High resolution characterization of HCV-specific B cell immunity reveals unique antibodies associated with infection outcome

Abstract

Hepatitis C virus (HCV) is a major public health problem, with over 71 million people infected worldwide and at risk for developing life-threatening liver disease. No vaccine for HCV is available, and immunity against the virus is not well understood. In most cases, HCV causes chronic infections, however 20-40% of infected individuals experience spontaneous recovery, suggested to be mediated by HCV-specific natural immunity. Therefore, using HCV as a model for comparing immune responses between spontaneous clearers (SC) and chronically infected (CI) individuals ~~(CI)~~, may empower identification of unique mechanisms dictating viral infection outcomes. Here, by using high-resolution approaches, we provide the first in-depth analysis of HCV-specific B-cell receptor repertoires, which identified unique antibody sequences correlating with infection outcome. We demonstrated that, SC patients develop a unique set of HCV-specific antibody clones with distinct properties~~, as~~ compared to CI patients~~, that developed different HCV-specific antibody clones~~. By using this clonotyping, we developed a machine learning method that accurately predicts infection outcome based on the BCR and TCR repertoires. ~~The~~ HCV-specific clones in SC were characterized by lower mutation~~s~~ number~~s~~ compared to clones from CI. Characterization of novel antibodies constructed by integrating the repertoire analysis with isolation of HCV-specific antibody sequences using combinatorial antibody phage display ~~library~~ technology, revealed that SC unique clones are characterized by high neutralization breadth.

This study demonstrates an innovative high-resolution approach for identifying and constructing unique antibodies correlating with infection outcome. It provides insights into the nature of effective immune responses underlying natural clearance versus progression to a chronic ~~carrier~~ state, ~~that~~which may have clinical implications for prognosis of future status of infection, and the design of immunotherap~~eutic~~ies and a vaccine for HCV.

Introduction

HCV is a major public health problem, with over 71 million people infected worldwide and at risk of developing life-threatening liver disease. HCV infection can lead to hepatitis, cirrhosis, liver failure and hepatocellular carcinoma (HCC) and is the leading cause of liver transplantation [1]. HCC is the fifth-most common cancer, and the third cause of cancer-related death worldwide. Its prevalence in the US and Western Europe is increasing [1]. No vaccine is available for HCV, and immunity against the virus is not well understood. Cure rates are expected to increase with the recent approval of Direct-Acting Antiviral Drugs (DAAs), yet despite this progress, many challenges still remain, such as limited implementation, efficacy and protection from reinfection [2]. Thus, global eradication of HCV by implementation of DAAs is currently not a feasible goal [3-6]. As vaccination is considered the most effective means of eradicating viral infections [5], a prophylactic HCV vaccine is an urgent, unmet medical need [3-6]. However, critical gaps in understanding the correlates of protective HCV immunity have hindered the design of anti-HCV vaccines and immunotherapeutics [3-6].

Unlike HIV-infections that are not spontaneously cleared, 20-40% of HCV-infected individuals experience spontaneous recovery [7]. A multitude of evidence suggests that induction of an efficient HCV-specific natural immunity can control the infection. Therefore, using HCV as a model for comparing immune responses between spontaneous clearers (SC) and chronically infected ~~patients~~ (CI) patients will enable identification of unique mechanisms that dictate human disease outcomes. Until recently, protection against persistent HCV infection was thought to be associated with a vigorous T cell response [8]. However, it is now widely accepted that neutralizing antibodies (nAbs) also play a key role in viral clearance [8-12]. Reports strengthen this point by demonstrating that natural clearance correlates not only with early development of nAbs [13] but also with nAbs exhibiting distinct epitope specificity [14]. Extensive characterization of monoclonal HCV-neutralizing antibodies (mnAbs), combined with the recently solved crystal structure of the core of the HCV envelope protein E2, that is the target of most HCV-nAbs, and of E2-mnAbs complexes, have provided valuable information regarding E2 antigenic landscape [15-19]. However, since most HCV mnAbs characterized to date were generated from CI patients ([12] and reviewed in [20, 21]), the nature and epitope specificities of mnAbs in spontaneously cleared HCV infections remain in question. Recent studies have demonstrated that early appearance of broadly-neutralizing nAbs (bnAbs) are associated with spontaneous clearance [13]. bnAbs also protected against and abrogated HCV infection in animal models [22-24]. Very recently, the first panels of bnAbs isolated from SC infections were developed [25, 26]. The panel reported by Bailey et al., displayed a low number of somatic mutations as compared to the well-characterized nAbs from chronic patients and higher

neutralization breadth, but were similar to nAns from chronic infections in terms of clonality and epitope specificities [26]. Whether different nAbs, with distinct features exist in SC patients, and the nature of the protective immunity they confer, remain to be revealed.

New emerging technologies empowering high-throughput direct screening for specific Abs have provided deep insight into the immunogens that elicit broad Ab responses [27, 28]. In the case of HIV, such technologies, e.g., isolation of HIV-specific B cells from or deep-sequencing of Ab repertoires, led to the generation of bnMAbs with significantly higher potency, breadth and novel epitope specificities (reviewed in [29]). These novel revolutionary methods of studying immune responses can offer important insights about the nature of immune response to infections. The Ab repertoire of an individual stores information about current and past threats that the body has encountered, and thus bears the potential to shed light on screening Abs and vaccine design [27, 30]. Comparing features of Ab repertoires between distinct patient populations may provide information that can be correlated with clinically relevant outcomes [31, 32]. Indeed, recent studies have found common Ab sequences in unrelated individuals following Dengue [33], influenza [34], and HIV [35] infections, as well as autoimmune diseases such as celiac [36] and pemphigus vulgaris [37]. In chronic lymphocytic leukemia, 30% of patients carry highly similar Abs [38]. Dramatic improvements in next generation sequencing (NGS) technologies now enable large-scale characterization of Ab repertoires [27, 28]. Here we utilized these high-resolution technologies to identify unique Abs that stratify between CI and SC HCV infection outcomes and demonstrate that Abs feature distinct properties indicative of mechanisms dictating infection outcome.

Materials and methods

Cell lines

Huh-7.5 cells (a generous gift from Charles Rice, Rockefeller University) and Huh7/FT3-7 cells (a generous gift from Stanley M Lemon, University of North Carolina at Chapel Hill) are human hepatoma cell lines that belongs to Huh-7 subline that is highly permissive for replication of subgenomic and full-length HCV RNAs [39]. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing, high glucose; 10% fetal bovine serum (FBS); 1% L-glutamine; 1% penicillin streptomycin; 1% non-essential amino acid. The cells ~~a~~were incubated in a humidified incubator at 37ºC containing 5% CO2.

The ~~of~~ irradiated 3T3-msCD40L feeder cells that express CD40L were obtained from (?) and cultured as previously described (ref).

Virus

Virus stocks from HJ3–5 chimeric virus (a generous gift from Stanley M Lemon, University of North Carolina at Chapel Hill) [39] and the other chimeric viruses containing E2 envelope protein from genotypes 1-7: HJ3-5/1a, H77C/1a, j6/2b, s52/3a, ED43/4a, sa13/5a, HK6A/6a, QC69/7a (a generous gift from ~~from~~ Prof. Jens Bukh [40]), were produced in Huh7/FT3-7 cells and viral titers were determined by FFU assay in Huh-7.5 cells, as described previously [39].

Antibodies

A panel of HCV mnAbs CBH-4B, CBH-4D, HC-1, HC-11, CBH-7, HC84.22, HC84.26, HC33.1 and HC33.4 that are representative E2 antigenic domain A-E Abs ([12] and reviewed in [20, 21]) were kindly provided from Dr. Steven Foung, Stanford.

Sample collection

מעוצב:גופן: 11 נק', נטוי, גופן עבור עברית ושפות אחרות: 12 נק', לא נטוי

מעוצב:גופן: 12 נק', גופן עבור עברית ושפות אחרות: 12 נק'

מעוצב:גופן: 12 נק', גופן עבור עברית ושפות אחרות: 12 נק'

All blood samples were collected from the Liver institute at Belinson and the Galilee Medical Center, Israel. In total we obtained blood samples from 80 individuals; among these, 18 were individuals that spontaneously cleared HCV infection, 52 with persistent chronic HCV infections and 10 from healthy controls. Subjects were defined as spontaneously cleared HCV if anti-HCV antibodies are detectable, with undetectable HCV RNA assessed by the Taqman reverse-transcription polymerase chain reaction (RT-PCR) quantitative assay. HCV chronic infections were defined as viremia for more than 1 year with detectable viral loads. Both cohorts were not treated with any anti-viral treatment. All blood samples were collected using protocols approved by the Institutional Review Boards and in accordance with the ethical standards of the Helsinki Declaration. Sample data is summarized in Table 1. For the isolation of peripheral blood mononuclear cells (PBMCs), 30-50 ml of whole blood from each donor was separated on Ficoll-Paque gradient (Lymphoprep™) according to the manufacturer's instructions.

Expression and purification of the E2 glycoprotein

מעוצב:צבע גופן: אדום

The H77 genotype 1a E2 sequence spanning residues 384- 661 (not containing the transmembrane domain) was amplifies by PCR pHJ3-5 (ref) using primers pSHOOTER-sec-E2-1a-SE and pSHOOTER-sec-E2-1a-As. The PCR product was digested with NotI and NcoI and cloned into plasmid pCMV-SEC-MBP (a generous gift from ~~Prof.~~ Itai Benhar, Tel-Aviv University) containing signal peptide for secretion, His and Myc tags and fused to maltose-binding protein (MBP) for higher expression and stabilization. The resulting plasmid ~~i~~was termed pCMV-SEC-MBP-E2-384-661-1a-His-Myc.

