recombinant production of endogenous antibodies. **A–C,** Overview of the method. **A,** Plasmablasts are single-cell sorted into 96-well plates, and cDNA is generated by reverse transcription (RT), during which a unique well ID barcode is added to all of the cDNA in an individual well. **B,** All of the barcoded cDNA from one plate is pooled and tagged with a unique plate ID. **C,** The cDNAs, which are now double-barcoded with well IDs and plate IDs, are pooled and sequenced using next-generation 454 sequencing. **D** and **E,** Specifics for barcoding and library preparation. **D,** During reverse transcription, an adapter containing a unique well ID barcode is incorporated at the 3′ end of the first-strand cDNA by using the 3′-tailing activity of a thermally stable RNase H^- reverse transcriptase. **E,** Plate IDs and 454 Titanium Primer A are added during polymerase chain reaction (PCR) amplification by using a barcoded adapter. A further round of nested PCR is performed, adding 3′ plate IDs and 454 Titanium primer B. **F,** For cloning and expression of an antibody, either PCR or gene synthesis is used. For PCR, a 5′ primer specific to a particular well ID is used to amplify only the desired pair of heavy chain (HC) and light chain (LC) cDNA from the pooled cDNA of the appropriate plate. Standard molecular biology is used to insert the antibody V(D)J or VJ sequences in frame into an expression vector containing the appropriate antibody constant region.

RT-PCR with DNA barcodes. We performed reverse transcription with a Moloney murine leukemia virus H reverse transcriptase that has 3′-tailing (20) and template-switching activity in order to synthesize complementary DNA (cDNA) and tag it with a "well ID barcode" sequence unique to each well (and hence unique to each plasmablast cell) (Figures 1A and D). We pooled the well ID–tagged cDNAs from each plate and tagged them with plate ID barcodes by using nested PCR and primers specific to the constant regions of the heavy and light chains. We used Phusion Hot Start II DNA polymerase (NEB/Fermentas) for both the first PCR (PCR1) and the nested PCR (PCR2). We recently described the primers and adapters used (15).

Sample preparation for 454 sequencing. For 454 sequencing, we pooled the amplified DNA, gel purified them, added Lib-L sequencing adapters, and then purified the amplicons with AMPure XP beads (Beckman Coulter). We determined DNA concentrations by using PicoGreen DNA assay kits (Invitrogen) and sent the amplicons to Roche, where they were sequenced by emulsion PCR at 1 copy per bead before being subjected to 454 Titanium sequencing.

Compound barcode assignment and assembly of sequences. Sequencing data were analyzed by 454 GS FLX data analysis software. Sff output files from 454 sequencing, containing sequences and quality scores for each nucleotide, were read into Python by using the Biopython package, and sequences were grouped and parsed into separate sff files on the basis of their compound ID (plate ID + well ID). We used Newbler 2.6 to assemble forward reads into consensus sequences by using the "-cdna", "-ud" and "-urt" options, with a minimum threshold of 8 reads. Where multiple assemblies occurred per well, which is common with oversampling (>100 reads), an assembly was accepted if it contained >50% of all reads in a well, and/or was 3 times more abundant than the next read. Otherwise, we assumed that the well contained more than one plasmablast, and we disregarded the sequence reads from that well.

V(D)J and clonal family assignment. We analyzed heavy- and light-chain sequences with IMGT HighV-QUEST (21), software that compares an antibody-chain sequence to a database of known alleles of germline sequences and predicts which germline alleles the antibody uses, how the germline sequences were recombined, and the number of mutations the antibody has relative to the germline sequence (broken down into "all mutations," "silent mutations," and "non-silent mutations"). Clonal family antibodies were defined as antibodies with shared heavy-chain VJ and light-chain VJ sequence usage and shared mutations as compared to the corresponding germline sequence. Heavy-chain D alleles were not used in the clonal family assignments, as their short length (<20 AA) and high mutation rates can make D-allele calling inaccurate. We also identified "singleton" antibodies, i.e., antibodies using heavy-chain and/or light-chain VJ sequences that were not used by any other antibody in an individual's plasmablast antibody repertoire.

