In Silico Biology 11 (2011/2012) 263–269
DOI 10.3233/ISB-2012-0460
IOS Press

Numerical detection, measuring and analysis of differential interferon resistance for individual HCV intra-host variants and its influence on the therapy response

Pavel Skums[a,*], David S. Campo[a], Zoya Dimitrova[a], Gilberto Vaughan[a], Daryl T. Lau[b] and Yury Khudyakov[a]
[a]*Division of Viral Hepatitis, Centers of Disease Control and Prevention, Atlanta, GA, USA*
E-mails: {kki8, fyv6, izd7, jiv9, yek0}@cdc.gov
[b]*Translational Liver Research, Division of Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA*
E-mail: dlau@bidmc.harvard.edu

Received 17 January 2012
Revised 14 June 2012
Accepted 7 July 2012

Abstract. Hepatitis C virus (HCV) is a major cause of liver disease world-wide. Current interferon and ribavirin (IFN/RBV) therapy is effective in 50%–60% of patients. HCV exists in infected patients as a large viral population of intra-host variants (quasispecies), which may be differentially resistant to interferon treatment. We present a method for measuring differential interferon resistance of HCV quasispecies based on mathematical modeling and analysis of HCV population dynamics during the first hours of interferon therapy. The mathematical models showed that individual intra-host HCV variants have a wide range of resistance to IFN treatment in each patient. Analysis of differential IFN resistance among intra-host HCV variants allows for accurate prediction of response to IFN therapy. The models strongly suggest that resistance to interferon may vary broadly among closely related variants in infected hosts and therapy outcome may be defined by a single or a few variants irrespective of their frequency in the intra-host HCV population before treatment.

Keywords: Quasispecies, HCV, interferon resistance

1. Introduction

Hepatitis C virus (HCV) infects 2.2% of the world's population and is a major cause of liver disease world-wide [1]. There is no vaccine against HCV and current interferon and ribavirin (IFN/RBV) therapy is effective in 50%–60% of patients [5]. Interferon therapy is expensive and poses a significant hardship for patients. To improve cost effectiveness and reduce patient hardship, it is important to be able to predict response to therapy at its beginning.

Upon IFN injection, the HCV load rapidly declines during the first phase lasting for 24–48 hours followed by a slower second phase. The initial rapid decline is defined by the rate of viral clearance and IFN effectiveness of blocking viral production. Speed and magnitude of reduction of HCV titer in the first hours of treatment are used in clinical practice to predict the outcome of therapy.

HCV exists in infected patients as a large viral population of intra-host variants (quasispecies)[2], which

*Corresponding author: P. Skums, Division of Viral Hepatitis, Centers of Disease Control and Prevention, 1600 Clifton Road NE, 30329, Atlanta, GA, USA. Tel.: +1 404 639 3803.

may be differentially resistant to IFN treatment and, therefore, should display variable temporal patterns during the first phase. This differential resistance is crucial for the outcome of therapy. Consider a hypothetical example: there are two patients, one of whom was infected with a single HCV variant that has a low sensitivity to IFN and the second patient was infected with a dominant IFN-sensitive variant and minor IFN-resistant variant. Even though both patients have similar HCV titers, which equally decrease during the first hours of treatment, the long-term outcome of therapy can be different in these 2 patients. If no IFN resistant mutations occurred, the first patient may clear the virus, while the IFN-resistant variant may grow dominant in the second patient causing therapy failure.

As this hypothetical example shows, detecting HCV quasispecies and measuring and analyzing their differential IFN resistance may be critical for understandingvariations in therapy outcomes. Next-generation sequencing technologies in conjunction with computational analysis allow for quantitative assessment of viral quasispecies, providing data on the intra-host dynamics of individual HCV variants and opportunity for numerical measuring of their IFN resistances.

In this paper we describe a mathematical model and algorithm for calculation of a numerical measure of the differential IFN resistance of quasispecies using sequence data obtained from deep sequencing. The calculated value is referred to as IFN-resistance coefficient. The intra-host HCV populations was analyzed in specimens collected at 9 time-points during the first 48 hours after the initiation of IFN therapy in 7 patients with chronic hepatitis C. The state of HCV population was assessed by deep sequencing of amplicons containing hypervariable region 1 (HVR1) located at amino acid positions 384–410 in the structural protein E2. Sequence variation in HVR1 correlates with neutralization escape and is associated with viral persistence during chronic infection [6,8]. The IFN-resistance coefficients were calculated for each intra-host HCV variant. We show that analysis of IFN resistance correctly predicts therapeutic outcomes for the examined patients.

