Drug Resistance of a Viral Population and Its Individual Intrahost Variants During the First 48 Hours of Therapy

DS Campo[1], P Skums[1], Z Dimitrova[1], G Vaughan[1], JC Forbi[1], CG Teo[1], Y Khudyakov[1] and DT-Y Lau[2]

Using hepatitis C virus (HCV) and interferon (IFN) resistance as a proof of concept, we have devised a new method for calculating the effect of a drug on a viral population, as well as the resistance of the population's individual intrahost variants. By means of next-generation sequencing, HCV variants were obtained from sera collected at nine time points from 16 patients during the first 48 h after injection of IFN-α. IFN-resistance coefficients were calculated for individual variants using changes in their relative frequencies, and for the entire intrahost viral population using changes in viral titer. Population-wide resistance and presence of IFN-resistant variants were highly associated with pegylated IFN-α2a/ribavirin treatment outcome at week 12 ($P = 3.78 \times 10^{-5}$ and 0.0114, respectively). This new method allows an accurate measurement of resistance based solely on changes in viral titer or the relative frequency of intrahost viral variants during a short observation time.

There has been great progress in the development of antiviral agents licensed for treatment of human immunodeficiency virus, herpes viruses, hepatitis viruses, and respiratory viruses.[1] The emergence of human immunodeficiency virus as a major human pathogen and the intensive use of antiretroviral compounds have also provided a better understanding of the genesis of antiviral resistance.[2] However, there is no simple method for directly measuring the effect that a drug has on a viral population and its individual intrahost variants. Such a measure could help screen promising drugs that affect the viral population and also detect the individual variants that are naturally more resistant. In this article, we use the effects of interferon (IFN) on hepatitis C virus (HCV) as a proof of concept, finding that our drug-resistance estimates, calculated during the first 48 h after IFN injection, are strongly associated with outcome of therapy at week 12.

HCV infects nearly 3% of the world's population and is a major cause of liver disease worldwide.[3] There is no vaccine against HCV and, until recently, the standard-of-care therapy involved the combined use of pegylated IFN (peg-IFN) and ribavirin. This combination therapy is expensive, effective in only 50–60% of patients, and can be associated with frequent and serious adverse side effects in more than 75% of patients.[4,5] To improve cost-effectiveness and ameliorate patient hardship, it would be desirable to predict the response at early onset of therapy.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

✓ There has been great progress in the development of antiviral agents. However, there is no simple method for directly measuring the effect that a drug has over a viral population and its individual intrahost variants.

WHAT QUESTION DID THIS STUDY ADDRESS?

✓ This study examined whether next-generation sequencing can be used to estimate the drug resistance of a viral population and its individual variants during early therapy.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

✓ Our estimates of population-wide resistance and presence of drug-resistant variants were highly associated with treatment outcome.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

✓ Because the model is based solely on changes in viral titer or the relative frequency of intrahost viral variants persisting during antiviral therapy, without consideration of other viral factors, host factors, or type of drug administered, it is not only applicable to the measurement and prediction of the drug response of HCV but also for other viral infections.

The first two authors contributed equally to this work.

[1]Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, Georgia, USA; [2]Liver Center, Division of Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA. Correspondence: DS Campo (fyv6@cdc.gov)

Received 7 November 2013; accepted 22 January 2014; advance online publication 2 April 2014. doi:10.1038/clpt.2014.20

IFNs are crucial components of the innate immune system. IFN-α acts by inducing production of IFN-stimulating genes to establish a nonspecific antiviral state within the cell with direct inhibition of viral replication. It is also known to exert immunomodulatory effects that enhance immune response and accelerate clearance of infected cells.[6] When exogenously administered as a single injection, IFN-α induces a decline of HCV RNA in blood in two phases: a rapid phase lasting 24–48 h, followed by a slower phase of decline over the ensuing weeks. The initial rapid decline is defined by the rate of viral clearance and the effectiveness of IFN in blocking viral production. Successful treatment results in sustained undetectable HCV RNA after completion of therapy. Treatment failure results either from nonresponse (minimal declines in viral titer during therapy) or relapse (robust initial responses, followed by rebounds of viral titers after termination of therapy).[6]

Several independent predictors of a sustained virological response to IFN/ribavirin therapy have been identified. These include HCV genotypes 2 and 3, low pretreatment viral load, Asian or Caucasian ethnicity, younger age, absence of advanced fibrosis or cirrhosis, and absence of steatosis.[7] More recently, genome-wide association studies have identified single-nucleotide polymorphisms near the *IL28B* gene (encoding IFN-λ3) as being particularly associated with spontaneous and treatment-induced clearance of HCV infection.[8,9] However, IL28B variations may account for only ~15% of interindividual variability of sustained virological response.[10] The interaction of these host factors determines the therapy's effect on the virus, which is directly evidenced by a decline in viral titer, the reason why the rate and magnitude of decline in the first weeks of treatment can predict the outcome of therapy.[6]

HCV exists in infected patients as a large viral population of intrahost variants, which may be differentially resistant to IFN treatment and therefore likely to display variable temporal patterns during the first phase of decline following IFN injection. Assessing the spectrum of HCV variants and measuring the IFN resistance of individual variants could be critical for understanding the variability in therapy outcomes. Next-generation sequencing technologies, in conjunction with computational analysis, allow for quantitative assessment of viral intrahost variants, providing data on the intrahost dynamics of individual HCV variants and an opportunity for measuring their resistance to IFN. The hypervariable region 1 (HVR1) region of HCV is used here as a tag or marker of individual intrahost viral strains for estimating their relative frequencies over a short period of time.

Because the model is based solely on changes in viral titer or the relative frequency of intrahost viral variants persisting during antiviral therapy, without consideration of other viral factors, host factors, or type of drug administered, it may be applicable to the measurement and prediction of HCV treatment response for other drug regimens, including the newly available direct-acting antiviral agents. In addition, the presented analytical framework should be applicable to the study of drug resistance of other viral infections.