For production of E2 protein, 293T cells were transfected with 12µg pCMV-SEC-MBP-E2-384-661-1a-His-Myc expression plasmid by PEI transfection reagent. 72h post transfection, medium containing the secreted protein was collected from cells for protein purification. The E2 protein was purified using Ni-NTA agarose beads (Qiagen) according to the manufacturer's instructions. Purified E2 glycoprotein was stored at -20°C. E2 glycoprotein containing fractions were analyzed on SDS-10 % polyacrylamide gels.

Construction of an immune anti-HCV antibody phage display library

We have constructed a phage display antibody library from a source of pooled PBMCs obtained from 10 SC patients. For library construction we designed a degenerative primer set by using the IMGT database (http://www.imgt.org/) (primers are listed in Table 2). Phage Ab library was produced using a protocol ~~as~~ we previously described [41]. In brief:, for each library, total RNA was extracted from 10^7 PBMCs using RNeasy mini kit (Qiagen). cDNA was produced from mRNA by reverse transcription using the AccuScript Hi-Fi cDNA Synthesis Kit (Agilent). Heavy and light chain variable domains were amplified from the RT-PCR cDNA product in PCR using a primer set we have designed. The heavy variable domains were amplified using the primer sets Hu-VH1-6-NcoI-BACK and Hu-JH1-6-FORF and the light variable domain was amplified using primer sets Hu-VK1-6-BACKF and Hu-JK1-5-NotI-FORF or Hu-VL1-10-BACKF and Hu-JL1-7-NotI-FORF. For the combinatorial assembly of the heavy and light chain variable domains into complete single-chain variable fragment (scFv), the fragments were mixed according to their natural frequencies, and PCR was performed using the assembly primer (Fow) and the primers set Hu-JK1-5-NotI-FORF for Kappa scFv or the primers set Hu-JL1-7-NotI-FORF for Lambda scFv (Rev). The amplified scFvs were cloned into the phagemid vector pCC16 (ref). The ligated DNA was purified by ethanol precipitation and then used for electroporation into electrocompetent XL-1 cells (Agilent technologies) with the following conditions: 2.5 kV, 200Ω, 25µF. In Total, we conducted 75 electroporations that yielded a total library size of 6*10^7 individual clones. To ensure vast diversity of libraries we conducted PCR from a colony that were?? digested by BstNI (NEB). Rescue of each library using helper phage and preparation of library stocks was performed essentially as described (ref).

Ensure what? **עם הערות: [PP1]**

מעוצב:כתב עילי

Biopanning and isolation of monoclonal anti-E2 phages

To enrich E2 specific phages, 5 cycles of biopanning was performed for the SC library essentially as described (ref). In brief:, phages were first rescued from library. Then, the first cycle of enrichment

was performed by coating the wells with E2 glycoprotein, and then 10^{11} phages ~~$1*10^{11}$~~ were added to the wells. Nonspecific phages were washed by PBST and ~~and~~ then specific phages were eluted with 100mM triethylamine. For neutralizing, 1M Tris•Cl pH 7.4 was added. Eluted phages were used for the next cycle of biopanning.

מעוצב:מיושר לשני הצדדים

Affinity selected phages were pooled and used for screening the libraries for specific phages. Phages were pooled from the 4^{th} and 5^{th} biopanning cycles. 96 colonies were picked from each cycle and rescued essentially as described (ref). Their specificity to E2 was screened by an ELISA as described below. For each biopanning selected cycle, ELISA was conducted 3 times.

Expression and purification of full length antibodies

To produce full length IgGs the heavy and light chains from scFvs were cloned into pMAZ-IgH and pMAZ-IgL vectors (a generous gift from ~~Prof.~~ Itai Benhar, Tel-Aviv University) that contain the constant regions of IgG (?) and a signal peptide for secretion. Heavy chain was amplified using primers (??) and light chains were amplified using primers (???). PCR products were digested with *BssH*II and *Nhe*I for heavy chains, *BssH*II and *Bsiw*I for the light Kappa chain and *BssH*II and *Avr*II for the light Lambda chain, and cloned into appropriate vectors.

For antibody production, 293T cells were transfected with pMAZ-IgH expressing the Heavy chain and ~~-~~of pMAZ-IgL expressing the Light chain. 72h post transfection, medium was collected from the cells ~~for~~ and antibodies were pu~~fir~~ified using Protein A Sepharose CL-4B beads (GE healthcare) according to the manufacturer's instructions. Purified antibodies were stored at -20°C. Antibodies containing fractions were analyzed on SDS~~/~~ 15 % polyacrylamide gels.

ELISA

For detecting specific antibodies in patient's sera: Each well of the ELISA plate was coated with 0.5µg of rE2 diluted in 100 µl of coating buffer and the plates were incubated at 4^0C overnight. The plates were washed twice with PBST and blocked with 3% skim milk in PBS for one hour at 37^0C. Next, the plates were washed twice with PBST and serum from different patients diluted 1:1000 were added to the wells followed by one-hour incubation at RT. The plates were washed three times with PBST and goat α human HRP conjugated antibody diluted 1:10000 was added for each well followed by one-hour incubation at RT. Then, 100µl of Tetramethylbenzidine (TMB) was added for each well and following incubation of 5-10 min the reaction was stopped by adding 50µl 0.5M of H_2SO_4 to each well. The plates were read in an ELISA plate reader in a wavelength of 450nm.

For detecting specific phages: ELISA was performed as previously described (Benhar, & Reiter, 2002). 96 wells ELISA plate were coated with 5µg of rE2 or negative control protein (BSA). Plates were incubated overnight, then washed ×3 with PBS, and blocking buffer was added to plate for 2h at 37^0C. Next, individual rescued phages were added from the master plate. Plates were incubated at RT 1h and washed ×3 with PBS. 1:5000 HRP conjugated to α M13 antibody was added. Then, 100µl of TMB was added and following incubation of 30 min the reaction was stopped by 50µl 0.5M of H_2SO_4 to each well. The plates were read in an ELISA plate reader in a wavelength of 410nm. Specific phages were picked by the strength of the O.D they exhibited to rE2 as opposed to BSA.

For antibodies specificity. For detecting antibodies binding to the rE2 ELISA plates were coated with 5µg of rE2. The plate was incubated and blocking buffer was added. Then, antibodies were added in different concentrations (?) and incubated for 1h at RT. HRP conjugated GαH was added at 1:10000 dilution and plate was incubated 1h at RT. TMB was added and following incubation of 5-10 min 50µl 0.5M of H_2SO_4 was added to each well. The plates were read in an ELISA plate reader in a wavelength of 450nm.

HCV-Neutralization assays

Huh7.5 cells were seeded on an eight chambers glass slide and incubated overnight at 37^0C. The next day, $5*10^{11}$ of each selected phage or different concentrations of purified IgGs was incubated for 1h with 100FFU/ml of HCVcc HJ3-5 chimeric virus or viruses containing E2 from genotypes 1-7 (1a (H77/JFH1); 2b (J8/JFH1); 3a (S52/JFH1); 4a (ED43/JFH1); 5a (SA13/JFH1); 6a (HK6a/JFH1); 7a (QC69/JFH1)) for 1h. Next, phages and virus mixtures were added to the cells. The mixtures were incubated in 37^0~~0~~C for one hour. The medium containing the virus and serum was added to the wells. The slides were incubated for 24 hours. 200µl of DMEM was added to each well and the slide was incubated for another 24 hours. The slide was washed twice with 200µl PBS. The PBS was gently removed and 100µl of Mathnol:Acetone 1:1 was added to each well followed by 10 minutes incubation ~~in~~at RT. Each well was washed twice with 200µl PBS. 7.5% BSA in PBS was added with serum from chronic infected HCV patient in a dilation of 1:1000 followed by 1 hour incubation in 370C. Each well was washed twice with 200µl PBS. 100µl of 7.5% BSA in PBS with goat anti-human antibody diluted 1:100 was added to each well followed by one hour incubation in RT. Each well was washed 3 times with 200µl PBS. Neutralization was measured by immunofluorescence microscopy and manual counting of foci.

Isolation of HCV-specific B cells

We have established platform for the propagation and isolation of HCV-specific B cells. PBMCs from CI and SC patients were isolated and CD19+ B cells were separated by a FACS sorter. B cells were then plated on feeder cells of irradiated 3T3-msCD40L cells that express CD40L, which induces proliferation, ~~and~~ Ab class switching and secretion [42]. B cells were activated with 5 µg/ml rE1E2 protein and a combination of IL2 (10000 U/ml) and IL21 (100 µg/ml) [43]. The cultures were incubated for 14 days and then HCV-specific B cells were isolated. After 14 days, activated B cells were incubated with rE2 and stained with CD19-PE, CD27-BV421 and tagged rE1E2 (anti-cMyc, alexa uor 633). Viable, CD19+, CD27+ and HCsAg+ were isolated by FACS. These HCV-specific B cells were then grown for one week, as described above. Supernatants were collected in each step and used in HCV-neutralization assays.

Sequencing B-cell repertoires

Bioinformatic analyses

~~28 individuals were enrolled: 7 healthy individuals as the control (C) group, 10 chronically infected with HCV (chronically infected, CI), and 11 that were infected with HCV and later on recovered spontaneously (spontaneously cleared, SC). Blood samples were...~~

Pre-processing of raw sequencing reads: Repertoire Sequencing TOolkit (pRESTO) (Vander Heiden et al., 2014) was applied to the raw reads using the following steps: a. Removal of low-quality reads (mean Phred quality score < 20). b. Removal of reads where the primer could not be identified or had a poor alignment score (mismatch rate > 0.1). c. Identification of sets of sequences with identical molecular IDs (corresponding to the same mRNA molecule). These will be collapsed into one consensus sequence per set, after removing sets with mean mismatch rate > 0.2. d. Assembly of the two consensus paired-end reads into a complete Ig sequence. Then, V(D)J segments were assigned for each of the Ig sequences using IMGT/HighV-QUEST (Alamyar et al., 2012). This was followed by quality control and additional filtering: a. Removal of non-functional sequences due to a stop codon or a reading frame shift between the V and the J gene. b. Sequences with Junction length < 12 nucleotides. c. Sequences whose junction length is not dividable by 3 and therefore cannot be completely translated into an amino acid sequence. Samples with unusually abundant single V-J junction length combination were excluded: samples CI4 and SC12 met this criterion.