2. Methods

2.1. Samples

We studied 7 patients with chronic hepatitis C, for which samples have been collected at 9 time-points during the first 48 hours after the initiation of IFN therapy. Ethical Review and Informed Consent approval was granted by the institutional review boards at Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA and the US Centers for Disease Control and Prevention.

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), and eluted with 50 ul of buffer according to the manufacturer's instructions. PCR quantification was determined by COBAS AmpliPrep/COBAS Taq-Man HCV Test (Roche Diagnostics, Mannheim, Germany), genotype was determined using VERSANT HCV Genotype 2.0 Assay (LiPA) (Innogenetic NV, Gent, Belgium). RNA was precipitated and reverse-transcribed using both random and specific primers as previously described [12]. The junction E1/E2 region (309 nt), which contains the HVR1 region, was amplified using the nested PCR protocol described in [12].

The amplicons generated during the first-round PCR were used as templates for a nested PCR using hybrid primers composed of the 454 primer adaptors, multiple identifiers and specific sequences complementary to the HCV genome. This 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). PCR products were pooled and subjected to pyrosequencing using the GS FLX Titanium Series Amplicon kit in 454/Roche GS FLX instrument. Sequencing of the reverse strand was conducted using the amp primer B. The initial reads were processed by matching to the corresponding identifier. Low quality reads were removed using the GS Run Processor v2.3 (Roche, 2010). The 454 files were post-processed using two error correction and haplotypes reconstruction algorithms KEC and ET [13]. These algorithms were shown to be highly accurate in finding true haplotypes, removing false haplotypes and estimating the frequency of true haplotypes. The error-corrected files were aligned using Muscle [3] and clipped to 293 bp. For each patient results of the algorithm, that lead to the calculation of the interferon resistance coefficient with the minimum error, were selected.

2.2. Mathematical model and calculations

In this subsection we describe the method for calculating differential IFN resistance coefficients for

quasispecies using HCV sequence data. This method is based on the quasispecies model and therefore agrees with the standard population genetic models, considering that the later is mathematically a special case of the former [14].

The data under consideration are HCV titer and sets of HVR1 sequences with their relative frequencies for each time point $t' \in P = \{0, 1, 3, 6, 12, 18, 24, 36, 48\}$ hours obtained by deep sequencing and post-processed by the error correction algorithm. It is well-known, that sequencing and error correction are subject to errors, which may consist in the loss of true sequences and introduction of false sequences. Taking it into account, in order to obtain the most reliable results we analyzed only sequences present in every time point. The set of these sequences was denoted as A. However, the remaining sequences may represent true HCV variants. Exclusion of these variants from consideration may lead to error in calculations. This error should be estimated and considered in the model.

vThe following model with minor changes follows the standard quasispecies model [4,9]. If the population consists of $n+1$ different viral variants, the system of differential equations describing the evolution of the viral quasispecies abundances has the following form:

$$\dot{x}_i(t) = \sum_{j=1}^{n+1} q_{j,i} f_j(t) x_j(t) - d_i(t) x_i(t);\ i = 1, ..., n+1. \tag{1}$$

Here $x_i = x_i(t)$, $f_i = f_i(t)$ and $d_i = d_i(t)$ are the abundance, the replication rate and the death rate of the variant i, respectively; $q_{j,i}$ is the probability that replication of genome j produces genome i. $q_{j,i}$ does not depend on time, $q_j = \sum_{i=1}^{n+1} q_{j,i}$ (we assume that all other mutations are deleterious and do not influence the population dynamics). By the substitution

$$v_i = v_i(t) = \frac{x_i(t)}{\sum_{i=1}^{n+1} x_i(t)} \tag{2}$$

equation (1) can be transformed into the equivalent equation describing the evolution of frequencies of quasispecies:

$$\dot{v}_i = \sum_{j=1}^{n+1} q_{j,i} f_j v_j - d_i v_i - v_i \sum_{j=1}^{n+1} (q_j f_j - d_j) v_j, \quad i = 1, ..., n+1. \tag{3}$$

The inverse transformation is

$$x_i = x_i(t) = v_i(t) exp\left(\int_0^t \sum_{j=1}^{n+1} g_j(\tau) v_j(\tau) d\tau \right), \tag{4}$$

where $g_j(\tau) = q_j f_j(\tau) - d_j(\tau)$. This implies that if $X = X(t)$ is the absolute size of the viral population, then

$$\dot{X}(t) = \left(\sum_{i=1}^{n+1} g_i(t) v_i(t) \right) X(t). \tag{5}$$

The fitness functions $g_i(t)$ reflect changes of quasispecies fitness under the selection pressure.