RESULTS

Outcomes of therapy

The demographics and clinical features of the 16 study patients are summarized in **Table 1**, which includes their gender, race, grade of hepatic inflammation and fibrosis, IL28 genotype (see also **Figure 2a**), and therapy outcome according to the clinical criteria of treatment success[11,12]: (i) rapid virological response (RVR) is defined as having undetectable HCV RNA by week 4 of therapy ($n = 4$); (ii) complete early virological response (cEVR): undetectable HCV RNA by week 12 of therapy ($n = 6$); (iii) partial early virological response (pEVR): a decrease of HCV RNA more than 2 $\log_{10}$ but still detectable after 12 weeks of therapy ($n = 4$); and (iv) nonresponse (NR): decrease of HCV RNA less than 2 $\log_{10}$ after 12 weeks of therapy ($n = 2$). The inefficiency of IL28B as the sole predictor of IFN-based therapeutic response is highlighted in **Figure 2a**, which displays the number of patients by IL28B status and therapy outcome.

Decline of viral titer during the first 48 h after single-dose injection

HCV RNA levels were measured at each time point (**Figure 1**) and interpolated over the entire 48 h. The titer declined in all patients, starting from an average of 4.76 h after IFN injection (SD: 2.23 h), with the lowest point being observed at 30.03 h (SD: 4.66 h), and an average decrease of 1.67 $\log_{10}$ IU/ml. The

Table 1 Characteristics of study patients

Patient	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	M	M	F	M	M	M	M	F	F	M	M	M	F	M	M	F
Race	B	W	W	W	W	W	W	W	B	W	W	W	W	W	B	W
Age (years)	51	56	60	59	22	54	49	27	56	47	25	30	46	50	36	50
Hepatic inflammation	1	2	2	2	2	2	1	1	2	2	2	1	2	2	2	2
Hepatic fibrosis	1	3	3	1	1	3	1	0	3	0	0	0	2	2	0	1
IL28B genotype	CC	CT	TT	CC	CC	CC	CC	CT	TT	CT	TT	CC	TT	CT	CT	TT
Outcome at week 12	NR	pEVR	NR	RVR	RVR	cEVR	cEVR	cEVR	pEVR	cEVR	RVR	RVR	pEVR	pEVR	cEVR	cEVR

B, black; cEVR, complete early virological response; F, female; M, male; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response; W, white.

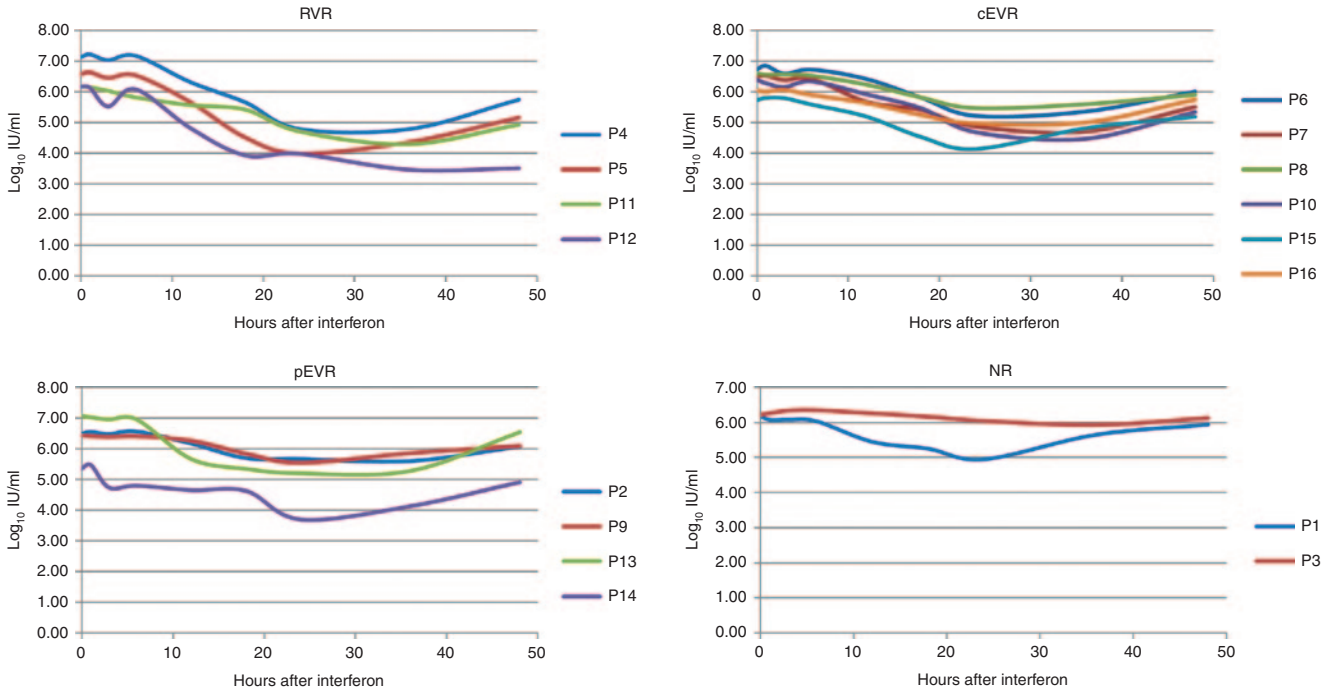

Figure 1 Hepatitis C virus RNA levels across all time points. Patients are grouped according to treatment outcomes. cEVR, complete early virological response; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response.

HCV RNA titer increased at 48 h in all patients, with an average increase of 0.86 $\log_{10}$ IU/ml as compared with the lowest titer, which likely corresponded to the gradual clearance of the administered IFN (**Figure 1**). All calculations conducted in this study used all time points, but we found four time points (0, 24, 36, and 48 h) to be particularly critical for the calculation of resistance. This is due to the fact that the most important changes in viral titer occur around these time points.

Association of population IFN resistance with outcome of treatment at week 12

The population IFN-resistance coefficient was calculated using the rate of titer change as described in Eq. 3. A positive coefficient value indicates IFN resistance, whereas a negative value indicates IFN sensitivity. **Figure 2b** shows the association of the coefficient values with therapy outcomes. The lowest coefficient value was calculated for RVR, followed by those for cEVR and pEVR, with the highest value calculated for NR. The average population IFN-resistance coefficients for the four outcomes were significantly different (P value = 3.78×10^{-5}), being 3.03 times higher for nonresponders (pEVR and NR) than for responders (RVR and cEVR; P value = 0.0040; **Figure 2b**). Thus, the resistance coefficients calculated using data from the first 48 h strongly correlated with therapy outcome at week 12.