Clustering of related B-cell sequences across all samples~~:~~

מעוצב:גופן: מודגש, ללא קו תחתון, גופן עבור עברית ושפות אחרות: מודגש

מעוצב:גופן: מודגש, גופן עבור עברית ושפות אחרות: מודגש

Sequences were first grouped by their V-gene, J-gene and junction length. For each group, the difference in amino acids between each pair of junctions was calculated by Hamming distance. Hierarchical clustering with complete linkage method was applied and sequences were clustered by genetic distance, using a threshold of 0.15, i.e. maximal dissimilarity between any two junction sequences in a cluster never exceeds 15% (ref).

As an additional quality control step, sequence clusters for which > 90% of sequences came from a single sample were removed.

מעוצב:גופן: מודגש, ללא קו תחתון, גופן עבור עברית ושפות אחרות: מודגש

Exploring enrichment of sequence groups from B or T cells in SC or CI

The frequencies per sample were calculated for our main sequence groups --V gene, J Gene, junction length and isotype. Then, for each separate group and several combinations, samples were grouped by clinical group and the difference was evaluated by t-test. Groups with p value < 0.05 were selected.

מעוצב:גופן: מודגש, ללא קו תחתון, גופן עבור עברית ושפות אחרות: מודגש

Naive detection of sequence clusters abundant in either SC or CI clinical groups of B cells

The frequency of each clustered sequence was calculated per sample. Then, samples were grouped by clinical group, and clusters that contained sequences from at least three additional SC/CI samples were defined as enriched in SC/CI, respectively. Selected clustered sequences were later applied the classification model that will be next described.

מעוצב:גופן: מודגש, ללא קו תחתון, גופן עבור עברית ושפות אחרות: מודגש

Naive detection of junction sequences abundant in either SC or CI clinical groups of T cells

The frequency of each junction sequence was calculated per sample. Then, samples were grouped by clinical symptom, and clustered sequences that contained sequences from at least three additional SC / CI samples were defined as enriched by SC/CI, respectively. Selected junction sequences were later applied the classification model that will be next described.

מעוצב:גופן: מודגש, גופן עבור עברית ושפות אחרות: מודגש

Comparing SC and CI clinical groups by AA conservation levels

1. Sequences with read number (CONSCOUNT) lower than two were removed. 2. Sequences were then grouped by V gene, J gene and junction length, to form gene groups. 3. The frequency of each AA at each junction position in each gene group in each sample was calculated. 4. Position frequencies were grouped by clinical groups. 5. Sum of frequency squares was calculated for each clinical group. 6. Gene-groups containing junction positions for which the subtraction of SC and CI

sum of frequencies was greater than 0.5, were defined as distinguishing gene groups.

מעוצב:גופן: מודגש, ללא קו תחתון, גופן עבור עברית ושפות אחרות: מודגש

Prediction models based on patients' repertoire~~:~~

A) Prediction model based on B cell repertoires:

1. Sequences were grouped to clusters as described above.
2. The frequency of each cluster per sample was calculated.
3. A classification model was applied:

a. The data set was randomly divided into 18 (~90%) and 2 samples (~10%) of training and test sets, respectively.

b. Feature selection was performed by Random forests model.

c. Logistic regression was applied to the remaining features, and the model was applied to the test set. Accuracy rate was measured.

d. The process was repeated 100 times, each time two different samples were taken as a test set.

e. Random predictions: to ensure that our results were not achieved randomly, clinical groups of samples were randomly shuffled. Then, steps a-d were applied to the randomly shuffled model.

B) Prediction model based on T cell repertoires:

1. Sequences from all samples were grouped by their unique AA/DNA junction sequence in model T_junc_AA_90/T_junc_90, respectively.
2. The frequency of each clustered sequence in each sample was then calculated.
3. Classification model was applied as described above in A-3.

Results

Screening for antibody binding and neutralization in resolved and chronic infections

Our overall approach is summarized in Figure 1, and included collection of blood samples from CI~~chronic~~ and ~~spontaneously cleared~~SC HCV infections~~,~~ in addition to healthy controls, and a screen to identify samples containing high levels of HCV-neutralizing antibodies. The selected samples were used for sequencing of total B and T cell~~s~~ repertoires~~ sequencing,~~ that were compared with~~to~~ HCV-specific B cell repertoires. This was~~,~~ followed by~~and~~ constructi~~on of~~ng monoclonal antibodies associated with infection clearance, based on repertoire data combined with antibody engineering tools (Figure 1).

First, we have collected PBMCs and sera from 80 individuals.~~,~~ ~~a~~Among these, 18 were individuals that spontaneously cleared HCV infection, 52 with persistent chronic HCV infections, and 10 from healthy controls. We aimed to identify the presence of neutralizing antibodies in sera from ~~chronic~~CI or ~~cleared~~SC HCV infections. First, we screened these sera for antibod~~y mediated~~ies able to bind~~ing~~ to a recombinant HCV envelope glycoprotein E2 (rE2) that we have produced, ~~(rE2)~~ by ELISA. While high levels of anti-rE2 were detected in chronic HCV infections, very low levels or no binding was detected in resolved HCV infections (Fig. 2A). This result is expected, since the ongoing infection in ~~chronic HCV~~CI ~~infection~~patients results in ~~the~~ generation of ~~antibodies~~ large numbers of anti-HCV antibodies from plasma ~~B~~ cells, while in resolved individuals, ~~the~~ anti-HCV antibodies are secreted from ~~the~~a low number of circulating HCV-specific memory ~~B~~ cells. Then, we screened these sera for HCV-neutralization ability by performing an HCVcc neutralization assay. Surprisingly, only approximately a two-fold drop in neutralization efficiency was observed in resolved infections (an average of 45%) compared to chronic infections (an average of 85%) (Fig. 2B).

To validate that we indeed measure HCV-specific immunity, we have collected two ~~chronic infected~~CI samples before and after successful anti-viral therapy~~, following achieving~~ (as determined by sustained virological response ~~(SVR~~) סיון, צריך לבדוק עם נלי מה היה הטיפול וכמה זמן אחרי טיפול קיבלנו (את הדוגמא). On these samples w~~W~~e again performed ELISA to detect binding to rE2, and HCV-neutralization assays. As expected, we observed a significant drop both in binding and neutralizing HCV following treatment (Figure 2C, D).

Collectively, these results suggest that although the anti-HCV antibodies in resolved infections are at low levels, they may be potent neutralizers. The samples that demonstrated high neutralization efficiency were selected for further analysis.

Repertoire sequencing approach

Previous studies suggested defining a successful immune response to HCV by studying SC versus CI [8, 9, 13, 26]. However, a deep insight into these responses is lacking. Here, we sought to use high-resolution technologies that will significantly increase the number of screened samples and the screening depth of each sample. We employed the Illumina platform that ~~is now suitable for sequencing variable Ab genes (V), as it provides longer reads and generates 10-100 fold more reads per run compared to the 454 traditional instrument. This approach~~ provides the added high sensitivity

necessary to identify ~~Ab~~ antibody sequences that are unique to or enriched in SC versus CI infections.

For amplification of the rearranged variable region of antibody ~~H~~heavy chains (VH), we purified total RNA from $5*10^6$ PBMCs from each sample. We performed RT-PCR reactions using 5' RACE that uses a universal primer for the 5' end, since this approach is not restricted to a predetermined primer set. For the 3', we used primers specific to the Ab constant regions- one primer for each Ab isotype [44]. In addition, single mRNA molecules were labeled with unique 17 base pair identifiers [45-48] that allows to avoid amplification and sequencing bias and accurately measures the original number of each sequence. DNA libraries were generated using the NEBNext Ultra DNA Library Prep kit (NEB) and run on Illumina MiSeq. ~~By using this approach w~~We sequenced Ab repertoires from 28 individuals; among these are 10 HCV ~~chronic infection~~ CI (CI10, 11, 13, 15, 16, 20, 21, 22, 26), ~~and~~ 11 ~~spontaneous clearance~~SC (SC1-3, 7-11, 14, 15) that demonstrated the highest neutralization efficiency as described above (Figure 2B), and 7 healthy control samples (C4-9).

Shouldn't this move to the methods? עם הערות: [PP2]

Need to decide if we are using Ab or antibody עם הערות: [PP3]

Pre-processing of raw sequencing reads was implemented to filter out low quality reads, sequences with identical molecular IDs (corresponding to the same mRNA molecule) and ~~Removal~~ removal of non-functional sequences. Samples CI4 and SC12 were removed from the analysis due to unusually abundant single V-J junction length combination. Following these filtering steps we identified between 10^4 to 10^5 unique full length heavy chain sequences for each sample (Figure 3A).