Let us denote viral variants by $s_1, \ldots, s_{n+1}$. Let the first n variants be sequences from the set A with the $(n+1)$th variant representing the aggregate of all other sequences. For each $i = 1,\ n+1$ the frequencies $v_i(t')$ at time-points $t' \in P$ are given. Let $V(t') = \sum_{j=1}^{n} v(t')$, $u_i(t') = \frac{v_i(t')}{V(t')}$, $i = 1, \ldots,\ n$. $v_i(t)$, $u_i(t)$ and $V(t)$ can be approximated on the whole segment $[0, T]$, $T = 48$ by cubic splines. Using the spline approximation, the approximations of $\dot{v}_i(t)$, $\dot{u}_i(t)$ and $\dot{V}(t)$, $t \in [0, T]$, $i = 1, \ldots, n$ can be calculated.

As mentioned above, we assume that only variants $s_1, \ldots,\ s_n$ represent real quasispecies, while all other sequences can be sequencing artifacts. Therefore for each time point we normalize the frequencies of first n variants obtained by error correction and haplotypes reconstruction algorithm and consider the system of differential equations describing dynamics of $u_i(t)$, $i = 1, \ldots, n$ as the basic model:

$$\dot{u}_i = \sum_{j=1}^{n} q_{j,i} f_j u_j - d_i u_i - u_i \sum_{j=1}^{n} (q_j f_j - d_j) u_j, \tag{6}$$

For each $t \in [0, T]$ (6) can be considered as a system of linear equations with variables $f_i(t)$, $d_i(t)$ and used to find $g_i(t)$. This system does not have a full rank. In order to give the system a full rank we add an additional equation, which can be deduced from the titer. As above, the titer $h(t')$ is known at each time-point $t' \in P$, and $h(t)$ and $\dot{h}(t)$ were found on $[0, T]$ using the cubic splines approximation. Assuming that absolute size of the population $X(t)$ is proportional to the titer, i.e. $X(t) = \alpha h(t)$ and $\dot{X}(t) = \alpha \dot{h}(t)$ and taking into account (5) an additional equation could be written as

$$\frac{\dot{h}(t)}{h(t)} = \sum_{i=1}^{n} u_i(t)(q_i f_i(t) - d_i(t)). \tag{7}$$

In this work we use the simplified case of (6). Since the number of different HCV quasispecies found by

deep sequencing and present in each time point $t' \in P$ is small and the probability of mutation between each two of them is at most 0.0032%, mutations should not have an effect on the dynamics of population under consideration during 48 hours of observation and can be ignored. Therefore, $q_{i,k} = 0$ can be substituted for $i \neq k$ and (6) can be rewritten as

$$\dot{u}_i(t) = g_i(t)u_i(t) - u_i(t)\sum_{j=1}^{n} g_j(t)u_j(t), \tag{8}$$

Taking into account (7) and the fact that according to the selection of qusispecies under consideration $u_i(t) \neq 0$ for every $t \in [0, T]$, $i = 1, \ldots, n$, we have

$$g_i(t) = \frac{\dot{u}_i(t)}{u_i(t)} + \frac{\dot{h}_i(t)}{h_i(t)}. \tag{9}$$

Suppose that t_0 is the time when IFN starts affecting the viral population. The average fitness of the variant i on the segment $[t_0, T]$ is

$$r_i = \frac{1}{T - t_0}\int_{t_0}^{T} g_i(t)dt. \tag{10}$$

This parameter is used as a measure of IFN resistance. We assume that the parameter t_0 is the time, from which the titer starts declining monotonically.