Stability of intrahost HCV subpopulations

Given the strong association between population IFN-resistance coefficient values and therapy outcome, we investigated whether this association could be attributed to resistance of individual HCV variants. The HCV E1/E2 region from all samples was subjected to deep sequencing, generating a total of 1,763,502 reads, with an average of 12,332 reads per time point (**Figure 2d**). Many low-frequency intrahost viral variants could not be detected at all time points, with numerous variants detectable only at a single time point. However, the spectrum of frequent intrahost HCV variants was stable. On average, 76.81% (SD: 17.66%) of all error-corrected reads found at 48 h were also present before therapy. The less persistent variants were usually found at very low frequencies, suggesting that their absence at some time points is probably due to stochastic sampling. The nucleotide diversity at each time point was calculated and was also found to be constant in all patients, with the SD of the nucleotide diversity ranging from 0.0001 in Patient 2 to 0.0139 in Patient 14.

Phylogenetic analysis showed that major viral subpopulations were constantly present during the entire sampling period in all patients. However, many low-frequency variants were detectable only at certain time points. Genetic differences between time points were measured for each patient using Φst. Viral populations found at the first and last time points were very similar in most patients; the average Φst was only 3.52%, with only one patient (Patient 6) showing a high Φst value (71.14%; **Table 2**).

Association between presence of IFN-resistant variants and therapy outcome

Sequences of intrahost HCV variants were not expected to accrue mutations within 48 h of observation. However, we found that the relative frequencies of persistent variants changed during this time period in all patients (**Figure 3**). The developed mathematical model allowed calculation of the IFN-resistance coefficient for each persisting intrahost HVR1 variant based on its relative frequency variations. IFN resistance varied broadly among intrahost variants in all patients.

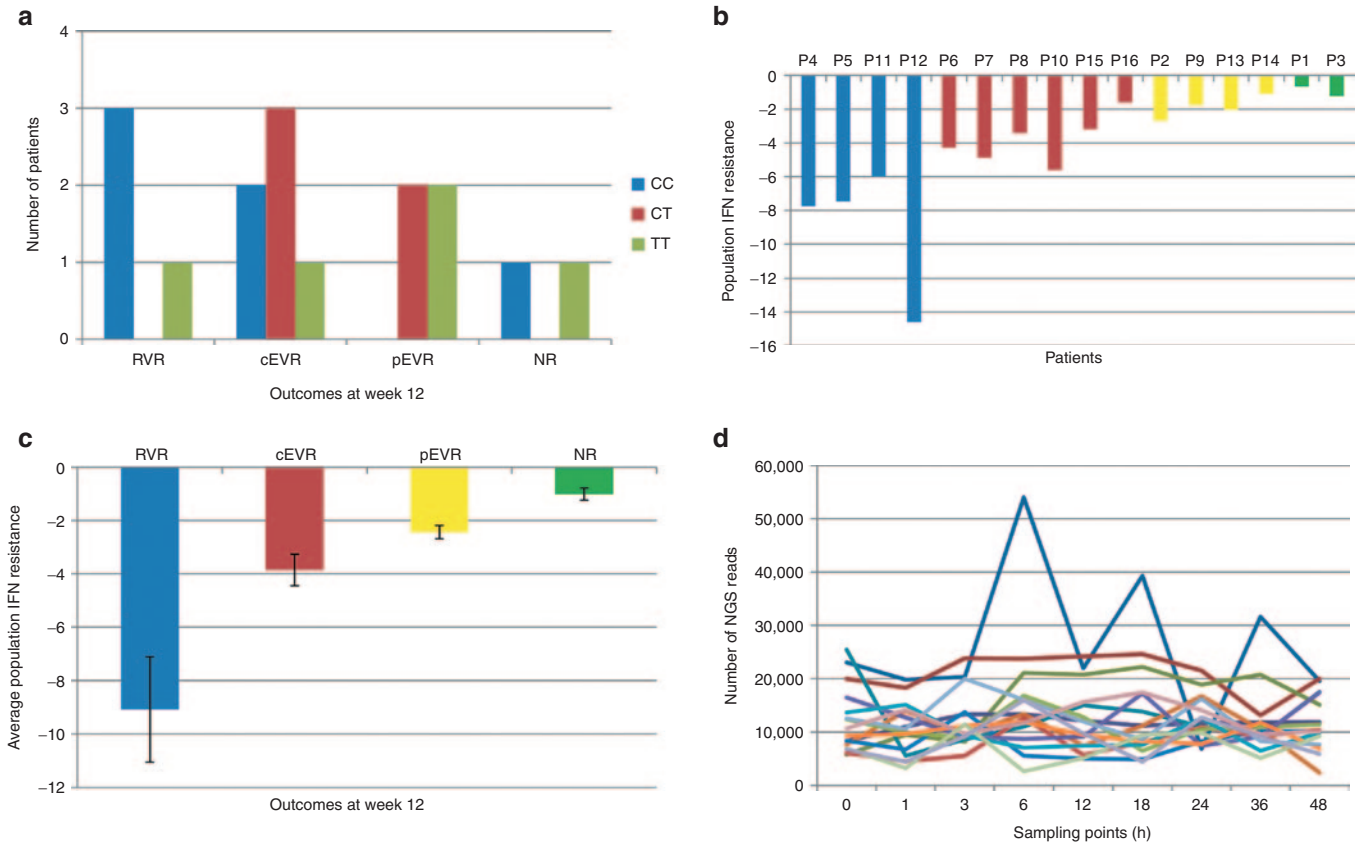

Figure 2 Description of patients and outcome group. (**a**) Number of patients by IL28B status and therapy outcome. (**b**) Population interferon (IFN) resistance for each patient. Blue, rapid virological response (RVR); red, complete early virological response (cEVR); yellow, partial early virological response (pEVR); and green, nonresponse (NR). (**c**) Average population IFN resistance of patients for each outcome; bars correspond to standard error of the mean. (**d**) Number of reads obtained by next-generation sequencing (NGS) from each patient and at each time point.

Not a single IFN-resistant variant was found in patients with RVR, and only one of six cEVR patients showed IFN-resistant variants (**Table 2** and **Figure 4a**). By contrast, patients with pEVR and NR were infected with an average number of 3.25 and 3.5 IFN-resistant intrahost variants, respectively (**Figure 4a**). The average number of IFN-resistant variants was significantly associated with the four outcomes ($P = 0.0114$), being 16.6 times lower for nonresponders (pEVR and NR) than for responders (RVR and cEVR; $P = 0.0014$). The average total fraction of reads with positive IFN-resistance values was also significantly different among the four outcomes ($P = 0.0014$; **Figure 4c**). These observations suggest that the absence or presence of IFN-resistant variants and their frequency before single-dose IFN injection are strong predictors of treatment outcome.