General characterization of B cell repertoires of resolved and chronic HCV infection

We aimed to identify specific and unique features of antibodies common to specific clinical groups representing different outcomes of HCV infections. The variable regions of BCRs are assembled by a complex process involving somatic recombination of a large number of germline encoded V, D and J gene segments, along with junctional diversity that is added at the boundaries where these segments are joined together [14]. ~~A second process that~~ Following antigen recognition, B cells undergo ~~is~~ an affinity maturation process in which their initial receptors (antibodies) are modified into higher affinity antibodies through cycles of somatic hypermutation and affinity- dependent selection [15]. ~~Therefore, t~~To identify features in B cell repertoires that are unique to ~~chronic~~CI or ~~spontaneous clearance~~SC HCV infections, we evaluated the usage frequency~~ies~~ of each V and J gene segment, junction length, as well as the mutation~~s~~ frequencies across the V genes. Sequences were grouped by their V- gene, J- gene and junction length, clustered by genetic distance, and the frequencies within

This is introduction עם הערות: [PP4]

and between the clinical groups were compared. We did not observe significant differences in jJunction lengths and J genes between the clinical groups (Fig 2B, D, respectively). However, ~~We however identified several V gene groups that presented significant frequency differences among the clinical groups (labels colored in red) (Fig 3C). T~~the usage frequencies of IGHV1-18, IGHV-13 and IGHV2-5 were ~~significantly~~ lower in SC compared to CI and C (Fig 3C, labels colored in red). Interestingly, ~~the frequency of~~ IGHV-169, that ~~was previously suggested to~~ encodes antibodies that binds ~~to~~ HCV in both ~~chronic and spontaneously cleared~~CI and SC infections (12, 34, jci insight), ~~appear also to be~~is lower in ~~spontaneously cleared~~SC infections compared to ~~chronic infections~~CI, ~~although this observation is not statistically significant~~. Several other IGHV genes were identified in higher frequency in SC compared to CI (such as IGHV3-23, IGHV3-9, IGHV4-39), suggesting they may function in antibodies able to neutralize and eliminate HCV ~~but the differences were not statistically significant~~.

[PP5] עם הערות: Need to decide if we call them IGHV or just V

~~A~~We performed a similar analysis ~~was performed~~ for β chains of TCRs ~~that we have sequenced~~from the same individuals. As demonstrated in Supplementary Figure 1, we did not observed ~~significant~~ differences in junction length, ~~I~~isotype frequency and J gene usage between SC and CI clinical groups (Supplementary Figure 1B, C, D). We did detect specific V genes with different frequencies between these clinical groups: TRBV4-1 and TRBV6-1 that showed lower frequency in SC compared to CI, and TRBV4-3 and TRBV6-2 that showed lower frequency in SC compared to CI (Supplementary Figure 1E).

Together, these data demonstrate that except for differences in usage of several V genes, no ~~significant~~ common differences in the general B ~~cell~~ and T- cell repertoires are observed between SC and CI clinical groups.

[PP6] עם הערות: מסקנה קצת מבאסת...

Signatures of VH clones in resolved and chronic HCV infections

Clonal grouping involves clustering ~~the~~ sets of BCR sequences ~~into B cell clones, which are defined as a group of cells~~ that are descended from a common ancestor. B cell clone members do not carry identical V(D)J sequences, but differ because of SHM (Chen, Collins, Wang, & Gata, 2010). Since most of the general features that we have analyzed, including J gene~~s~~ usage and junction length, were similar between SC and CI, we sought to explore the possibility that specific clones are uniquely abundant in either SC or CI groups. We grouped the ~~I~~gantibody sequences by ~~different gene combinations:~~either V X J (Figure 3E), ~~or~~ V X J X junction length (Figure 3F), ~~and~~or ~~sequence~~

~~clusters that were grouped by~~ V~~,~~ X J ~~and~~ X junction similarity (Figure 3G). The frequency of each clustered sequence was calculated per sample and grouped by clinical group. Figure 3 summarizes the ~~significant~~ clones ~~for~~with different explored patterns ~~(P value <0.05)~~. Six V X J genes combinations ~~(Figure 4A)~~ and two V X J X junction length combinations ~~(Figure 4B)~~ demonstrated ~~significantly~~ different abundance between SC and CI ~~(0.01<P<0.05 for V X J and 0.001<P<0.01 for V X J X Junction length groups). All were at lower frequency in SC clinical group than the CI clinical group.~~

~~It is likely that subgroups of V and J genes combinations may contain sequences with higher abundance for one of the clinical groups. Interestingly, w~~We also identified 13 combinations of V, J and junction similarity that are ~~significantly~~ different between the clinical groups. Among these, 11 we~~a~~re unique to SC ~~and were not detected in CI~~, and one ~~is~~ unique to CI ~~(Figure 4C)~~. In total, we identified 165 clones that were enriched in SC samples and 172 clones that were enriched in CI samples.

Maybe we should add a concluding sentence here about the possible meaning of these results **[PP7] עם הערות:**

Mutation signatures in VH sequences in resolved and chronic HCV infections

The first BCRs produced in the antibody response are IgM and IgD, that undergo~~es~~ isotype switching to IgG or IgA ~~and somatic hypermutation in~~during the process of affinity maturation ~~followed by selection of clones with the highest affinity to the Ag~~ (De Silva & Klein, 2015). Therefore, to evaluate mutation~~s~~ frequencies ~~among~~between the clinical groups we first subdivided the sequences into IgM, IgD, IgG or IgA isotypes for each clinical group. ~~–~~No significant differences in the frequencies of the different isotypes were observed between the clinical groups (Supplementary Figure ~~5~~2A). Supplementary Figure ~~4~~2B display~~s~~ a violin plot comparing the distribution of somatic mutation frequencies across IgA, IgD, IgG and IgM between the clinical groups. As expected, higher mutation numbers were observed in IgG and IgA isotypes, compared to IgM and IgD isotypes ~~(Figure 5B)~~. No significant differences were observed in mutation~~s~~ numbers within each isotype between the clinical groups (Supplementary Figure ~~5C~~2C).

We also compared mutation numbers for each isotype across V genes between the clinical groups. Fourteen isotype-specific V genes were significantly different between the clinical groups (Supplementary Figure ~~5D~~2D). Of these, four demonstrated higher mutation number~~s~~ i~~s~~n SC compared to CI, including IGHV3-53, IGHV2-70, IGHV1-8 and IGHV3-33. The remaining ~~10~~

I think we need also here a conclusion sentence **עם הערות: [PP8]**

~~isotype-specific~~ V genes demonstrated lower mutation numbers in SC compared to CI~~, including IGHV3-49, IGHV3-11, IGHV1-69, IGHV3-22, IGHV5-10, IGHV4-28, IGHV7-4~~.

Model for prediction of HCV infection outcome based on frequencies of VH clones

In order to understand whether a combination of features, rather than one at a time, would provide better insight on the unique ~~I~~gantibody sequences that participate in the response to HCV, we utilized a machine learning approach, which predicts the clinical group based on a combination of the given features. This approach can be utilized not only as a prediction model, but it may also be used as a tool to identify significant features that did not rise in the single-feature analysis.

For feature selection~~ and prediction model for B cell repertoires~~, we calculated frequency per-sample for each cluster of sequences. To avoid false clusters which may occur due to grouping of several erroneous sequences with correct ones, we removed rare clusters that appeared ~~in~~at low frequencies and/or in less than four samples. Then, two samples were left out as a test set, and we trained the model on the remaining samples.

We applied Random Forests model to extract the best 18 clusters (equals to the size of the training set), followed by utilization of Logistic regression on the selected clusters to generate the prediction model. Finally, we applied the model to the remaining two samples and accuracy was calculated. The process of sampling and training was repeated 100 times, to ensure that the model was not biased towards specific samples.

The final predication results summarized in Figure ~~6~~4, demonstrate 91% accuracy of prediction of SC vs. CI~~, by using the machine-learning algorithm~~. ~~Moreover~~As a control, when we randomly shuffled the clinical groups and trained our model, prediction rates were 49% and 35% for SC and CI groups, respectively (Figure ~~4~~6A), suggesting that we did not achieve our high accuracy predictions due to over fitting or other random bias of any specific sample. Therefore, we identified clones that can accurately stratify between SC and CI samples (termed as "stratifying clones"). Of the 10 best junction sequences (Figure ~~6~~4C), four sequences, IGHV3-15*IGHJ4*8**130, IGHV4-34*IGHJ6*14**103, IGHV3-23*IGHJ4*10**707 and IGHV3-23*IGHJ6*20**367, were also previously found in the naive enrichment comparison (Figure ~~4C~~3G).

Switch the order of the figure or the text. Now 4c comes before 4b. **עם הערות: [PP9]**

Training the model for T cell repertoires was very similar to the one for B cell repertoires, except that the data was categorized by DNA Junction sequences. Average accuracy for junction DNA sequence was ~79% and 85% for SC and CI groups, compared to 50% by using random sequences (Figure ~~6B~~4B)~~.~~, ~~that~~This is less accurate than prediction using B cell repertoires. Of the 10 best

junction sequences (Figure ~~6D~~4D), two sequences, TGTGCCAGCAGCACAGCGGGACAGGGTTTGACTGAAGCTTTCTTT and TGTGCCAGCAGTTTGGGGACCCCCAATGAGCAGTTCTTC, were also found in the naive enrichment comparison.

Isolation of HCV-specific B cells from resolved and chronic HCV infections

Deciphering the humoral immune response to HCV has been challenging due to the immense heterogeneity of B cell immunity. Previous studies have reported ~~that~~ frequencies of circulating, antigen-specific B cells in humans ~~are~~of up to 1% of the overall B cell population [49]. Therefore, the polyclonal nature of the immune response may impose significant background noise that interferes with characterizing the HCV-specific immune response. Thus, we sought to isolate specifically the HCV-specific B- cell population, and characterize their properties.

Since the HCV-specific B- cells may be rare, we first sought to establish a platform for the in vitro propagation of the HCV-specific B cells before their isolation. For propagation of B- cells, irradiated 3T3-msCD40L cells expressing CD40L, which induces proliferation, ~~and Ab~~ class switching and secretion, were used as feeder cells (58). For the activation of B cells, we first isolated CD19+ B cells by FACS, ~~that t~~Then, the ~~B-~~cells were incubated with rE2 protein that we have produced, and a combination of IL2 and IL21, for 14 days in vitro. The combination of CD40L feeder cells and addition of cytokines IL2, IL21 can successfully stimulate switched -memory B cells to produce high concentrations of IgG ~~in~~to the supernatant (Berglund et al., 2013). Supplementary Figure ~~2~~ 3 demonstrates the successful propagation of memory B cells by using this method. Separation of CD19+ B cells from a healthy individual, and evaluation of CFSE staining following 14 days of ~~in vitro~~ culture~~ing~~ in stimulated ~~and~~or non-stimulated conditions, demonstrate~~s~~ CFSE fading only ~~in~~under stimulated conditions~~,~~. ~~which~~This indicates proliferation of the activated culture (Supplementary Figure 3~~2~~A). Moreover, in the activated culture, 23% of the population is memory B cells that are positive for CD27+, compared with very low numbers of CD27+ cells in the non-activated culture (Supplementary Figure 3~~2~~B). For evaluating the ability of B cells to differentiate and produce IgG, we measured by ELISA the concentrations of IgG secreted to the culture medium 3 or 8 days following B cell activation ~~were measured by ELISA~~. As shown in Supplementary Figure 32C, the activation induced IgG secretion, ~~which is enhanced with longer time of stimulation and higher number of CD19+ B cells~~in a time and cell number dependent manner.