Clustering of sequences. Heavy- and light-chain sequences were clustered based on their V-region sequences, binned according to their heavy-chain V-gene usage, and then concatenated and aligned with Muscle (22). They were clustered with PhyML (23) maximum-likelihood clustering, and rooted by their germline heavy-chain V gene. Each V-gene phylogenetic tree was then arranged by heavy-chain V-gene

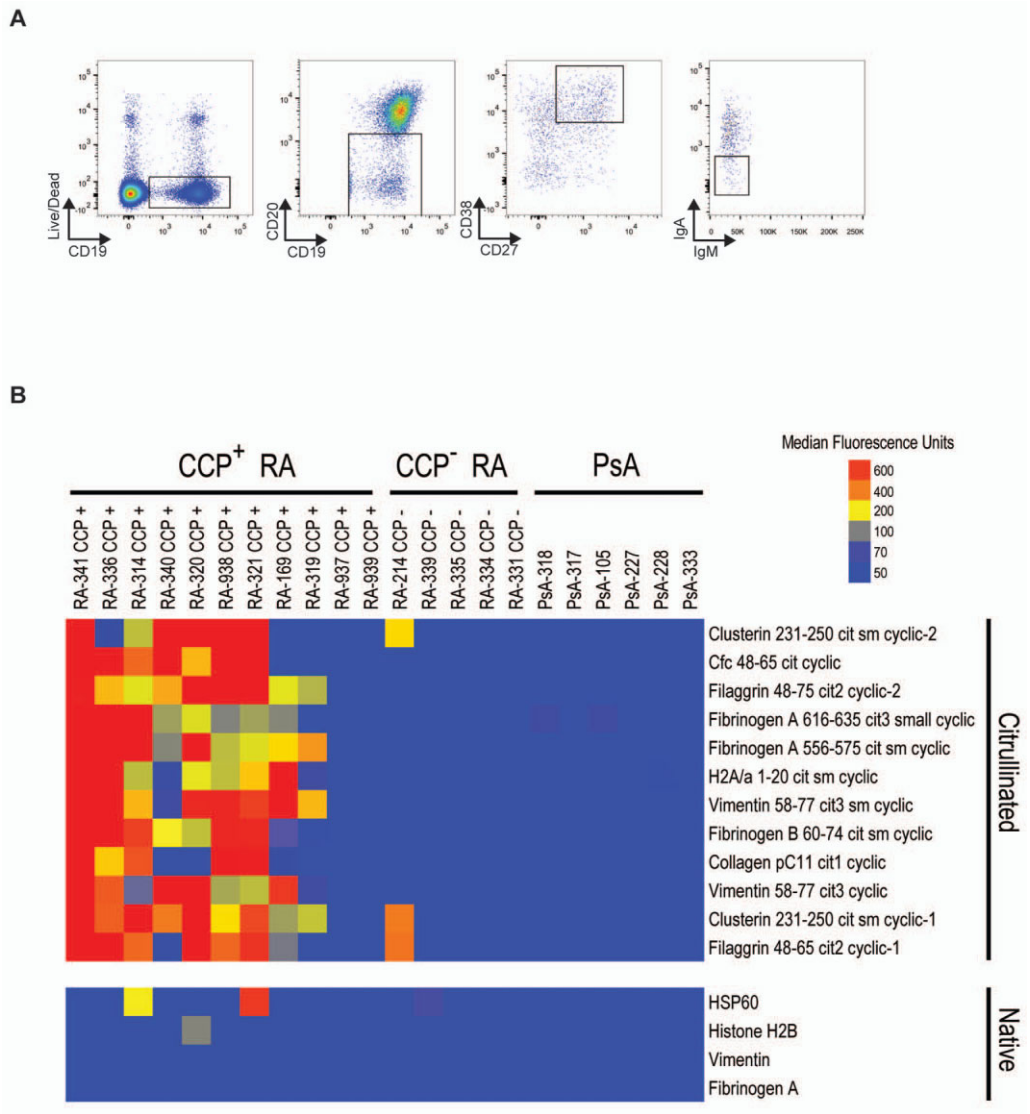

Figure 2. Anti–citrullinated protein antibody production by rheumatoid arthritis (RA) plasmablasts. **A,** Fluorescence-activated gating strategy for single-cell sorting of IgG-secreting plasmablasts from peripheral blood mononuclear cells (PBMCs) obtained from individuals with RA. Plasmablasts were first gated on live, single PBMCs and then gated on surface staining for IgG+ plasmablasts, as shown. **B,** Antigen array screening of the binding to citrullinated peptides of antibodies produced by bulk cultured plasmablasts from patients with anti–cyclic citrullinated peptide (anti-CCP)–positive RA, patients with anti-CCP− RA, and patients with psoriatic arthritis (PsA). Antibody-binding values were normalized as described in Materials and Methods. Blue represents no reactivity, yellow represents moderate reactivity, and orange and red represent high reactivity. Cfc = cyclic citrullinated filaggrin peptide; sm = small; H2A = histone H2A.

families, generating the displayed phylogenetic trees. Therefore, the distance between each V-gene subdendrogram is arbitrary. Trees were drawn with ETE (24).