The value of r_i may inaccurately reflect the dynamics of quasispecies from the set A, since, as mentioned above, the unknown proportion of true sequences may be excluded from consideration. In order to evaluate the possible error, note that the titer could reflect the size of the aggregate population of all $n + 1$ variants $s_1, \ldots, s_{n+1}$. Denoting the fitness functions and average fitness coefficients of variants $s_1, \ldots, s_n$ obtained in that case by $g_i'(t)$ and r_i', respectively, the basic model can be presented as the system

$$\dot{v}_i(t) = g_i'(t)v_i(t) - v_i(t)\sum_{j=1}^{n+1} g_j'(t)v_j(t), i = 1, ..., n + 1. \tag{11}$$

Equation (7) is replaced by the following equation:

$$\frac{\dot{h}(t)}{h(t)} = \sum_{i=1}^{n+1} v_i(t)g_i'(t) \tag{12}$$

and the formula for $g_i'(t)$ could be written as

$$g_i'(t) = \frac{\dot{u}_i(t)}{u_i(t)} + \frac{\dot{h}_i(t)}{h_i(t)} + \frac{\dot{V}(t)}{V(t)}. \tag{13}$$

or

$$g_i'(t) = g_i(t) + \frac{\dot{V}(t)}{V(t)}. \tag{14}$$

By integrating both sides of (14) we have

$$|r_i' - r_i| = \frac{1}{T - t_0}\left|\int_{t_0}^{T} \frac{\dot{V}(t)}{V(t)}dt\right| = \varepsilon. \tag{15}$$

So, ε could be used as the upper bound for the error. Note, that $\dot{x}_i(t) = g_i(t)x_i(t)$ and therefore

$$x_i(t) = x_i(0)e^{\int_0^t g_{i(t)dt}}. \tag{16}$$

Therapy is considered to be successful if after 12 weeks = 2016 hours of treatment the viral RNA concentration declines ≥ 100 times from the initial concentration [10,11]. Using this criteria, the threshold fitness coefficient r^* can be defined as the maximal constant fitness associated with the 100-fold reduction of concentration after 2016 hours. According to (16) $r^* = -0.0023$. Using fitness coefficients r_i, all HCV quasispecies were divided into 3 categories:

1. $r_i \leq r^*$ – strongly IFN-sensitive variants;
2. $r^* < r_i < 0$ – weakly IFN-sensitive variants;
3. $r_i \geq 0$ – IFN-resistant variants.

For convenience we use values $100 * r_i$ and $100 * r^* = -0.23$ and refer them as IFN-resistance coefficients and threshold IFN-resistant coefficient, respectively.

All calculations were made in MATLAB R2010b (The MathWorks, Inc).

3. Results and discussions

The model described in Methods was used to analyze HCV quasispecies from 7 patients with different outcomes of IFN therapy. Patients are divided into 4 groups according to the clinical criterion of treatment success [10,11]:

1. NR – nonresponder ($< 2log_{10}$ decrease of HCV RNA after 12 weeks of therapy);
2. RVR – rapid virological responder (undetectable HCV RNA by week 4 of therapy);
3. cEVR – complete early virological responder (undetectable HCV RNA by week 12 of therapy);
4. pEVR – partial early virological responder ($> 2log_{10}$ decrease of HCV RNA but still detectable HCV RNA after 12 weeks of therapy).

If patient is NR or pEVR, then the IFN therapy is considered unsuccessful.

The clinical data for patients are summarized in Table 1 in the column "Virologic Response".

3.1. Analysis of differential IFN resistance of quasispecies and model predictions

First of all, for each patient the average resistance coefficient of quasispecies $100\sum_{i=1}^{n} r_i v_i$ was calculated (here v_i is the average frequency of the variant i over time). The results are summarized in Table 1 in the column "Average resistance coefficient". The table shows that the average resistance coefficients of patients correlate with their virological responses.

The quasispecies populations of each patient were analyzed using the model.