[PP10] עם הערות: לא מבינה את המשפט ☹
מה זה הפלוס שבסוף של כל הנוגדנים?
מה עושה אנטי מיק?
ולמה עוד פעם CD19?
זה מצד אחד מפורט מדי, ומצד שני לא ברור. אולי פשוט נכתוב שבודדנו תאי זיכרון ספציפיים לוירוס?

Next, we isolated HCV-specific B- cells following in vitro propagation of B-cells isolated from CIchronic and SCspontaneously cleared HCV infections. After 12 days of stimulation in vitro, the isolated CD19 positive B-cells were incubated with labeled anti-Myc+ Ab, for detection of rE2 binding to the HCV-specific BCR, and with labeled anti-CD19+ and anti-CD27+ Abs for separation of memory B-cells. The HCV E2+ positive gate was set relative to the CD19+ cells. Background was compared to healthy individuals stained and gated as the tested samples (Supplementary Figure 7A4). In total, the HCV E2+-specific populations were separated from six chronicCI and three SCspontaneously cleared HCV infections (Figure 7B5A). The fold enrichment of cells isolated from HCV-specific B cells ranged between 2 to 466 (Figure 7B5A). To validate the enrichment of HCV-specific B-cells, the isolated HCV-specific B cells were grown in vitro with stimulation of IL2, IL21 and feeder cells 3T3-CD40L. After four days the supernatant containing antibodies were collected and were used for HCV-neutralization assay, demonstrating higher neutralization in CI and SC samples compared with healthy controls, which were further enhanced following separation of HCV-specific B-cells (Figure 7C5B).

Collectively, these results demonstrate that the novel platform that we have developed is successful infor in vitro propagation and separation of HCV-specific B cells that can be detected in and isolated from individuals with chronic and spontaneously cleared HCV infections.

Identifying unique features in HCV-specific B- cell repertoires

To identify unique features in HCV-specific B-cell immune response, Tthe HCV-specific B cells that we have isolated were sequenced as described above. of the general repertoire to identify unique features in HCV-specific B-cell immune response. First, we evaluated the genomic distance of the VDJ region between the binding sequences from between the different samples by Levenshtein distance. Interestingly, some of the most closely related sequences were originated from different samples (Figure 8A5C). This observation may implyies that similar antibodies convergently evolve in different patients to bind HCV. To compare the specific binding sequences repertoire with the general repertoire, we first detectedsearched for sequences clones in the general repertoire that contain Ig sequences are that are similar to the specific binding sequences. Similarity was defined as having same V gene, J gene, and junction sequence with similarity not lower than 75%. In total, we detected 5447 different clones that were similar in the HCV-specific B-cell repertoire that were similar to clones in and the general repertoire. Interestingly, in the specific repertoire we identified 17 clones that were enriched in SC samples in the general repertoire, and 15 clones that were enriched in CI samples in the general repertoire. The lists of these clones are presented in Supplementary Tables 3 and 4. Comparison between these two lists reveals that except for the V and -J gene combinations IGHV3-33*IGHJ4, thatwhich is abundant in both lists, different HCV-binding clones are enriched betweenin the two clinical groups.

Another feature that we have analysed in the general repertoire compared to the specific repertoire is the mutations numbers. Against each specific sequence, one non-specific sequence was randomly sampled from the general repertoire. The sampled sequence contained the same V and J gene as the corresponding specific sequence. Then, sequences were grouped by iIsotype, and mutations numbers in the V gene waswere compared. Both for IgA and IgG ~~sequences, that are the Ig isotypes that undergoes somatic hypermutation process and affinity maturation,~~ we detected significantly higher mutations numbers in specific compared to non-specific repertoires. For IgM, we observed an opposite trend (Mann Whitney test, IGA p=3.488873e-07, IGG p=6.849511e-08, IGM p=3.764229e-04) (Figure ~~8B~~5D). This observation validates that the HCV-specific B-cell repertoire we sequenced is indeed antigen specific, HCV in this case. We then evaluated the mutation number in the HCV-specific repertoire in SC compared to CI. All specific sequences of SC samples were unified into one bulk ~~whil~~, ande CI samples were unified in a second bulk. Then, the sequences were grouped by Iisotype and mutations numbers in the V genes w~~as~~ere compared. Strikingly, the number of mutations in the SC specific repertoire bulk was higher compared to the number of mutations in the CI specific repertoire, with a significant difference for IgA (IGA p=0.000574, IGG p=0.435930) (Figure ~~8C~~5E).

The CDR3 of VH is the most diverse region in antibody sequences. Therefore, conservation of amino acids in theis region ~~sequence of BCRs particularly in the VH CDR3~~ can highlight the evolution of somatic hypermutation in affinity maturation process to maintain specific amino acid positions that are probably specialized in antigen binding. Here we searched for conserved amino acid in CDR3 region in HCV-specific repertoire compared to general repertoire. Against each VXJXCDR3-length binder sequence, we selected random sequence with similar V, J and CDR3 length from our Repertoires - defined as non-binder. Then, Junctions AAs that were very conserved for binders sequences but not very conserved for non-binders were selected. We identified three VXJXCDR3-length combinations containing conserved amino acids is specific position in CDR3 in binder compared to non-binders sequences (Figure ~~8D~~5F). Interestingly, IGHV4-39–IGHJ6–7 contained a stretch of nine conserved positions in CDR3 and was observed in 3 (CI56H, CI57H and CI59H ?עוז האם נכון) different samples. These results imply that these clones are specific for HCV binding and were evolved independently in different subjects.

Identifying binder Ab sequences associated with HCV infection clearance

[PP11] עם הערות: מה זה בעצם אומר? שצריך יחסית יותר מוטציות בשביל HCV מאשר בשביל פתוגנים אחרים?

[PP12] עם הערות: למה?

[PP13] עם הערות: צריך כאן משפט מסקנה או הסבר מה זה אומר

[PP14] עם הערות: אנחנו מצפים לתגובה של IgA לוירוס הזה?

Our ultimate goal was to translate the information obtained via repertoire analysis to construct HCV-neutralization antibodies that will enable to 1) validate that the unique features we observed in SC or CI infections indeed represented differences in the HCV-specific immune response between these clinical groups, and 2) explore the properties of ~~obtain unique HCV-neutralizing~~ antibodies that are associated with infection clearance ~~and explore their properties~~. One limitation of the construction of mAbs directly from repertoire analysis is pairing of heavy and ~~VH with its V~~ light chains ~~(VL)~~. While ~~VH~~heavy and ~~VL~~ light chains have been successfully paired based on their frequencies in the B cell repertoire [30], ~~still,~~ this approach requires extensive testing of many different combinations. We appl~~y~~ied a different approach for matching ~~VH~~heavy with ~~VL~~ light chains, by construction of a phage display Ab library. The advantage of such libraries is that they allow for the selection of Ab binders that are coupled to their Ab sequence. Since these Abs contain the variable regions of both heavy and light chains as single chain (scFv), it enables direct engineering of full Abs [50].

השתנה קוד שדה

השתנה קוד שדה

Since we specifically focused on neutralizing antibodies associated with HCV ~~infection~~ clearance, we have constructed a phage display antibody library from a source of pooled PBMCs obtained from 10 SC patients (Sample data is summarized in Table 1). The scFv library was constructed by amplification of VH and VL genes separately, and then their combinatorial assembly and cloning into a phagemid vector. In total we obtained a library size of $6*10^{\Delta 7}$ individual clones. Screening for HCV E2 binders was performed by 5 cycles of biopanning using rE2. Following the 4^{th} and 5^{th} cycles, 96 colonies were picked from each cycle and their specificity to E2 was screened by ~~an~~ ELISA. Positive phages were validated.

מעוצב:כתב עילי

[PP15] עם הערות: Need to be consistent with numbers, are we writing them in words or numbers? Usually lower than 10 is in words and higher is in numbers.

We have identified and validated six different phages that demonstrated at least 2 fold and up to 15 fold specific binding to rE2 compared to binding to BSA as background (Figure ~~9A~~6A). Comparing the VH sequences of these isolated scFv with sequences from the general repertoire, we identified a cluster of sequences that were similar. We calculated the distances from the group of similar sequences and the VH from the scFv to identify the most similar sequences in the repertoire. scFv SC11 demonstrated the highest similarity to sequences from the general repertoire~~, and we could detected almost identical sequences~~. Similar sequences were identified also for the other scFvs (Figure ~~9B~~6B). Interestingly, these scFvs (except for scFv SC93) that were isolated from a library constructed from a source of SC patients, were indeed more similar to sequences in the SC general repertoire than the CI general repertoire (Supplementary Figure ~~3~~5). Two of the scFv wea~~r~~e grouped to clones that were detected as uniquely abundant in SC repertoires (Figure ~~4C~~3G). The scFv SC11 ~~i~~was clustered to clone IGHV4-39*IGHJ4, that was detected in the repertoires of 7 out of 9 SC samples, and scFv SC28 wa~~i~~s clustered ~~co~~ to clone IGHV6-1*IGHJ6*17**20, that was detected in

[PP16] עם הערות: הייתי שמה פה איזה מספר שמעיד על רמת הדמיון

repertoires of 5 out of 9 SC samples. Both clones were not detected is CI repertoires. Interestingly, clone IGHV6-1*IGHJ6*17**20 was also enriched in the HCV-specific repertoire (Supplementary Table 4). Lineage trees ~~generated for each clone~~ that demonstrate the ancestral relationship between the sequences ~~and the clone~~ and temporal ordering of mutations, reveal~~s~~ed that the VH sequences of SC11 and SC28 are positioned relatively high in the tree (Figure ~~9C~~6C, D). This suggests that these sequences are common ancestors to many sequences in these specific clones.