Cloning of antibody genes. To select antibodies representative of clonal families, we aligned the sequences of antibodies within clonal families and selected the most common sequence for expression. Singleton antibodies were selected at random. We inserted κ and λ light chains into vector pEE12.4 (Lonza), and the γ heavy chain into pEE6.4 (Lonza). Heavy- and light-chain sequences were cloned by PCR or gene synthesized, and the heavy- and light-chain variable regions were inserted into the vector containing the 5′ end of the antibody chains (the leader and V(D)J sequences) along with the appropriate antibody constant regions. For PCR isolation of the variable regions, we used well ID–specific forward primers and constant region–specific reverse primers. For gene

synthesis, heavy- and light-chain variable region sequences were e-mailed to and gene synthesized by Lake Pharma.

Expression of monoclonal antibodies. We performed transient, dual transfections of paired light-chain–containing pEE12.4 and heavy-chain–containing pEE6.4 constructs in 293T cells by using Lipofectamine 2000. We purified the expressed antibodies from the culture supernatants by using Protein A Plus–Agarose Beads (Pierce).

Anti-CCP assay. Anti–CCP-2 ELISA was performed according to the instructions of the manufacturer (Euro-Diagnostica).

RA antigen microarrays. RA antigen microarrays were printed and probed, and data sets analyzed, as previously described (25–27).

RESULTS

Development of a DNA barcoding method for sequencing the cognate heavy and light chains of antibodies expressed by individual B cells. We developed a novel method for tagging all cDNA generated from each individual antibody-expressing cell with a unique DNA barcode before sequencing the cDNA; we then used these unique barcodes to match the heavy- and light-chain antibody sequences that derive from the same cell (15). In this way, we uncovered the sequences of the specific heavy and light chains that make up each antibody, information that we used to generate phylogenetic trees of the antibody response and to clone and express the antibodies encoded by these sequences.

We single-cell sorted plasmablasts into separate wells of 96-well plates. For each plate of 96 wells, we included 8 control wells that remained empty (i.e., we did not sort cells into those 8 wells). Therefore, for each plate, we could expect at most 88 sequences. On average, we achieved ~44 heavy-chain sequences per plate, and 65 light-chain sequences per plate. After quality control and pairing of sequences on the basis of barcodes, we obtained on average 41 antibody sequences (paired heavy and light chains) per plate.

Production of ACPAs by RA peripheral blood plasmablasts. We collected PBMCs from individuals with active RA, and bulk-sorted IgG+ plasmablasts on the basis of cell-surface staining for CD19+CD20−CD27+CD38+IgA−IgM− (Figure 2A). Bulk-sorted IgG+ plasmablasts were cultured in vitro, and their culture supernatants were analyzed on RA antigen microarrays (Figure 2B). (Additional results are available from the author upon request.) Antibodies derived from the supernatants of peripheral blood plasmablasts from individuals with anti-CCP+ RA reacted to multiple citrullinated peptides contained on RA antigen microarrays (25,26), including citrullinated peptides derived from fibrinogen α, vimentin, histone H2A, and clusterin. In contrast, those from individuals with anti-CCP− RA had minimal reactivity, and those from individuals with psoriatic arthritis had no reactivity. This observation, that plasmablasts circulating in the peripheral blood of individuals with RA produce ACPAs (Figure 2B), is consistent with the findings of others (28), and suggests that analyzing plasmablast antibody repertoires in active RA could uncover the specific ACPAs and other key autoantibodies involved in RA pathogenesis.

Sequencing the antibody repertoires of RA peripheral blood plasmablasts. We single-cell sorted IgG+ plasmablasts from PMBCs derived from 4 anti-CCP+ RA patients into 96-well plates on the basis of cell-surface staining for CD19+CD20−CD27+CD38+IgA−IgM− (Figure 2A). Following the system shown in Figure 1, we synthesized and barcoded cDNA from the plasmablasts and sequenced it by 454 sequencing. We used our bioinformatic analyses method to obtain sequences of cognate heavy- and light-chain pairs and determined their V(D)J usage. Using maximum-likelihood clustering, we then performed heavy chain–weighted clustering of the resulting paired heavy- and light-chain sequences to generate phylogenetic trees of the antibody responses in anti-CCP+ RA patients and identify clonal families of antibodies that use the same heavy-chain V(D)J and light-chain VJ sequences as well as singleton antibodies (Figure 3). Nucleotide sequence alignments of antibodies contained in clonal families show evidence of somatic hypermutation due to affinity maturation, and an example of a representative heavy-chain alignment is available from the author upon request.