1. Non-responders.

   Patient 1. The most frequent variant is on the border between being weakly and strongly IFN sensitive. Variants 2 and 3 are IFN resistant with the aggregate frequency of 45.53%. The variant 3 is IFN resistant even taking into account the possible error. Thus, presence of IFN-resistant variants of high frequency explains failure of therapy to clear the virus. [see Table 2]

   Patient 2. The main HCV variant is weakly IFN sensitive, suggesting that IFN was stalling replication or caused slow decline of this variant. However, the HCV population contains a highly IFN-resistant variant with $r_3 = 0.5939$ of high frequency, 11.60%. This variant has a potential to become dominant over time. This observations explains the negative outcome of therapy. [see Table 3]

2. Partial early virological responder

   Patient 3. Two most frequent variants with the aggregate frequency 65.17% are weakly and strongly IFN sensitive, respectively; while variant 3 is IFN resistant. The model suggests that the size of HCV population should initially decline due to the extinction of the major variants. However, the rate of decline should be moderate because of the growth in frequency of variant 3. Thus, the model explains negative outcome of therapy despite the initial decline in the titer, which corresponds to pEVR. [see Table 4]

3. Complete early virological responders

   Patient 4. The dominant variant is weakly IFN sensitive. All other variants are IFN sensitive. There is no IFN resistant variants. The model explains that the size of viral population should slowly decline, which is consistent with cEVR. [see Table 5]

   Patient 5. Variants 1 and 3 are strongly IFN sensitive. The highly frequent variant 2 is weakly IFN sensitive. If only first 3 variants are taken into account the

Table 1
Patients data

Patient	Virologic response	Average resistance coefficient
1	NR	−0.0773
2	NR	−0.1107
3	pEVR	−0.1744
4	cEVR	−0.1980
5	cEVR	−0.3543
6	RVR	−0.4694
7	RVR	−0.5073

Table 2
Patient 1

Sequence	Average frequency, %	$100 * r_i$
1	49.15	−0.2410
2	32.95	0.1038
3	12.58	0.1999
4	3.18	−0.5491
5	1.47	−0.1478
6	0.66	0.2144
error		0.1302

Table 3
Patient 2

Sequence	Average frequency, %	$100 * r_i$
1	58.36	−0.1222
2	14.45	−0.5404
3	11.60	0.5939
4	7.90	0.0627
5	7.68	−0.4578
error		0.0629

Table 4
Patient 3

Sequence	Average frequency, %	$100 * r_i$
1	34.72	−0.2102
2	30.45	−0.3393
3	20.35	0.1115
4	10.77	−0.0672
5	2.56	−0.2392
6	1.15	−0.6490
error		0.0188

Table 5
Patient 4

Sequence	Average frequency, %	$100 * r_i$
1	69.17	–0.0402
2	14.84	–0.9531
3	9.44	–0.0360
4	6.55	–0.3879
error		0.0016

Table 6
Patient 5

Sequence	Average frequency, %	$100 * r_i$
1	42.00	–0.6240
2	37.83	–0.0565
3	13.89	–0.7841
4	2.80	0.3569
5	2.54	1.1165
6	0.93	–0.0274
error		0.0373

conclusion is the same as for Patient 4. Variants 4 and 5 have a positive resistance coefficients. However, accuracy of the resistance coefficient is strongly dependent on the estimates of quasispecies frequencies. The coefficient should be most sensitive to variation in frequency of low-prevalence variants, for which the estimate error may be on the same scale as the frequency itself. However, considering that treatment with INF is not the only factor capable of restricting viral replication, low frequency may be an indication of the overall reduced fitness for a variant. Thus, such low-frequency variants may have a limited contribution to the INF resistance of the entire intra-host HCV population, making estimation of their frequency less relevant for accurate explanation of the therapy outcomes. This consideration is consistent with the observation of possible low impact of the minority INF-resistant variant on the therapy outcome in Patient 5. Alternatively, this observation may reflect overrating of resistance for this variant owing to the inaccurate frequency measurement, indicating importance of the measurement accuracy. Application of advanced sequencing-error correction and frequency-estimation algorithms should significantly improve evaluation of INF resistance. [see Table 6]

4. Rapid virological responders.

   All variants in Patient 6 are strongly IFN sensitive. Except for the low-frequency, 1.20%, variant 4, which is weakly IFN sensitive, all other variants

Table 7
Patient 6

Sequence	Average frequency, %	$100 * r_i$
1	58.01	–0.4052
2	38.25	–0.5908
3	2.54	–0.2738
4	1.20	–0.12
error		0.0195

Table 8
Patient 7

Sequence	Average frequency, %	$100 * r_i$
1	58.53	–0.3234
2	14.62	–0.5481
3	8.21	–0.3038
4	8.10	–0.7106
5	3.41	–2.3263
6	2.84	–0.3720
7	1.98	–1.0353
8	1.28	–1.0677
9	1.22	–1.2127
10	1.22	–1.5234
11	0.60	–0.7568
error		0.2314

in Patient 7 are strongly IFN sensitive. Considering the potential problem with accurate estimating frequencies for rare variants, all variants in these 2 patients remain IFN sensitive. The model explains the fast therapy success. [see Table 7]