These results point for scFvs SC11 and SC28 as potential antibody sequences that are abundant and unique to the SC clinical group, and therefore were selected as candidates for construction of full length Abs and characterization of their properties.

Construction of broadly neutralizing antibodies associated with HCV infection clearance

We sought to construct full length Abs from the sequences identified as binders by the phage display library and compare these to Abs we construct directly from sequences identified by the repertoire analysis. Therefore, we have cloned the heavy and light chains from scFvs SC11 and SC28 into pMAZ-IgH and pMAZ-IgL vectors that contain the constant regions of IgG1 for heavy and light chains respectively. In addition, we have cloned the VH sequences from the general repertoires that are the nearest to scFv SC11 and scFv SC28 into pMAZ-IgH vector. Eventually, we obtained two pairs of Abs. One pair is the Ab SC11-generated from the VH and VL of the scFv SC11, and its corresponding Ab VH16618 containing the same VL but its VH was synthesized from the sequence identified in the repertoire. The second pair the Ab SC28-generated from the VH and VL of the scFv SC11, and its corresponding Ab VH510520 containing the same VL but its VH was synthesized from the sequence identified in the repertoire. These four antibodies were expressed in 293T cells and purified from the supernatant by using Protein A Sepharose.

First, we evaluated the specificity of binding of these four antibodies to rE2, compared to binding to BSA in ELISA. We observed more than 35 fold higher binding signals in Abs VH16618 and VH510520 than Abs SC11 and SC28 (Figure ~~10A~~7A). To further characterize the binding capacity of these Abs VH16618 and VH510520 to HCV rE2, we compared binding of these antibodies to binding signals of a well characterized panel of mnAbs, including CBH-4B, CBH-4D, HC-1, HC-11, CBH-7, HC84.22, HC84.26, HC33.1 and HC33.4, that are representative E2 antigenic domain A-E Abs ([12] and reviewed in [20, 21]). ELISA results with rE2 demonstrated comparable binding capacity of Abs VH16618 and VH510520 compared to the well-defined panel of Abs (Figure

~~10B~~7B). To evaluate neutralization breath, we performed neutralization assays with these Abs across all HCV genotypes using a panel of infectious HCVcc containing envelope proteins from genotypes 1-7 [40]. Abs VH16618 and VH510520 efficiently neutralized all 7 HCV genotypes, including genotype 3 which was not efficiently neutralized by previous panels of HCV nAbs, including the recently reported panel from SC patients [26], implying for their high neutralization breadth (Figure ~~10C~~7C, D).

Discussion

This study provides the first in-depth analysis of HCV-specific immune response and identifies unique features that correlate with infection outcome. Except for differences in frequency of usage of few V BCR and TCR gene segments, in general, the landscapes of B and T cells repertoires featuring frequency of usage of particular J gene segments, Junction length and mutations number did not significantly differ between SC and CI clinical groups. The most prominent differences between the SC and CI are specific clones enriched in a specific group identified in general and HCV specific B-cell repertoires. Strikingly, we found that enrichment of specific clones in SC or CI predicts outcome of infection, with accuracy of over 90% for B cell repertoires and 80% for T cell repertoires. The comparison of general B cell repertoires with HCV-specific B-cell repertoires revealed lower mutation number in SC specific repertoire compared to CI. Finally, the construction of two Abs identified in the enrichment analysis, by combining phage display antibody library technology, validated that the clones we identified as unique to SC are indeed HCV-neutralizing antibodies and revealed their exceptional potential of broad neutralization breath.

Our finding that specific clones are distinctive to clearance of HCV infection, while others are distinctive to progression to chronic infection demonstrate that similar antibodies evolved in different individuals. The identification of fraction of these clones in the HCV-specific repertoire validate that they are provoked in response to the infection and binds the virus. Previous publications suggested that VH1-69 antibody gene segment is enriched in clones identified in both SC and CI, based on isolation of a panel of HCV-neutralizing antibodies (ref). However, our high-resolution approach that provides a wide overview on the general repertoire and HCV-specific repertoire demonstrate that this gene segment is more abundant in CI repertoires that SC, and is not enriched in HCV-specific clones. The stratifying clones were grouped by junction similarities that contain the CDR3 which is the most diverse region in the VH sequence (ref). Sharing similar CDR3 in different individuals was suggested to be very rare (ref), although such immunological signatures were recently reported in

viral specific responses (ref). These discoveries raise an intriguing question as what governs these pronounced similarities in the antibody response to HCV in different individuals, which is indicative of infection outcome. Similar genetic background was previously reported not to lead to induction of identical antibody sequences in response to the same antigen, even in identical twins (ref). Indeed, the lack of general differences in features and usage of antibody V and J gene segments implies that that these clinical groups share specific sequences not due to similarities in their BCR or TCR gene alleles. With other possibilities such as the specificity of the antigens or the individual's natural history of antigen exposure, this remains an open question. Nonetheless, our identification of an immune signature of stratifying clones that are able to predict infection outcome with over 90% accuracy may have important clinical relevance. It may have prognostic value for the outcome of an active infection. In the DAAs era, when availability to effective HCV therapy is limited by the high costs (ref), using the platform we have established may be indicative for clinical decisions for treatment.

An important finding of this study is the lower frequency of mutations we observed in HCV-specific BCR clones in SC compared to CI. Indeed, recent study demonstrated that a panel of HCV-neutralizing antibodies isolated from spontaneously cleared infections contained lower number of mutations compared to HCV-neutralizing antibodies isolated from chronic infection (ref). Our study confirms this finding and expands its relevance to a broad HCV-specific antibody sequences and to multiple individuals. Moreover, we validated the broad neutralization potential of two of the identified HCV-specific clones unique to clearance. The broadly neutralizing antibodies were suggested to be induced in early stages of infection in spontaneously cleared infection, while chronic infections were associated with induction of such antibodies later in infection and require higher mutations number to achieve broad neutralization to the variable quasispecies population of viruses that evolved in these later stages. Thereby, the broad neutralizing antibodies with relatively low number of mutations were suggested to be associated with viral clearance. In contrast, in the case of HIV infection, broadly neutralizing antibodies require high mutation number and therefore evolve only years after infection. Indeed, the ability to provoke such immune response of broad neutralization with low number of mutations in HCV infection and not in HIV is translated to proximately 30% spontaneous clearance in HCV infection and no such outcome in HIV infections.

Our findings that HCV-specific Abs in SC are characterized by low mutations number and high neutralization breath compared to Abs in CI combined with the context of these sequences within different clonotypes points to the conclusion that the immune response to HCV infection provoked in SC are essentially different than that provoked in CI. Therefore, we provide the first evidence that the nature of the immune response is associated with infection outcome and not only timing of

appearance of broadly neutralizing Abs as was suggested previously (ref). It will be highly interesting to determine whether Abs that are unique to SC are also characterized with distinct epitopes. Similar epitope specificities were demonstrated for the recent panel of nAbs isolated from SC infections [26]. Still, it has been suggested that nAbs with distinct epitope specificities do exist but remain to be discovered [12]. Discovery of novel antigenic epitopes will point to a new mechanism driving infection outcome.

In summary, this study provides a novel high resolution insight into the nature of HCV-specific immune response and demonstrates for the first time that the outcome of infection is determined by unique features of the immune response induced following infection. Our innovative approach combined B-cell repertoire analysis and Ab engineering tools that provided high sensitivity necessary to identify Ab sequences are unique to or enriched in SC versus CI infections and use this information to produce full antibodies with unique features. This data may have translational information for designing rational prophylactic vaccine and for developing therapeutic vaccine. Passive immunization with combinations of mnAbs possessing well-defined epitope specificities, may overcome virus resistance [51], confer a prophylactic effect (such as in liver transplantation [52], where re-infection of the transplant is rapid [53]) and may also prove effective in treating existing HCV infection [24]. In addition, our identification of Ab signatures that stratify between CI and SC infections may have prognostic value for predicting the outcome of infection. In a more general point of view, the in-depth analysis of immune repertoires as demonstrated here may open a realm of possibilities for advancing monoclonal Ab discovery and engineering strategies which bear many potential clinical implications. It may advance the capacity to scan gene repertoires of individuals to identify past infections and immunological weaknesses, and to tailor personalized vaccines.

References

.1 Parkin DM: **The global health burden of infection-associated cancers in the year 2002.** *Int J Cancer* 2006, **118:**3030-3044.

.2 Hayes CN, Chayama K: **Why highly effective drugs are not enough: the need for an affordable solution to eliminating HCV.** *Expert Rev Clin Pharmacol* 2017, **10:**583-594.

.3 Trucchi C, Orsi A, Alicino C, Sticchi L, Icardi G, Ansaldi F: **State of the Art, Unresolved Issues, and Future Research Directions in the Fight against Hepatitis C Virus: Perspectives for Screening, Diagnostics of Resistances, and Immunization.** *J Immunol Res* 2016, **2016:**1412840.

.4 Hagan LM, Schinazi RF: **Best strategies for global HCV eradication.** *Liver Int* 2013, **33 Suppl 1:**68-79.

.5 Walker CM, Grakoui A: **Hepatitis C virus: why do we need a vaccine to prevent a curable persistent infection?** *Curr Opin Immunol* 2015, **35:**137-143.