Using plasmablast cDNA generated from 4 patients, we sequenced a total of 557 antibodies (paired heavy and light chains). We identified 38 clonal families, consisting of a total of 105 antibodies. The remaining 452 antibodies we identified as singletons.

Identification of recombinant RA plasmablast antibodies that bind CCP. We selected antibodies representative of 11 different clonal families from the 4 patients (RA3–RA7 and RA9–RA14), and 3 singleton antibodies that do not belong to a clonal family (RA1, RA2, and RA8) for PCR cloning or gene synthesis and recombinant expression (Figure 3). Using a commercial anti–CCP-2 assay, we found that 4 of the recombinant

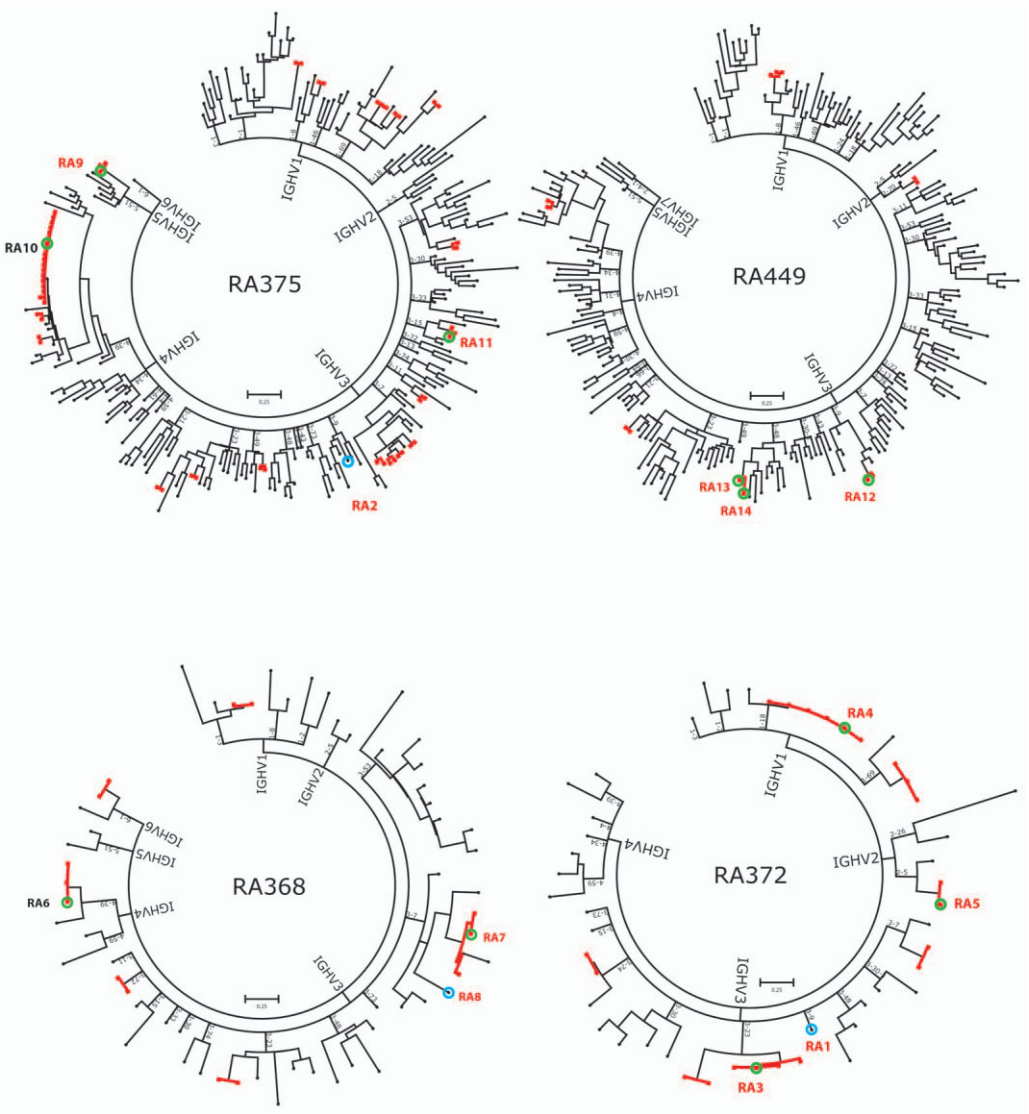

Figure 3. Sequencing of the plasmablast antibody repertoires in individuals with RA. DNA barcoding and next-generation sequencing were used to analyze the plasmablast antibody repertoire of 4 individuals with active, anti-CCP+ RA. A phylogenetic tree of the plasmablast antibody repertoire was generated by concatenating and clustering the heavy- and light-chain sequences of the plasmablast antibodies and arranging them by heavy-chain V-gene family usage. Each peripheral node represents a single antibody. Red lines indicate clonal families of antibodies. The letters RA followed by a number denote an antibody that was selected for cloning and expression, and red font denotes antibodies that exhibited reactivity in the anti–CCP-2 assay, as described in the Figure 4A legend, or on RA antigen microarrays, as described in the Figure 4B legend. See Figure 2 for definitions.