In all patients, the variant with the highest initial frequency (henceforth referred to as the major variant) showed a strong sensitivity to IFN, with an average value of −4.75 for all patients (**Table 2**). The average IFN resistance of the major variants for the four outcomes was significantly different ($P = 0.0071$), being 3.68 times higher for nonresponders (pEVR and NR) than for responders (RVR and cEVR; $P = 0.0028$; **Figure 4b**). There was no association between therapy outcome and the initial frequency of the IFN-sensitive major variant ($P = 0.1535$). The maximum IFN resistance found in each patient was significantly different among the four outcomes ($P = 0.0188$), being 5.25 times higher for nonresponders (pEVR and NR) than for responders (RVR and cEVR; $P = 0.0084$; **Figure 4d**).

Phylogenetic analysis of sensitive and resistant variants

For all patients showing one or more IFN-resistant variants, we performed a phylogenetic analysis to ascertain the genetic relatedness between the sensitive and resistant variants. For each patient, resistant variants were always different from the sensitive variants in the same cluster by one or two nucleotides. In addition, resistant sequences did not cluster together but were mostly scattered across other subpopulations (**Figure 5**).

DISCUSSION

Viral kinetic modeling has played an important role in analysis of HCV decay in peripheral blood (as measured by serum or plasma HCV RNA) after initiation of antiviral therapy.[13–15] Kinetic models usually include such parameters as viral load, number of target cells, number of infected cells producing virions, target cell production rate, target cell death rate, *de novo* infection rate constant, infected cell death rate, HCV virion production rate, and HCV virion clearance rate. A recent study[15] of 2,100 patients from two clinical trials further developed a kinetic

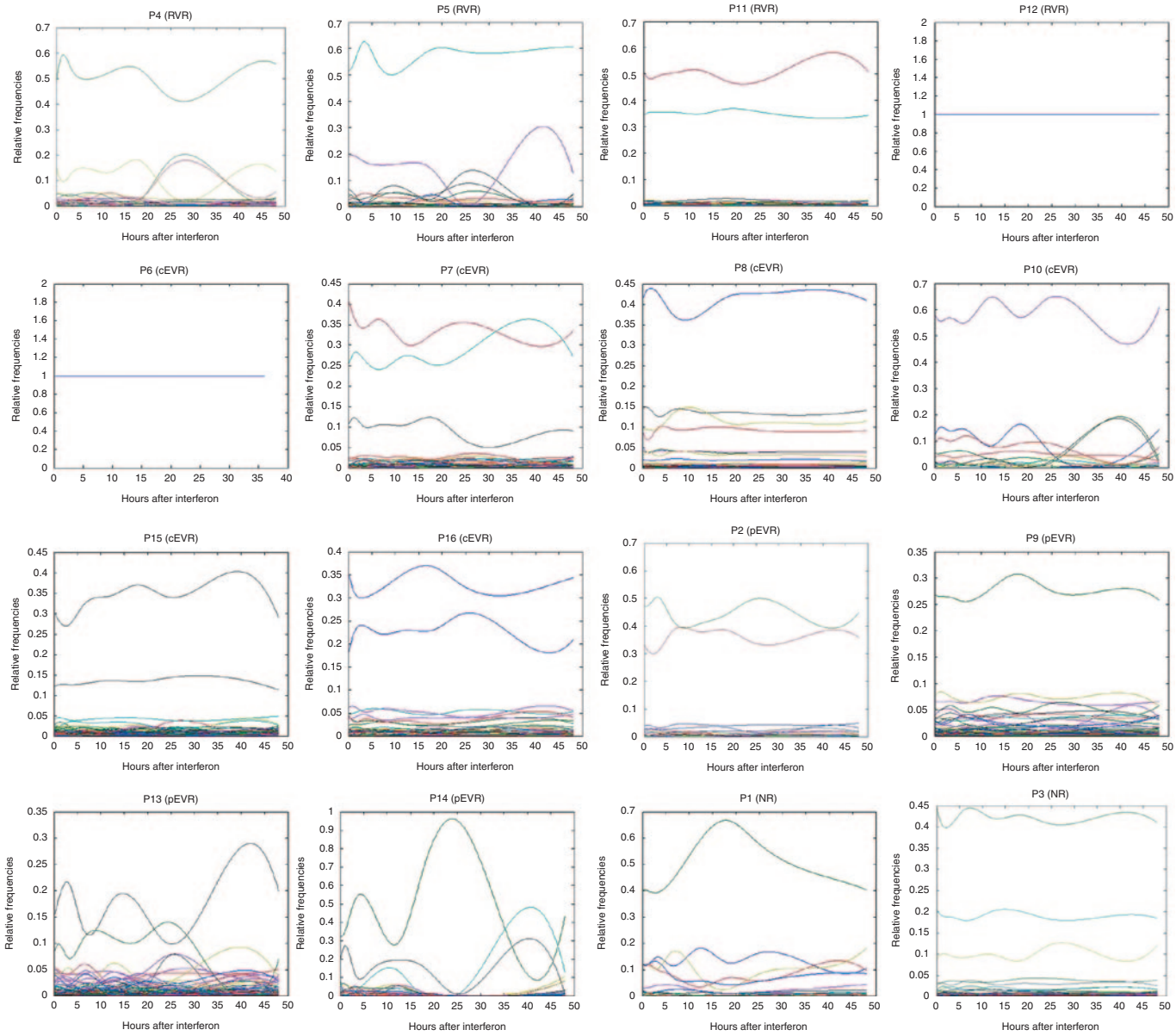

Figure 3 Relative frequencies of persistent variants over time in all patients. cEVR, complete early virological response; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response.

analysis based on a model that incorporates variables such as liver regeneration, HCV RNA below the threshold of detection, and a cure boundary. Although useful for understanding HCV pathogenesis and replication, such models require substantial *a priori* knowledge and assumptions, in addition to data collected over 24–72 weeks of therapy.