.6 Honegger JR, Zhou Y, Walker CM: **Will there be a vaccine to prevent HCV infection?** *Semin Liver Dis* 2014, **34:**79-88.

.7 Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, Kidd J, Kidd K, Khakoo SI, Alexander G, et al: **Genetic variation in IL28B and spontaneous clearance of hepatitis C virus.** *Nature* 2009, **461:**798-801.

.8 Cashman SB, Marsden BD, Dustin LB: **The Humoral Immune Response to HCV: Understanding is Key to Vaccine Development.** *Front Immunol* 2014, **5:**550.

.9 Dustin LB, Cashman SB, Laidlaw SM: **Immune control and failure in HCV infection--tipping the balance.** *J Leukoc Biol* 2014, **96:**535-548.

.10 Houghton M: **Hepatitis C: the next 25years.** *Antiviral Res* 2014, **110:**77-78.

.11 Walker CM: **Designing an HCV vaccine: a unique convergence of prevention and therapy?** *Curr Opin Virol* 2017, **23:**113-119.

.12 Pierce BG, Keck ZY, Lau P, Fauvelle C, Gowthaman R, Baumert TF, Fuerst TR, Mariuzza RA, Foung SK: **Global mapping of antibody recognition of the hepatitis C virus E2 glycoprotein: Implications for vaccine design.** *Proc Natl Acad Sci U S A* 2016.

.13 Osburn WO, Snider AE, Wells BL, Latanich R, Bailey JR, Thomas DL, Cox AL, Ray SC: **Clearance of Hepatitis C Infection Is Associated With the Early Appearance of Broad Neutralizing Antibody Responses.** *Hepatology* 2014, **59:**2140-2151.

.14 Ndongo N, Berthillon P, Pradat P, Vieux C, Bordes I, Berby F, Maynard M, Zoulim F, Trepo C, Petit MA: **Association of anti-E1E2 antibodies with spontaneous recovery or sustained viral response to therapy in patients infected with hepatitis C virus.** *Hepatology* 2010, **52:**1531-1542.

.15 Kong L, Giang E, Nieusma T, Kadam RU, Cogburn KE, Hua Y, Dai X, Stanfield RL, Burton DR, Ward AB, et al: **Hepatitis C virus E2 envelope glycoprotein core structure.** *Science* 2013, **342:**1090-1094.

.16 Sabahi A, Uprichard SL, Wimley WC, Dash S, Garry RF: **Unexpected structural features of the hepatitis C virus envelope protein 2 ectodomain.** *J Virol* 2014, **88:**10280-10288.

.17 Khan AG, Whidby J, Miller MT, Scarborough H, Zatorski AV, Cygan A, Price AA, Yost SA, Bohannon CD, Jacob J, et al: **Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2.** *Nature* 2014, **509:**381-384.

.18 Meola A, Tarr AW, England P, Meredith LW, McClure CP, Foung SK ,McKeating JA, Ball JK, Rey FA, Krey T: **Structural flexibility of a conserved antigenic region in hepatitis C virus glycoprotein E2 recognized by broadly neutralizing antibodies.** *J Virol* 2015, **89:**2170-2181.

.19 Freedman H, Logan MR, Law JL, Houghton M: **Structure and Function of the Hepatitis C Virus Envelope Glycoproteins E1 and E2: Antiviral and Vaccine Targets.** *ACS Infect Dis* 2016, **2:**749-762.

.20 Kong L, Jackson KN, Wilson IA, Law M: **Capitalizing on knowledge of hepatitis C virus neutralizing epitopes for rational vaccine design.** *Curr Opin Virol* 2015, **11:**148-157.

.21 Ball JK, Tarr AW, McKeating JA: **The past, present and future of neutralizing antibodies for hepatitis C virus.** *Antiviral Res* 2014, **105:**100-111.

.22 Law M, Maruyama T, Lewis J, Giang E, Tarr AW, Stamataki Z, Gastaminza P, Chisari FV, Jones IM, Fox RI, et al: **Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge.** *Nat Med* 2008, **14:**25-27.

.23 Morin TJ, Broering TJ, Leav BA, Blair BM, Rowley KJ, Boucher EN, Wang Y, Cheslock PS, Knauber M, Olsen DB, et al: **Human monoclonal antibody HCV1 effectively prevents and treats HCV infection in chimpanzees.** *PLoS Pathog* 2012, **8:**e1002895.
.24 de Jong YP, Dorner M, Mommersteeg MC, Xiao JW, Balazs AB, Robbins JB, Winer BY, Gerges S, Vega K, Labitt RN, et al: **Broadly neutralizing antibodies abrogate established hepatitis C virus infection.** *Sci Transl Med* 2014, **6:**254ra129.
.25 Merat SJ, Molenkamp R, Wagner K, Koekkoek SM, van de Berg D, Yasuda E, Bohne M, Claassen YB, Grady BP ,Prins M, et al: **Hepatitis C virus Broadly Neutralizing Monoclonal Antibodies Isolated 25 Years after Spontaneous Clearance.** *PLoS One* 2016, **11:**e0165047.
.26 Bailey JR, Flyak AI, Cohen VJ, Li H, Wasilewski LN, Snider AE, Wang S, Learn GH, Kose N, Loerinc L ,et al: **Broadly neutralizing antibodies with few somatic mutations and hepatitis C virus clearance.** *JCI Insight* 2017, **2**.
.27 Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR: **The promise and challenge of high-throughput sequencing of the antibody repertoire.** *Nat Biotechnol* 2014, **32:**158-168.
.28 Benichou J, Ben-Hamo R, Louzoun Y, Efroni S: **Rep-Seq: uncovering the immunological repertoire through next-generation sequencing.** *Immunology* 2012, **135:**183-191.
.29 Burton DR, Poignard P, Stanfield RL, Wilson IA: **Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses.** *Science* 2012, **337:**183-186.
.30 Reddy ST, Ge X, Miklos AE, Hughes RA, Kang SH, Hoi KH, Chrysostomou C, Hunicke-Smith SP, Iverson BL, Tucker PW, et al: **Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells.** *Nat Biotechnol*, **28:**965-969.
.31 Gibson KL, Wu YC, Barnett Y, Duggan O, Vaughan R, Kondeatis E, Nilsson BO, Wikby A, Kipling D, Dunn-Walters DK: **B-cell diversity decreases in old age and is correlated with poor health status.** *Aging Cell* 2009, **8:**18-25.
.32 Watson CT, Breden F: **The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease.** *Genes Immun* 2012, **13:**363-373.
.33 Parameswaran P, Liu Y, Roskin KM, Jackson KK, Dixit VP, Lee JY, Artiles KL, Zompi S, Vargas MJ, Simen BB, et al: **Convergent antibody signatures in human dengue.** *Cell Host Microbe* 2013, **13:**691-700.
.34 Pappas L, Foglierini M ,Piccoli L, Kallewaard NL, Turrini F, Silacci C, Fernandez-Rodriguez B, Agatic G, Giacchetto-Sasselli I, Pellicciotta G, et al: **Rapid development of broadly influenza neutralizing antibodies through redundant mutations.** *Nature* 2014, **516:**418-422.
.35 Sok D ,Laserson U, Laserson J, Liu Y, Vigneault F, Julien JP, Briney B, Ramos A, Saye KF, Le K, et al: **The effects of somatic hypermutation on neutralization and binding in the PGT121 family of broadly neutralizing HIV antibodies.** *PLoS Pathog* 2013, **9:**e1003754.
.36 Di Niro R, Mesin L, Zheng NY, Stamnaes J, Morrissey M, Lee JH, Huang M, Iversen R, du Pre MF, Qiao SW, et al: **High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions.** *Nat Med* 2012, **18:**441-445.
.37 Steinsbo O, Henry Dunand CJ, Huang M, Mesin L, Salgado-Ferrer M, Lundin KE, Jahnsen J, Wilson PC, Sollid LM: **Restricted VH/VL usage and limited mutations in gluten-specific IgA of coeliac disease lesion plasma cells.** *Nat Commun* 2014, **5:**4041.
.38 Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, Davis Z, van Gastel-Mol EJ, Tresoldi C, Chu CC, et al: **Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies.** *Blood* 2012, **119:**4467-4475.
.39 Yi M, Ma Y, Yates J, Lemon SM: **Compensatory mutations in E1, p7, NS2, and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus.** *J Virol* 2007, **81:**629-638.
.40 Gottwein JM, Scheel TK, Jensen TB, Lademann JB, Prentoe JC, Knudsen ML, Hoegh AM, Bukh J: **Development and characterization of hepatitis C virus genotype 1-7 cell culture systems: role of CD81 and scavenger receptor class B type I and effect of antiviral drugs.** *Hepatology* 2009, **49:**364-377.