antibodies—RA1, RA2, RA3, and RA8—can bind to CCP (Figure 4A). Recombinant antibodies RA1, RA2, and RA3 bound weakly, whereas RA8 bound relatively strongly. Our observation that several of these antibodies bind only weakly to CCP may be because CCP is an artificial mimic of the unknown endogenous citrullinated antigens present in RA synovium (1,4,26).

Identification of citrullinated fibrinogen, histones H2A and H2B, and α-enolase as targets of recombinant RA plasmablast antibodies. We next used RA antigen microarrays containing putative RA autoantigens (25,27) to screen the antigen specificity of the recombinant antibodies. We found that antibody RA4 bound strongly to the α-enolase peptide 403–442, while

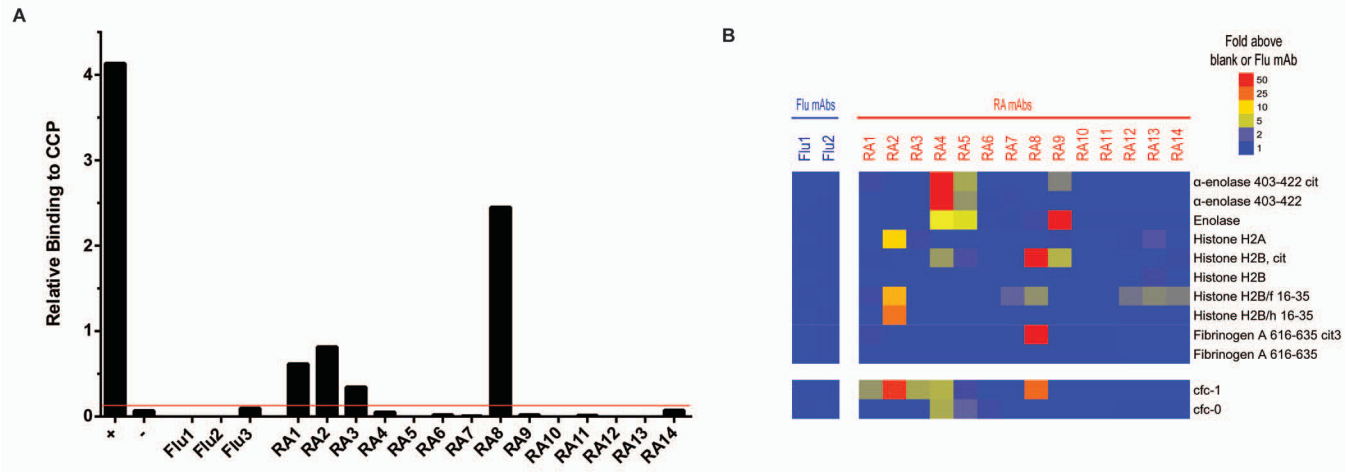

Figure 4. Identification of recombinant antibodies that bind to citrullinated peptides or proteins. **A,** Anti–cyclic citrullinated peptide 2 (anti–CCP-2) enzyme-linked immunosorbent assay (ELISA) analysis of recombinant antibody binding to CCP. Three influenza (Flu)–specific antibodies (Flu1, Flu2, and Flu3) were used as isotype controls. The red line indicates 3 SEM above the reactivity of the 3 negative control influenza-binding antibodies tested. **B,** Antigen microarray analysis of recombinant monoclonal antibody (mAb) binding to candidate citrullinated and native synovial peptides and proteins. Reactivity of the recombinant antibodies to Cfc-1, a citrullinated filaggrin peptide used in early anti-CCP assays, is also displayed.

RA5 and RA9 bound to it weakly (Figure 4B). Alpha-enolase is present in the synovium, and autoantibody targeting of α-enolase is thought to be a pathogenic link between environmental and genetic risk factors for the development of RA (29,30). RA8 bound strongly to citrullinated histone H2B and to the citrullinated fibrinogen α peptide 616–635. RA2 bound to histone H2A (which is constitutively citrullinated as evidenced by mass spectrometric analysis; data not shown) and to the native histone H2B peptides H2B/f 16–35 and H2B/h16–35. RA8, RA12, RA13, and RA14 all bound to the native histone H2B peptide H2B/f 16-35.