It is the interaction of several host factors that determines the therapy's effect on the virus, but such effect is directly evidenced by a decline in viral titer, the reason why the rate and magnitude of decline in the first weeks of treatment can predict the outcome of therapy.[6] The decline in viral titer observed among the patients during the first 48 h after a single IFN injection indicates that their intrahost viral populations were under considerable selective pressure after the medication. Although the spectrum of intrahost HCV variants was relatively stable during that period, we found that the frequencies of the persistent HCV variants were variable. This allowed the degree of IFN resistance to be calculated for the individual intrahost HCV variants. We took advantage of the capacity of next-generation sequencing to produce a massive number of sequences, facilitating accurate and robust assessment of the frequency of viral variants.

All calculations conducted here used nine time points, but the collection and sequencing of serum samples at such frequency is difficult for routine research or eventual patient evaluation. However, we found four time points—0, 24, 36, and 48 h—to be sufficient for the calculation of resistance. Viral titer at these time points reflects most closely the effect of IFN on viral population. It should be noted that the specific time points that are most critical for estimation of resistance coefficients may differ for other drugs from those found in this study, owing to

Table 2 IFN resistance of the individual variants in each patient

Patient	Outcome	Population IFN resistance	Φst	*n*	*n* Resistant	Frequency of major variant	IFN resistance of major variant	Fraction of resistant reads	Minimum IFN resistance	Maximum IFN resistance
P4	RVR	−7.9	0.72	29	0	0.48	−7.63	0	−11.26	−5.86
P5	RVR	−7.58	0.49	23	0	0.55	−7.33	0	−10.05	−5.08
P11	RVR	−5.97	0.07	24	0	0.51	−5.86	0	−7.49	−3.43
P12	RVR	−14.85	13.08	1	0	1	−14.63	0	−14.63	−14.63
P6	cEVR	−4.27	71.14	1	0	1	−11.45	0	−11.45	−11.45
P7	cEVR	−5.03	4.31	40	0	0.41	−5.21	0	−7.08	−2.84
P8	cEVR	−3.4	0.29	42	0	0.41	−3.49	0	−6.94	−0.4
P10	cEVR	−5.66	0.82	12	0	0.66	−5.43	0	−7.42	−2.56
P15	cEVR	−3.14	1.37	75	0	0.31	−3.02	0	−5.62	−0.19
P16	cEVR	−1.6	0	41	2	0.36	−1.3	0.0048	−4.01	1.45
P2	pEVR	−2.72	0.06	15	0	0.48	−2.56	0	−5.7	−1.05
P9	pEVR	−1.74	0.48	85	7	0.27	−1.72	0.0077	−5.09	1.44
P13	pEVR	−2.41	3.93	86	3	0.15	−1.94	0.0074	−6.47	1.88
P14	pEVR	−2.88	22.85	25	3	0.34	−2.35	0.1473	−9.59	1.78
P1	NR	−0.79	0.86	14	4	0.41	−0.77	0.1465	−1.72	0.34
P3	NR	−1.22	0	50	3	0.45	−1.31	0.0065	−4.49	0.74

cEVR, complete early virological response; IFN, interferon; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response.

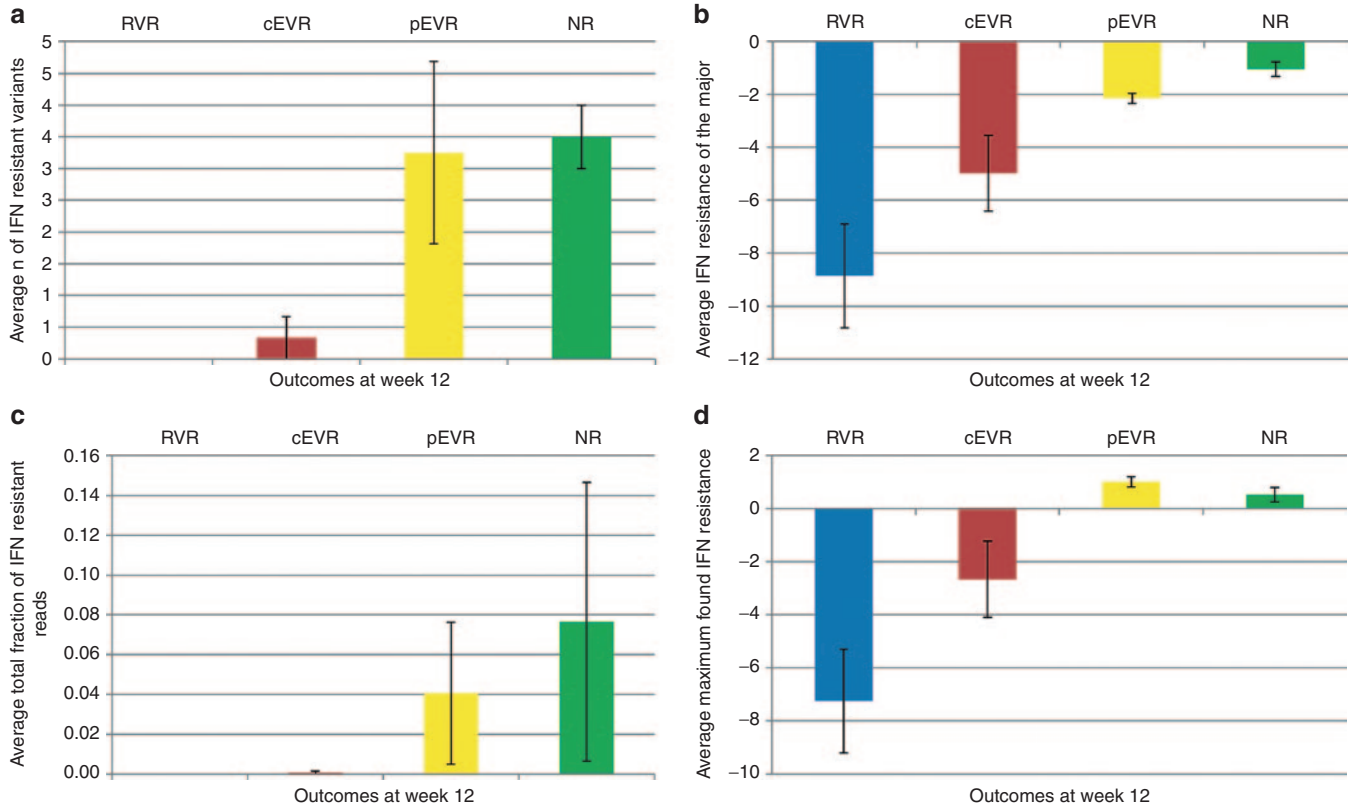

Figure 4 Average interferon (IFN) resistance calculated for the patients of each therapy outcome. (**a**) Number of IFN-resistant variants. (**b**) IFN-resistance coefficient of the major variant. (**c**) Fraction of the total number of reads that are IFN resistant. (**d**) Maximum IFN resistance found. Bars correspond to standard error of the mean. cEVR, complete early virological response; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response.