4. Conclusion

Mathematical models developed in this study strongly suggest that individual intra-host HCV variants are differentially resistant to IFN. A novel coefficient of IFN resistance calculated in these models varies broadly among intra-host HCV variant in patients. The presence of HCV variants with the coefficient value > 0 is a strong indication of IFN resistance. The developed models allow for the evaluation of the HCV IFN resistance within the first 48 hours from initiation of treatment. The data suggest a novel framework for prediction of outcomes of IFN therapy.

The developed in this study quantitative measure of IFN resistance for individual intra-host HCV variants opens a novel opportunity for devising computational models for prediction of the therapy outcome. Success of the recently developed Bayesian networks models associating IFN resistance to the genetic

structure of small HCV genomic regions [7] suggests that similar models based on the INF resistance coefficient proposed here can be developed to predict the outcome of therapy from temporal frequency patterns of intra-host HCV. Additionally, the data obtained in this study can be used for modeling the association between genetic variations of the relevant HCV genomic regions such as HVR1 with the quantitative measure of IFN resistance for individual intra-host HCV variants, thus potentially improving accuracy of the previous models based on consensus sequences [7].

All calculations conducted here used molecular data sampled from 9 close time-points. Collection of serum samples and sequencing at such frequency may be unsuitable for the routine patient management. However, contribution of different time-points to estimation of the resistance coefficient is variable, with some points being seemingly dispensable, which warrants further investigation.

References

[1] M. Alter, "Epidemiology of hepatitis C virus infection", World J Gastroenterol, vol. 13, no. 17, 2007, pp. 2436–2441.

[2] E. Domingo, "Biological significance of viral quasispecies", Viral Hepatitis Rev. 2, 1996, pp. 247–261.

[3] R. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput", Nucleic Acids Res., 32(5), 2004, pp. 1792–1797.

[4] M. Eigen, J. McCaskill and P. Schuster, "The molecular quasi-species", Advances in Chemical Physics 75, 1989, pp. 149–263.

[5] J. Heathcote and J. Main, "Treatment of hepatitis C", J. Viral Hepat, vol. 12, no. 3, 2005, pp. 223–235.

[6] M. Isaguliants, "Hepatitis C virus clearance: the enigma of failure despite an impeccable survival strategy", Curr Pharm Biotechnol, vol. 4, no. 3, 2003, pp. 169–183.

[7] J. Lara, G. Xia, M. Purdy, Y. Khudyakov. "Coevolution of the hepatitis C virus polyprotein sites in patients on combined pegylated interferon and ribavirin therapy", J. Virol., 85(7), 2011, pp. 3649–3663.

[8] F. Lopez-Labrador, et al., "Genetic variability of hepatitis C virus NS3 protein in human leukocyte antigen-A2 liver transplant recipients with recurrent hepatitis C", Liver Transpl, vol. 10, no. 2, 2004, pp. 217–227.

[9] M. Novak, "Evolutionary dynamics", Belknap Press of Harvard University Press, 2006.

[10] E. Paulon, N.V. Naoumov, "Individualization of antiviral treatment regimens for chronic hepatitis C", European Journal of Gastroenterology and Hepatology, vol. 18, N. 4, 2006, pp. 321–326.

[11] J.M. Pawlotsky, "Molecular diagnosis of viral hepatitis", Gastroenterology, 122(6), 2002, pp. 1554–1568.

[12] S. Ramachandran, G.L. Xia, L.M. Ganova-Raeva, O.V. Nainan, Y. Khudyakov. "End point limiting dilution real time PCR assay for evaluation of HCV quasispecies in serum", J. Virol Meth, 151(2), 2008, pp. 217–224.

[13] P. Skums, et al., "Efficient error correction for deep sequencing of viral amplicons", Proc. 7th International Symposium on Bioinformatics Research and Applications (ISBRA 2011), Changsha, China, May 27–29, 2011, pp. 133–137.

[14] C. Wilke, "Quasispecies theory in the context of population genetics", BMC Evolutionary Biology, 5:44, 2005.