.41 Nahary L, Benhar I: **Design of a human synthetic combinatorial library of single-chain antibodies.** *Methods Mol Biol* 2009, **525:**61-80, xiv.
.42 Wykes M: **Why do B cells produce CD40 ligand?** *Immunol Cell Biol* 2003, **81:**328-331.
.43 Berglund LJ, Avery DT, Ma CS, Moens L, Deenick EK, Bustamante J, Boisson-Dupuis S, Wong M, Adelstein S, Arkwright PD, et al: **IL-21 signalling via STAT3 primes human naive B cells to respond to IL-2 to enhance their differentiation into plasmablasts.** *Blood* 2013, **122:**3940-3950.
.44 Turchaninova MA, Davydov A, Britanova OV, Shugay M, Bikos V, Egorov ES, Kirgizova VI, Merzlyak EM, Staroverov DB, Bolotin DA, et al: **High-quality full-length immunoglobulin profiling with unique molecular barcoding.** *Nat Protoc* 2016, **11:**1599-1616.
.45 Stern JN, Yaari G, Vander Heiden JA, Church G, Donahue WF, Hintzen RQ, Huttner AJ, Laman JD, Nagra RM, Nylander A, et al: **B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes.** *Sci Transl Med* 2014, **6:**248ra107.
.46 Shiroguchi K, Jia TZ, Sims PA, Xie XS: **Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes.** *Proc Natl Acad Sci U S A* 2012, **1**.09:1347-1352
.47 Jiang N, He J, Weinstein JA, Penland L, Sasaki S, He XS, Dekker CL, Zheng NY, Huang M, Sullivan M, et al: **Lineage structure of the human antibody repertoire in response to influenza vaccination.** *Sci Transl Med* 2013, **5:**171ra119.
.48 Bashford-Rogers RJ, Palser AL, Idris SF, Carter L, Epstein M, Callard RE, Douek DC, Vassiliou GS, Follows GA, Hubank M, Kellam P: **Capturing needles in haystacks: a comparison of B-cell receptor sequencing methods.** *BMC Immunol* 2014, **15:**29.
.49 Hammond PW: **Accessing the human repertoire for broadly neutralizing HIV antibodies.** *MAbs*, **2:**157-164.
.50 Kuhn P, Fuhner V, Unkauf T, Moreira GM, Frenzel A, Miethe S, Hust M: **Recombinant antibodies for diagnostics and therapy against pathogens and toxins generated by phage display.** *Proteomics Clin Appl* 2016, **10:**922-948.

Figure legends

Figure 1. Scheme of work-flow. The work flow included the following steps: collection of blood samples from SC or CI infections, isolation of HCV-specific B-cells from chronic or cleared infections, sequencing total B-cell repertoires and HCV-specific B-cell repertoires from chronic or cleared infections, analyzing Abs repertoires and identifying Ab clones associated with viral clearance, construction of antibody phage display library, isolation of a panel of HCV-binding antibody sequences that associate with cleared infections, integration of all data to construct and cheacterize HCV-nAbs unique to infection clearance.

Figure 2. Characterization of sera from HCV-infected individuals. (A) HCV Abs binding to E1E2 protein performed by ELISA of CI (n=52) and SC (n=18). Each dot represents the OD per patient. Background of binding to BSA was subtracted from all samples. (B) HCVcc neutralization assay were performed with sera of CI (n=52) and SC (n=18) to screen for Abs that can neutralize HCV infection. Each dot represents the percentage of neutralization capacity per patient, compared to neutralization of sera from healthy control. (C,D) Characterizing HCV-Abs binding and neutralizing in sera obtained from two patients (CI21 and CI22) before and after anti-HCV treatment and following SVR by ELISA (P<?) (C) and by HCVcc neutralization assay (P<?) (D).

Figure 3. Identifying unique feature in the B cell repertories of resolved and chronic HCV infection. (A) Number of sequences per sample after pre-processing with pRESTO. The range of the sequences is between 1800 to 100000 per patient. Sequences were grouped by their IGHV-gene, IGHJ-gene and junction length, clustered by genetic distance, and frequencies within and between the clinical groups were compared. (B) Junction length distribution across the IGHV gene (C) Overview of the collective IGHV genes that were used by healthy control, SC and CI (P value <0.05). (D) IGHJ-genes frequency. Each clinical group represent by different color. סטטיסטיקה? (E-G) The Ig sequences were grouped by different gene combinations: IGHV X IGHJ, IGHV X IGHJ X junction length, and sequence clusters that were grouped by IGHV, IGHJ and junction similarity (P value <0.01). Dots indicate the gene's frequency per sample. Genes combinations that their abundance significantly differ between SC and CI groups (Ftest with P value < 0.05) is presented for: (E) IGHV X IGHJ genes (0.01<P<0.05). (F) IGHV X IGHJ X Junction length (0.001<P<0.01). (G) Combination of IGHV X IGHJ X Junction length, that are significantly different between the clinical groups (P<?).

Figure 4. Model for prediction of HCV infection outcome based on frequencies of IGHV clones using the machine-learning algorithm. (A) Prediction model based on B cells repertoire. Ordered

data represent clustered sequences base on identical IGHV and IGHJ from known clinical group were in the random bars the clinical groups is randomly shuffled and as a results the prediction rates significantly decreases (by half) in all clinical groups. (B) Prediction model based on T cells repertoire. The training for the T cell repertoires model is very similar to the B cell model, except that the data was categorized by DNA Junction sequences. (C,D) Identification of stratifying clones by means of clones that can accurately stratify between SC and CI samples. (C) In B cells clusters. (D) In T cells clusters.סטטיסטיקה

Figure 5. Isolation of HCV-specific B cells from resolved and chronic HCV infection. (A) The fold enrichment of cells isolated from six CI and three SC HCV-specific B cells. (B) HCVcc-neutralization assays using supernatant of cultured B-cells from healthy, SC17 and CI58 samples following 2 weeks activation in vitro and after isolation of HCV-specific B-cells. (C-F) Identifying unique features in HCV-specific B-cell repertoires. (C) Genomic distances between the binding sequences from the different samples by Levenshtein. The sequences obtained following sequencing of HCV binders from HCV-specific B cell repertoires from six chronic and three spontaneously cleared HCV-infections. (D) Mutations number in the general repertoire compared to the specific repertoire. Each specific sequence was compared to non-specific sequence which was randomly sampled from the general repertoire, the specific and non-specific sequences contain the same IGHV and IGHJ. The sequences were grouped by Isotype and mutations number in IGHV gene were compared, Mann Whitney test IGA p=3.488873e-07, IGG p=6.849511e-08, IGM p=3.764229e-04. (E) The specific sequences from the SC and CI were unified in different bulks following grouping by isotype and then the mutations number in the IGHV were compared, IGA p=0.000574, IGG p=0.435930. (F) Conserved amino acid in CDR3 region in HCV-specific repertoire compared to general repertoire. Against each VXJXCDR3-length binder sequence, random sequence with similar V, J and CDR3 length were selected from our Repertoires - defined as non-binder. Junctions AAs that were very conserved for binder sequences but not very conserved for non-binders were selected.

Figure 6. Identifying binder Ab sequences associated with HCV infection clearance. (A) Binding of phage displayed Abs to E2 protein by ELISA. Each bar indicates the fold change in the OD compared to BSA binding. (B) The distance between similar sequences from the general repertoire and the IGHV from the scFv. Differentiation and maturation of (C) IGHV6-1*IGHJ6*17**20, (D) IGHV4-39*IGHJ4*13**861. Phylogenetic trees, which illustrate the evolvement SHM in the IGHV gene of B cell clones from a putative germline sequence

Figure 7. Construction and characterization of unique antibodies correlating with infection clearance. (A) Binding of Abs VH510520 and VH16618 compared to the phage display Abs SC28 and SC11 by ELISA using 16mg/ml Ab. Each bar indicates the fold change in the OD binding (B) Binding of Abs VH16618 and VH 510520 to E1E2 compared to a well-defined panel of nAbs by ELISA using 16mg/ml Ab. (C and D) HCVcc neutralization assays were carried out with genotypes 1-7 using 50mg/ml of Abs VH16618 and VH510520.

Supplementary Figures

Supplementary Figure 1. General characterization of T cell repertoires of resolved and chronic HCV infection. (A) Number of sequences per sample after pre-processing with pRESTO. (B) Junction length distribution per sample, colored by clinical group.(C) The frequent isotype antibodies distribution of IgA, IgD, IgE, IgG and IgM, per clinical group. (D) Frequency of usage of J genes of TCR, colored by clinical groups ($P<0.005$). (E) Frequency of usage of V genes of TCR, colored by clinical groups ($P<0.005$).

Supplementary Figure 2. Mutation signatures in IGHV sequences in resolved and chronic HCV infections. Sequences were first subdivided into isotypes and then observed differences in the frequencies for each clinical groups. (A) The frequent isotype antibodies distribution of IgA, IgD, IgG and IgM, per clinical group. (B) Mutation number on IGHV genes within each isotype IgA, IgM, IgG and IgD. (C) Mutations numbers within each isotype between the clinical groups. (D) Differences in mutation numbers across significant IGHV genes between the clinical groups. סטטיסטיקה

Supplementary Figure 3. Enrichment of HCV-specific B cell population in vitro. For the in vitro proliferation of B cells, $CD19^+$ cells were isolated from PBMCs of healthy donor using a FACS sorter. Isolated B cells were labeled with CFSE, cultured in the presence of IL2, IL21 and feeder irradiated 3T3-msCD40L cells, and activated with a pool of positive peptides (ref) for 8 days. (A) CFSE prole of $CD19^+$ B cells. CFSE fading indicates proliferation of the activated culture. (B) Evaluating proliferation of memory B cells. In the activated culture, 23% of the population is memory B cells that are positive for $CD27^+$ (right panel), compared with very low numbers of $CD27^+$ cells in the non-activated culture (left panel). (C) Evaluating the ability of B cells to differentiate and produce IgGs. Concentration of IgG secreted to the culture medium 3 or 8 days following B cell activation were measured by ELISA.

Supplementary Figure 4. Isolation of HCV-specific B-cells from chronic HCV and healthy donor by FACS. Gating for E1E2 binding CD19+, CD27+ B cells from healthy control resolved and chronic HCV infection patients. $CD19^+$ B cells from SC17 and CI58 were grown with feeder irradiated 3T3-msCD40L cells and activated with E1E2 protein, IL2 and IL21 for 13-14 days. After 14 days, activated B cells were incubated with E1E2 and stained with CD19-PE, CD27-BV421 and tagged E1E2 (anti-cMyc, alexa fluor 633). Viable, $CD19^+$, $CD27^+$ and $HCsAg^+$ were isolated by FACS.

Supplementary Figure 5. Distance between scFv Abs sequences and clones from B-cell repertoires of resolved and chronic HCV infection. Each bar represent the distances between -scFv Ab sequence and clone of B cell repertoire from healthy control, CI and SC. The lower the distance the most similar the scFv Ab sequence. סטטיסטיקה