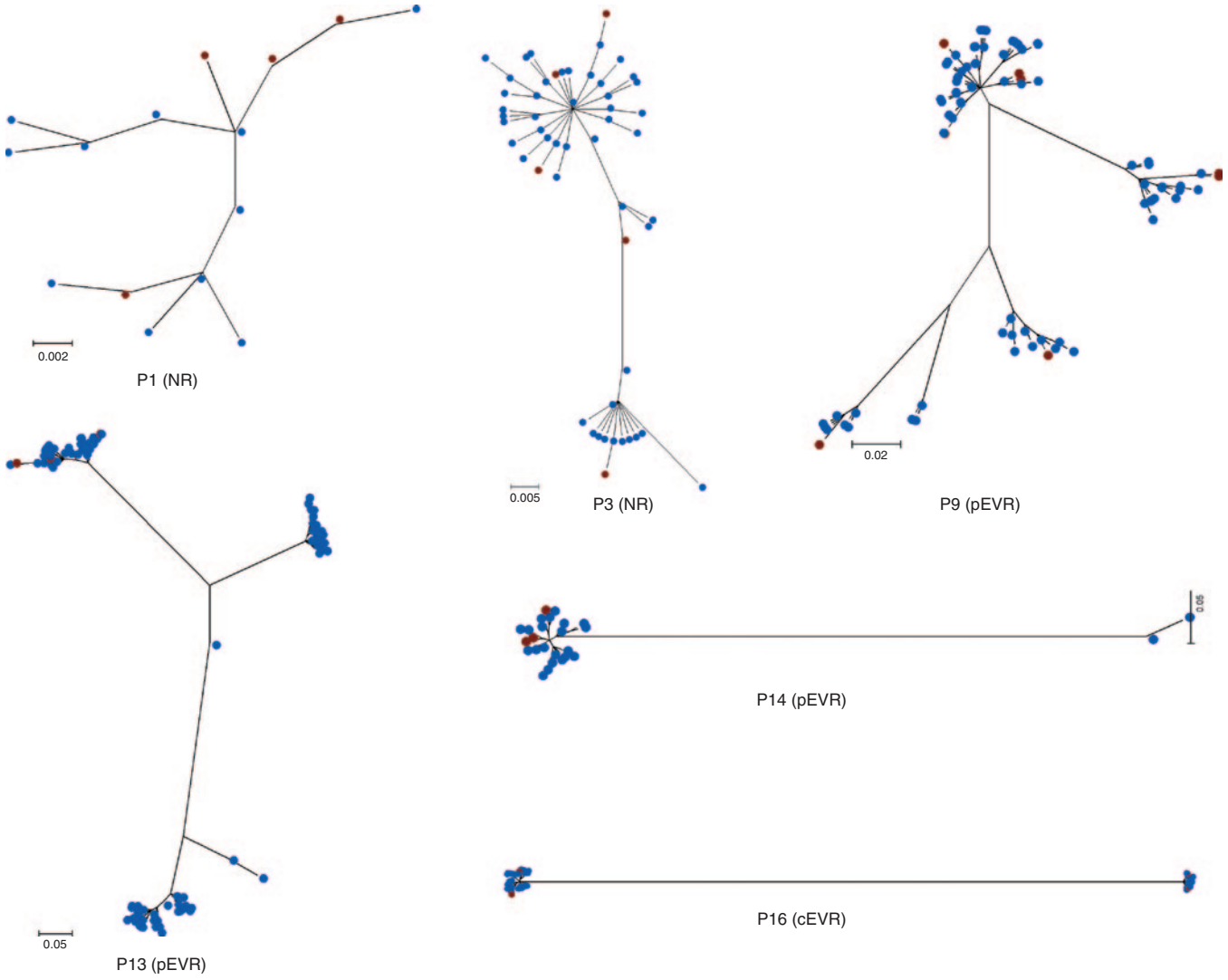

Figure 5 Maximum-likelihood trees of all the sequences from patients with resistant variants. Red, interferon (IFN)-resistant variants; blue, IFN-sensitive variants. cEVR, complete early virological response; NR, nonresponse; pEVR, partial early virological response; RVR, rapid virological response.

differences in mechanisms of action and clearance kinetics of drugs.

HVR1 is used here as a tag or marker of individual intrahost viral strains for estimating the relative frequency of intrahost HCV genomic variants. However, this relatively short region may be shared by more than one genetically distinct genomic variant. Nonetheless, HVR1 is the most heterogeneous region of the viral genome, so the probability of such sharing is minimized. Furthermore, we previously demonstrated that coordinated evolution among sites from the entire HCV genome is strongly associated with IFN resistance, and that sequence polymorphisms in short genomic regions, including HVR1, can accurately reflect that association.[16]

Phylogenetic analysis showed that IFN-resistant variants were different from sensitive variants by one or two nucleotides, but both resistant and sensitive variants belonged to the same cluster, rather than being segregated. The association between variation in the HVR1 sequence and therapy outcome is therefore unlikely to be mediated by the few mutations in HVR1 but rather by other regions of the HCV genome, with HVR1 heterogeneity reflecting genetic variation in these regions, as we have previously shown.[16,17]

Patient 16 was the only one from the group of responders (four RVR and six cEVR patients) who carried IFN-resistant variants with a titer value similar to those of pEVR and NR patients. The reasons for this discrepancy between the outcomes and our estimates are difficult to ascertain, although it was observed that this patient lost 62.5% of the titer within 2 days before commencement of IFN treatment. The results raise the interesting possibility that this patient cleared the HCV infection independently of the IFN treatment.