Although recombinant antibodies RA1 and RA3 did not bind to any of the putative citrullinated synovial joint antigens contained on our RA antigen microarrays, they did bind to cfc-1 (a citrulline-containing peptide of the synthetic CCP (5) (Figure 4B) (additional results are available from the author upon request), confirming the results of the anti–CCP-2 assay (Figure 4A).

Thus, our method enabled direct generation of monoclonal antibodies sequenced from peripheral blood plasmablasts derived from anti-CCP+ RA patients, and these antibodies were shown to target putative citrullinated and native autoantigens (26,29–33). Our results demonstrate that ACPAs that target citrullinated epitopes present on fibrinogen and histone H2B are produced by the active B cell response (e.g., the plasmablast response) in RA, suggesting that these ACPAs and their corresponding autoantigen targets may contribute to the pathogenesis of RA.

DISCUSSION

We describe the development of a DNA barcoding technology for the large-scale sequencing of the cognate heavy- and light-chain immunoglobulin genes expressed by individual B cells, and apply this approach to sequence the plasmablast antibody repertoires of individuals with active, anti-CCP+ RA. This method accurately and rapidly yields sequence information at a breadth and depth that allows one to dissect the antibody response and map its evolution with high resolution. Bioinformatic identification of groups of antibodies that use the same heavy-chain V(D)J and light-chain VJ sequences allows for rational, bioinformatic selection, and hence efficient production, of the antibodies that are likely to be integral to the active immune response.

Plasmablasts develop during an immune response (from newly activated naive as well as memory B cells), express antibodies that are affinity matured, and circulate transiently in the bloodstream (16–18). Therefore, by analyzing the antibody repertoires of circulating plasmablasts, we can home in on those antibodies that are produced specifically as part of an ongoing immune response—in this case, the relevant autoantibodies in active RA. Studying antibody responses when antigen

exposure is not ongoing would require analysis of other antibody-expressing cells, such as antigen-specific memory B cells or tissue-resident plasma cells, and our method can equally be applied to the analysis of such cell types. Indeed, our method can be used to study the antibody repertoires of any B cell lineage cells, and thus represents a versatile tool that can be applied to diverse fundamental, translation, and clinical questions.

In using our DNA barcoding method to characterize the peripheral blood plasmablast antibody repertoires in RA (28,34), we generated phylogenetic trees representing the antibody repertoires of these individuals. The phylogenetic trees revealed multiple distinct clonal families of antibodies with shared heavy-chain V(D)J and light-chain VJ sequences, indicating ongoing B cell activation and affinity maturation in individuals with anti-CCP+ RA.

From the RA plasmablast antibody repertoires, of the 14 recombinant antibodies we expressed, 4 were reactive in the anti–CCP-2 assay and 1 bound to citrullinated epitopes contained on RA antigen microarrays, i.e., a total of 5 of 14 antibodies had reactivity against citrullinated epitopes. In addition, we identified 3 antibodies that bound to α-enolase, multiple antibodies that bound to histone H2B, and one antibody that bound to histone H2A, i.e., they bound to antigens that are present in synovial joints and whose autoantibody targeting could contribute to the pathogenesis of RA. In addition, representative recombinant antibodies from several clonal families had RF binding activity (data not shown), and characterization of their binding and functional properties are ongoing.

One recombinant antibody bound strongly to a citrullinated fibrinogen peptide and to citrullinated histone H2B. We previously demonstrated that the citrullination of fibrinogen increases the potency with which it binds to Toll-like receptor 4 (TLR-4) by 10-fold, such that immune complexes comprising citrullinated fibrinogen and anti–citrullinated fibrinogen antibodies potently dual-stimulate macrophages through TLR-4 and Fc receptors to produce tumor necrosis factor (35). In addition, transfer of an anti–citrullinated fibrinogen antibody exacerbated arthritis in a mouse model (33). Our finding that peripheral blood plasmablasts derived from anti-CCP+ RA patients produce ACPAs that target citrullinated fibrinogen suggests that these plasmablast ACPAs are the product of an ongoing B cell response in RA. This suggests that citrullinated fibrinogen and anti–citrullinated fibrinogen ACPAs (31–33) could play a role in driving synovitis and the pathogenesis of RA.

In addition, 50% (7 of 14) of the recombinant antibodies bound to epitopes present on histone proteins or histone-derived peptides, specifically to histone H2B peptides or protein, and to histone H2A protein. Histones are one of the constitutively citrullinated proteins (4,36). Although histones are generally intracellular (and often nuclear) proteins, neutrophils can extrude neutrophil extracellular traps (NETs) containing citrullinated antigens, thereby exposing the antigens to the immune system and triggering autoimmune responses in RA and other autoimmune diseases (37,38). NET-mediated exposure of histones may account for the presence of antihistone autoantibodies in the RA patients in our study.