This study is the first to apply a mathematical model using empirically derived viral kinetic data with viral sequences linked to the frequency of their persistence. We found that frequency changes of the viral population associate strongly with treatment outcome. Because the model is based solely on changes in viral titer or the relative frequency of intrahost viral variants persisting during antiviral therapy, without consideration of other viral

factors, host factors, or type of drug administered, it is applicable to the measurement and prediction of HCV treatment response for other drug regimens, including the newly available direct-acting antiviral agents. The analytical framework developed here should also be applicable to the drug response of other viral infections, as long as informative samples are collected at time points reflecting variation in viral titer associated with the action of specific drugs on the virus under investigation.

METHODS

Patients. Sixteen treatment-naive patients with genotype 1 chronic hepatitis C were studied. Blood samples were collected from each patient over 48 h (at time points 0 (baseline), 1, 3, 6, 12, 18, 24, 36, and 48 h) after a single subcutaneous injection of IFN-α (10 MU). Patients then received a course of peg-IFN-α2a and ribavirin starting 48 h after the single dose of IFN-α. Therapy continued for up to 48 weeks per standard-of-care recommendation (see Lau *et al.*[18] for more details). Ethical review and informed consent were granted by the institutional review boards of Beth Israel Deaconess Medical Center and the Centers for Disease Control and Prevention.

Nucleic acid extraction. Total nucleic acids from the specimens were extracted from serum using the Roche MagNA Pure LC instrument and the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche Diagnostics, Mannheim, Germany). RNA was precipitated and reverse-transcribed using both random and specific primers as previously described.[19] PCR quantification was conducted by the COBAS AmpliPrep/COBAS Taq-Man HCV Test (Roche Diagnostics), and the HCV genotype was determined using the VERSANT HCV Genotype 2.0 Assay (LiPA) (Innogenetics NV, Ghent, Belgium).

HVR1 complementary DNA amplification. The E1/E2 junction of the HCV genome (309 nt), which contains the HVR1 region, was amplified using our nested PCR protocol as previously described.[19] The amplicons generated during first-round PCR were used as templates for nested PCR using hybrid primers composed of primer adapters, multiple identifiers, and specific sequences complementary to the HCV genome. This strategy allowed for multiplexing and downstream pyrosequencing. Resulting amplicons were quantified using the Picogreen kit (Invitrogen, Carlsbad, CA). Integrity of each fragment was evaluated using Bioanalyzer 2100 (Agilent, Santa Clara, CA).

Next-generation sequencing. PCR products were pooled and subjected to pyrosequencing using the GS FLX System and the GS FLX Titanium Sequencing Kit (454 Life Sciences, Roche, Branford, CT). Low-quality reads were removed using the GS Run Processor v2.3 (Roche). Initial reads were processed by matching to the corresponding identifier. The 454 files were processed using the error correction algorithms KEC and ET,[20] which have been validated as highly accurate in finding true haplotypes, removing false haplotypes, and estimating the frequency of true haplotypes. The error-corrected files were aligned using Muscle,[21] and the HVR1 sequences were clipped to 293 bp.

Genetic structure analysis. Unbiased estimates of nucleotide diversity were calculated according to Nei[22] using ARLEQUIN.[23] We also measured the genetic differences between the first time point (0 h) and the last time point (48 h) according to Excoffier *et al.*[24] by means of ARLEQUIN.[23] The genetic structure was analyzed with consideration of the molecular differences between sequences, in addition to differences in their frequencies, resulting in estimates of Φst, a measure of the percentage of genetic heterogeneity due to differences between two samples. Significance of the differences was estimated by use of a permutation test (n = 10,000). A maximum-likelihood tree was built for the persistent variants of each patient by means of the software HyPhy,[25] using the General Time Reversible model. A median-joining network was also built for each patient using the program NETWORK 4.6 (ref. 26). The median-joining network method begins computing the minimum spanning trees (a graph that connects all sequences with the minimum total length of the branches), and then all these graphs are combined within a single (reticulate) network.[26]

Nucleotide sequence accession numbers. The E1/E2 sequences produced in this study have been deposited into the National Center for Biotechnology Information GenBank database under accession numbers KC562310–KC562900.

IFN-resistance coefficients. The data sets under consideration are the HCV titer and the sets of E1/E2 sequences, along with their relative frequencies for each time point. To obtain reliable results in the calculation of IFN resistance for individual variants, only sequences present at every time point were analyzed. The average fraction of persistent variants for all patients was 61.4% (ranging from 10.9% for Patient 6 to 94.9% for Patient 8).

The dynamics of the viral population consisting of n variants was described by the following system of differential equations[27]:

$$\dot{u}_i(t) = g_i(t)u_i(t) - \frac{u_i(t)(\dot{h}(t))}{h(t)},\ i = 1,\ldots,n\,. \tag{1}$$

Here, $u_i(t)$ and $g_i(t)$ are the frequency and the fitness of a variant i at time t, respectively, and $h(t)$ is the titer. The functions, $u_i(t)$, $i = 1,\ldots, n$ and $h(t)$ were interpolated on the whole segment from 0 to 48 h using cubic splines, and the derivatives $\dot{u}_i(t)$ and $\dot{h}(t)$ were calculated using the obtained spline approximations.

The average fitness during the studied time interval was used as a measure of the IFN resistance of each variant. Using the expressions for $g_i(t)$ obtained from Eq. (1), the individual variant IFN-resistance coefficient r_i for each variant, i, was calculated using the following formula:

$$r_i = \frac{1}{T - t_0}\int_{t_0}^{T}\left(\frac{\dot{u}_i(t)}{u_i(t)} + \frac{\dot{h}(t)}{h(t)}\right)dt \tag{2}$$

Here T = 48 h and t_0 is the time when IFN begins affecting the viral population, which is estimated as the point when the titer starts declining monotonically in the obtained spline approximation. The population IFN-resistance coefficient was calculated by the formula

$$r = \frac{1}{T - t_0}\int_{t_0}^{T}\frac{\dot{h}(t)}{h(t)}dt \tag{3}$$

To evaluate the estimates' quality of individual variants resistance, it was assumed that the correctly estimated fitnesses, $u_i(t)$, changed smoothly at the interval [0, 48]. Abrupt changes of $u_i(t)$ may indicate errors in the estimation of variant frequencies due to PCR bias, sequencing inaccuracies, or both. The total variation of fitness on the interval $[0,T]$ for each variant was used as a measure of the possible error and was calculated using the following formula:

$$V_0^T(g_i) = \int_0^T \left|\dot{g}_i(t)\right| dt \tag{4}$$

The variants with a variation greater than mean fitness variation over the whole population plus 2 SDs were excluded from analysis.

The justification of the model and derivations of formulas have been described in detail by Skums *et al.*[27] All calculations were made in MATLAB R2010b (MathWorks, Natick, MA).

Statistical analysis. We tested the differences in IFN resistance among therapy outcomes by means of a multiresponse permutation procedure.[28,29] Multiresponse permutation procedure is a permutation version of the t-test, a nonparametric test for testing the hypothesis of no difference between two groups of paired samples. Permutation tests

represent ideal situations in which exact probabilities associated with a test statistic may be derived, rather than approximated from, common probability distributions such as t, F, and χ^2 (ref. 30). In the majority of studies, the population distribution is unknown and assuming a normal distribution is inappropriate for many biological data sets, which often are skewed, discontinuous, and multimodal. We used the multiresponse permutation procedure implemented in BLOSSOM,[28] using exact probabilities (all possible permutations) and V parameter equal to 1.

ACKNOWLEDGMENTS

This study was supported by National Institutes of Health grants R01 DK068598-01 (to D.T.-Y.L.), M01-RR-01032, and M01-RR-01032 (to the General Clinical Research Center).

AUTHOR CONTRIBUTIONS

Y.K. wrote the manuscript. D.S.C., C.G.T., Y.K., and D.T.-Y.L. designed the research. D.S.C., P.S., Z.D., G.V., and J.C.F. performed the research. D.S.C., P.S., and Z.D. analyzed the data. D.T.-Y.L. contributed new reagents/analytical tools.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

1. Said, Z. & Abdelwahab, K. Antiviral replication agents. In: *Viral Replication.* 127–144 (InTech, 2013) (doi: 10.5772/53707).
2. Zambon, M.C. Surveillance for antiviral resistance. *Influenza Other Respir. Viruses* **7**(suppl. 1), 37–43 (2013).
3. Mohd Hanafiah, K., Groeger, J., Flaxman, A.D. & Wiersma, S.T. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* **57**, 1333–1342 (2013).
4. Hsu, C.S., Liu, C.J., Lai, M.Y., Chen, P.J., Kao, J.H. & Chen, D.S. Early viral kinetics during treatment of chronic hepatitis C virus infection with pegylated interferon alpha plus ribavirin in Taiwan. *Intervirology* **50**, 310–315 (2007).
5. Aghemo, A., Rumi, M.G. & Colombo, M. Pegylated IFN-alpha2a and ribavirin in the treatment of hepatitis C. *Expert Rev. Anti. Infect. Ther.* **7**, 925–935 (2009).
6. Feld, J.J. & Hoofnagle, J.H. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature* **436**, 967–972 (2005).
7. Kau, A., Vermehren, J. & Sarrazin, C. Treatment predictors of a sustained virologic response in hepatitis B and C. *J. Hepatol.* **49**, 634–651 (2008).
8. Chevaliez, S. & Asselah, T. Mechanisms of non-response to antiviral treatment in chronic hepatitis C. *Clin. Res. Hepatol. Gastroenterol.* **35**(suppl. 1), S31–S41 (2011).
9. Lange, C.M. & Zeuzem, S. IL28B single nucleotide polymorphisms in the treatment of hepatitis C. *J. Hepatol.* **55**, 692–701 (2011).
10. Thompson, A.J. *et al*. Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* **139**, 120–9.e18 (2010).
11. Paulon, E. & Naoumov, N.V. Individualization of antiviral treatment regimens for chronic hepatitis C. *Eur. J. Gastroenterol. Hepatol.* **18**, 321–325 (2006).
12. Pawlotsky, J.M. Molecular diagnosis of viral hepatitis. *Gastroenterology* **122**, 1554–1568 (2002).
13. Zeuzem, S. The kinetics of hepatitis C virus infection. *Clin. Liver Dis.* **5**, 917–930 (2001).
14. Neumann, A.U. *et al*. Hepatitis C viral dynamics *in vivo* and the antiviral efficacy of interferon-alpha therapy. *Science* **282**, 103–107 (1998).
15. Snoeck, E. *et al*. A comprehensive hepatitis C viral kinetic model explaining cure. *Clin. Pharmacol. Ther.* **87**, 706–713 (2010).
16. Lara, J., Xia, G., Purdy, M. & Khudyakov, Y. Coevolution of the hepatitis C virus polyprotein sites in patients on combined pegylated interferon and ribavirin therapy. *J. Virol.* **85**, 3649–3663 (2011).
17. Campo, D.S., Dimitrova, Z., Mitchell, R.J., Lara, J. & Khudyakov, Y. Coordinated evolution of the hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **105**, 9685–9690 (2008).
18. Lau, D.T. *et al*. Innate immune tolerance and the role of kupffer cells in differential responses to interferon therapy among patients with HCV genotype 1 infection. *Gastroenterology* **144**, 402–413.e12 (2013).
19. Ramachandran, S., Xia, G.L., Ganova-Raeva, L.M., Nainan, O.V. & Khudyakov, Y. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: performance under optimal and suboptimal conditions. *J. Virol. Methods* **151**, 217–224 (2008).
20. Skums, P. *et al*. Efficient error correction for next-generation sequencing of viral amplicons. *BMC Bioinformatics* **13**(suppl. 10), S6 (2012).
21. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
22. Nei, M. *Molecular Evolutionary Genetics.* (Columbia University Press, New York, 1987).
23. Schneider, S., Roessli, D. & Excoffier, L. ARLEQUIN. (Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland, 2000).
24. Excoffier, L., Smouse, P.E. & Quattro, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491 (1992).
25. Pond, S.L., Frost, S.D. & Muse, S.V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
26. Bandelt, H.J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
27. Skums, P., Campo, D.S., Dimitrova, Z., Vaughan, G., Lau, D.T. & Khudyakov, Y. Numerical detection, measuring and analysis of differential interferon resistance for individual HCV intra-host variants and its influence on the therapy response. *In Silico Biol.* **11**, 263–269 (2011).
28. Cade, B. & Richards, J. *User manual for BLOSSOM statistical software.* (Midcontinent Ecological Science Center US Geological Survey, 2001).
29. Mielke, P. & Berry, K. *Permutation Methods: A Distance Function Approach.* (Springer-Verlag, New York, 2001).
30. Cai, L. Multi-response permutation procedure as an alternative to the analysis of variance: an SPSS implementation. University of North Carolina Department of Psychology, 2004.