Thus, our method can be used to isolate monoclonal autoantibodies and hence the autoantigens relevant to the pathogenesis of RA and other autoimmune diseases for which the autoantigens remain obscure.

Limitations of the present study include the small number of individuals from whom plasmablasts were obtained, the limited depth of sequencing performed, and the limited comparisons made between singleton antibodies and antibodies belonging to clonal families. Further studies will be necessary to fully assess whether bioinformatically identifying clonal families is useful in identifying the functional, disease-relevant antibodies. It is possible that certain immunodominant antigens elicit the generation of large clonal families of antibodies that bind to a particular antigen but do not have any functional effect. Nevertheless, the ability to use clonal families to "normalize" the antibody repertoires facilitates selection of antibodies representative of both large and small clonal families for recombinant expression and characterization.

In conclusion, we developed a method for DNA barcoding the cDNA generated from an individual cell and used this method to perform large-scale sequencing of the paired heavy- and light-chain immunoglobulin genes expressed by individual peripheral blood plasmablasts derived from humans with anti-CCP+ RA. Bioinformatic analysis of the resulting sequence data sets revealed clonal expansion and affinity maturation of the plasmablast response and enabled us to bioinformatically select antibodies representative of clonal families and singletons for recombinant expression. This approach yielded recombinant antibodies that bound citrullinated epitopes contained in fibrinogen, histone

H2B, and α-enolase, autoantigens implicated in the pathogenesis of RA (1,4,26,29,35).

High-resolution analysis of the clonality and evolution of the antibody response could be useful for identifying, understanding, and monitoring pathogenic immune responses in autoimmune diseases (39). By enabling the rapid, comprehensive, and accurate characterization of ongoing human antibody responses, our DNA barcode–enabled method could provide a critical tool for uncovering the pathogenic autoantibodies produced in RA and other human autoimmune diseases for which the autoantigens remain obscure. Large-scale characterization of the autoantibody repertoires could provide further insights into pathogenesis of, and facilitate development of next-generation diagnostics and therapeutics for, RA and a variety of autoimmune diseases. Our method also has broad applicability for characterizing protective antibody responses against microbial infections, in response to vaccination, and in cancer.

ACKNOWLEDGMENTS

We thank Drs. Tito Serafini, Wayne Volkmuth, and Daniel Emerling for insightful discussions and input.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Robinson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Tan, Sokolove, Robinson.

Acquisition of data. Tan, Kongpachith, Blum, Ju, Lahey, Cai, Wagner.

Analysis and interpretation of data. Tan, Kongpachith, Blum, Lu, Lindstrom, Sokolove, Robinson.

ADDITIONAL DISCLOSURES

Author Tan is currently an employee of Atreca, Inc.

REFERENCES

1. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med 2011;365:2205–19.
2. Firestein GS. Evolving concepts of rheumatoid arthritis. Nature 2003;423:356–61.
3. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101:273–81.
4. Van Venrooij WJ, van Beers JJ, Pruijn GJ. Anti-CCP antibodies: the past, the present and the future. Nat Rev Rheumatol 2011;7:391–8.
5. Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000;43:155–63.
6. Klein F, Halper-Stromberg A, Horwitz JA, Gruell H, Scheid JF, Bournazos S, et al. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature 2012;492:118–22.
7. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, Jarrossay D, et al. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science 2011;333:850–6.
8. Scheid JF, Mouquet H, Feldhahn N, Seaman MS, Velinzon K, Pietzsch J, et al. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature 2009;458:636–40.
9. Scheid JF, Mouquet H, Ueberheide B, Diskin R, Klein F, Oliveira TY, et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science 2011;333:1633–7.
10. Traggiai E, Becker S, Subbarao K, Kolesnikova L, Uematsu Y, Gismondo MR, et al. An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med 2004;10:871–5.
11. Wrammert J, Smith K, Miller J, Langley WA, Kokko K, Larsen C, et al. Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature 2008;453:667–71.
12. Fischer N. Sequencing antibody repertoires: the next generation. MAbs 2011;3:17–20.
13. DeKosky BJ, Ippolito GC, Deschner RP, Lavinder JJ, Wine Y, Rawlings BM, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nat Biotechnol 2013;31:166–9.
14. Cheung WC, Beausoleil SA, Zhang X, Sato S, Schieferl SM, Wieler JS, et al. A proteomics approach for the identification and cloning of monoclonal antibodies from serum. Nat Biotechnol 2012;30:447–52.
15. Tan YC, Blum LK, Kongpachith S, Ju CH, Cai X, Lindstrom TM, et al. High-throughput sequencing of natively paired antibody chains provides evidence for original antigenic sin shaping the antibody response to influenza vaccination. Clin Immunol 2014;151:55–65.
16. Radbruch A, Muehlinghaus G, Luger EO, Inamine A, Smith KG, Dorner T, et al. Competence and competition: the challenge of becoming a long-lived plasma cell. Nat Rev Immunol 2006;6:741–50.
17. Fink K. Origin and function of circulating plasmablasts during acute viral infections. Front Immunol 2012;3:78.
18. Shlomchik MJ, Weisel F. Germinal center selection and the development of memory B and plasma cells. Immunol Rev 2012;247:52–63.
19. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
20. Schmidt WM, Mueller MW. CapSelect: a highly sensitive method for 5′ CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. Nucleic Acids Res 1999;27:e31.
21. Alamyar E, Giudicelli V, Li S, Duroux P, Lefranc MP. IMGT/HighV-QUEST: the IMGT web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. Immunome Res 2012;8:1:2. URL: http://www.imgt.org/PDF/ImmunomeRes/8_1-15_2012.pdf.
22. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32:1792–7.
23. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-

likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59:307–21.
24. Huerta-Cepas J, Dopazo J, Gabaldon T. ETE: a python Environment for Tree Exploration. BMC Bioinformatics 2010;11:24.
25. Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, Fries JF, et al. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. Arthritis Rheum 2005;52:2645–55.
26. Sokolove J, Bromberg R, Deane KD, Lahey LJ, Derber LA, Chandra PE, et al. Autoantibody epitope spreading in the preclinical phase predicts progression to rheumatoid arthritis. PLoS One 2012;7:e35296.
27. Robinson WH, DiGennaro C, Hueber W, Haab BB, Kamachi M, Dean EJ, et al. Autoantigen microarrays for multiplex characterization of autoantibody responses. Nat Med 2002;8:295–301.
28. Kerkman PF, Rombouts Y, van der Voort EI, Trouw LA, Huizinga TW, Toes RE, et al. Circulating plasmablasts/plasmacells as a source of anticitrullinated protein antibodies in patients with rheumatoid arthritis. Ann Rheum Dis 2013;72:1259–63.
29. Mahdi H, Fisher BA, Kallberg H, Plant D, Malmstrom V, Ronnelid J, et al. Specific interaction between genotype, smoking and autoimmunity to citrullinated α-enolase in the etiology of rheumatoid arthritis. Nat Genet 2009;41:1319–24.
30. Lundberg K, Bengtsson C, Kharlamova N, Reed E, Jiang X, Kallberg H, et al. Genetic and environmental determinants for disease risk in subsets of rheumatoid arthritis defined by the anticitrullinated protein/peptide antibody fine specificity profile. Ann Rheum Dis 2013;72:652–8.
31. Masson-Bessiere C, Sebbag M, Girbal-Neuhauser E, Nogueira L, Vincent C, Senshu T, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the α- and β-chains of fibrin. J Immunol 2001;166:4177–84.
32. Zhao X, Okeke NL, Sharpe O, Batliwalla FM, Lee AT, Ho PP, et al. Circulating immune complexes contain citrullinated fibrinogen in rheumatoid arthritis. Arthritis Res Ther 2008;10:R94.
33. Kuhn KA, Kulik L, Tomooka B, Braschler KJ, Arend WP, Robinson WH, et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J Clin Invest 2006;116:961–73.
34. Owczarczyk K, Lal P, Abbas AR, Wolslegel K, Holweg CT, Dummer W, et al. A plasmablast biomarker for nonresponse to antibody therapy to CD20 in rheumatoid arthritis. Sci Transl Med 2011;3:101ra92.
35. Sokolove J, Zhao X, Chandra PE, Robinson WH. Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fcγ receptor. Arthritis Rheum 2011;63:53–62.
36. Rohrbach AS, Slade DJ, Thompson PR, Mowen KA. Activation of PAD4 in NET formation. Front Immunol 2012;3:360.
37. Khandpur R, Carmona-Rivera C, Vivekanandan-Giri A, Gizinski A, Yalavarthi S, Knight JS, et al. NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. Sci Transl Med 2013;5:178ra40.
38. Knight JS, Carmona-Rivera C, Kaplan MJ. Proteins derived from neutrophil extracellular traps may serve as self-antigens and mediate organ damage in autoimmune diseases. Front Immunol 2012;3:380.
39. Boyd SD, Marshall EL, Merker JD, Maniar JM, Zhang LN, Sahaf B, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. Sci Transl Med 2009;1:12ra23.