Table 2a. Variations in each Amino Acid Position and SVR rate

| Position | | Group 1 (n = 43) | *P* value | Group 2 (n = 35) | *P* value | Combined (n = 78) | *P* value |
|---|---|---|---|---|---|---|---|
| Core aa 110 | T | 100% (19 / 19) | 0.01 | 92.9% (13 / 14) | 0.004 | 97% (32 / 33) | 5E - 05 |
| | non T | 70.8% (17 / 24) | | 42.9% (9 / 21) | | 57.8% (26 / 45) | |
| p7 aa 773 | V | 77.4% (24 / 31) | 0.16 | 53.6% (15 / 28) | 0.03 | 66.1% (39 / 59) | 0.002 |
| | non V | 100% (12 /12) | | 100% (7 / 7) | | 100% (19 / 19) | |
| NS5A aa 2099 | R | 92.9% (13 / 14) | 0.40 | 91.7% (11 / 12) | 0.01 | 92.3% (24 / 26) | 0.01 |
| | non R | 79.3% (23 / 29) | | 47.8% (11 / 23) | | 65.4% (34 / 52) | |
| NS5B aa 3013 | L | 78.9% (26 / 33) | 0.17 | 47.8% (11 / 23) | 0.01 | 66.1% (37 / 56) | 0.008 |
| | non L | 100% (10 / 10) | | 91.7% (11 / 12) | | 95.5% (21 / 22) | |

Table 2b. Number of Amino Acid Substitutions in each Region and SVR rate

| Region | | Group 1 (n = 43) | *P* value | Group 2 (n = 35) | *P* value | Combined (n = 78) | *P* value |
|---|---|---|---|---|---|---|---|
| E2 aa 400-403 | mutation ≥2 | 89.3% (25 / 28) | 0.22 | 100% (11 / 11) | 0.002 | 92.3% (36 / 39) | 0.0005 |
| | mutation 0-1 | 73.3% (11 / 15) | | 45.8% (11 / 24) | | 56.4% (22 / 39) | |
| E2 aa 724-743 | mutation ≥1 | 100% (28 / 28) | 0.0002 | 72% (18 / 25) | 0.12 | 86.8% (46 / 53) | 0.0006 |
| | no mutation | 53.3% (8 / 15) | | 40% (4 / 10) | | 48% (12 / 25) | |
| ISDR(aa 2213-2248) | mutation ≥2 | 100% (15 / 15) | 0.08 | 86.7% (13 / 15) | 0.02 | 93.3% (28 / 30) | 0.003 |
| | mutation 0-1 | 75% (21 / 28) | | 45% (9 / 20) | | 62.5% (30 / 48) | |
| NS5A aa 2258-2306 | mutation ≥5 | 100% (19 / 19) | 0.01 | 84.2% (16 / 19) | 0.006 | 92.1% (35 / 38) | 0.0006 |
| | mutation 0-4 | 70.8% (17 / 24) | | 37.5% (6 / 16) | | 57.5% (23 / 40) | |

Table 3. Multivariate Logistic Regression Analysis

| Factor | Odds (95% CI) | *P* value |
|---|---|---|
| Age | 1.01 (0.91-1.13) | 0.85 |
| HCV RNA | 1.00 (1.00-1.00) | 0.09 |
| Fibrosis score ≥3/0-2 | 2.37 (0.21-26.7) | 0.48 |
| RVR achievement | 3.46 (0.54-22.1) | 0.19 |
| Ribavirin dose ≥ 80% | 16.0 (1.66-153) | 0.02 |
| Core aa 110 T | 24.7 (1.72-353) | 0.02 |
| NS5A aa 2258-2306 mutations 0-4/≥5 | 11.5 (1.23-108) | 0.03 |

Table 4a. Baseline Characteristics of Patients with NS5A aa 2258-2306 mutations 0-4 or ≥5 (Group 1 and 2)

| Characteristic | Mutation 0-4 (n = 40) | Mutation ≥5 (n = 38) | *P* value |
|---|---|---|---|
| Gender (Male/Female) | 22 / 18 | 16 / 22 | NS† |
| Age (yrs) | 54.3 ± 11.4* | 53.5 ± 11.5 | NS‡ |
| ALT (IU/l) | 73.8 ± 70.3 | 85.3 ± 78.7 | NS‡ |
| Platelet ($x10^4/mm^3$) | 18.0 ± 5.9 | 21.0 ± 5.7 | 0.03‡ |
| Fibrosis score (0-2 / ≥3)§ | 33 / 5 | 33 / 2 | NS† |
| HCV RNA (KIU/ml) | 1100 (99 - 30000) ** | 380 (12 - 5000) | 0.02‖ |
| IFN dose (≥80% / 60-80%)¶ | 31 / 8 | 33 / 2 | NS† |
| Ribavirin dose (≥80% / 60-80%)¶ | 25 / 14 | 28 / 7 | NS† |
| RVR rate (%) | 65.8 | 62.9 | NS† |
| EVR rate (%) | 94.7 | 100 | NS† |
| Relapse rate (%) | 35.9 | 7.9 | 0.002† |
| SVR rate (%) | 57.5 | 92.1 | 0.0006† |

*: mean ± SD, †: Fisher's exact probability test , ‡: Student t test, §: Mutation 0-4 : n = 38, mutation ≥5 : n = 35,

**: median (range), ‖: Mann-Whitney's U test, ¶: Mutation 0-4 : n = 39, mutation ≥5 : n = 35

Table 4b. Baseline Characteristics of Patients with Core 110 T or N/S (Group 1 and 2)

| Characteristic | Core 110 T (n = 33) | Core 110 N/S (n = 45) | *P* value |
|---|---|---|---|
| Gender (Male/Female) | 18 / 15 | 20 / 25 | NS† |
| Age (yrs) | 50.4 ± 13.0* | 56.4 ± 9.5 | 0.032‡ |
| ALT (IU/l) | 64.5 ± 48.2 | 88.8 ± 86.2 | NS‡ |
| Platelet ($x10^4/mm^3$) | 19.3 ± 4.9 | 19.5 ± 6.6 | NS‡ |
| Fibrosis score (0-2 / ≥3)§ | 30 / 1 | 36 / 6 | NS† |
| HCV RNA (KIU/ml) | 580 (54 - 3600) ** | 980 (12 - 30000) | NS‖ |
| IFN dose (≥80% / 60-80%)¶ | 26 / 3 | 38 / 7 | NS† |
| Ribavirin dose (≥80% / 60-80%)¶ | 23 / 6 | 30 / 15 | NS† |
| RVR rate (%) | 72.4 | 59.1 | NS† |
| EVR rate (%) | 100 | 95.5 | NS† |
| Relapse rate (%) | 3.0 | 38.6 | 9E-05† |
| SVR rate (%) | 97.0 | 57.8 | 5E-05† |

*: mean ± SD, †: Fisher's exact probability test , ‡: Student t test, §: Core 110 T : n = 31, Core 110 N/S : n = 42,

**: median (range), ‖: Mann-Whitney's U test, ¶: Core 110 T : n = 29

**Figure 1**

Click here to download high resolution image



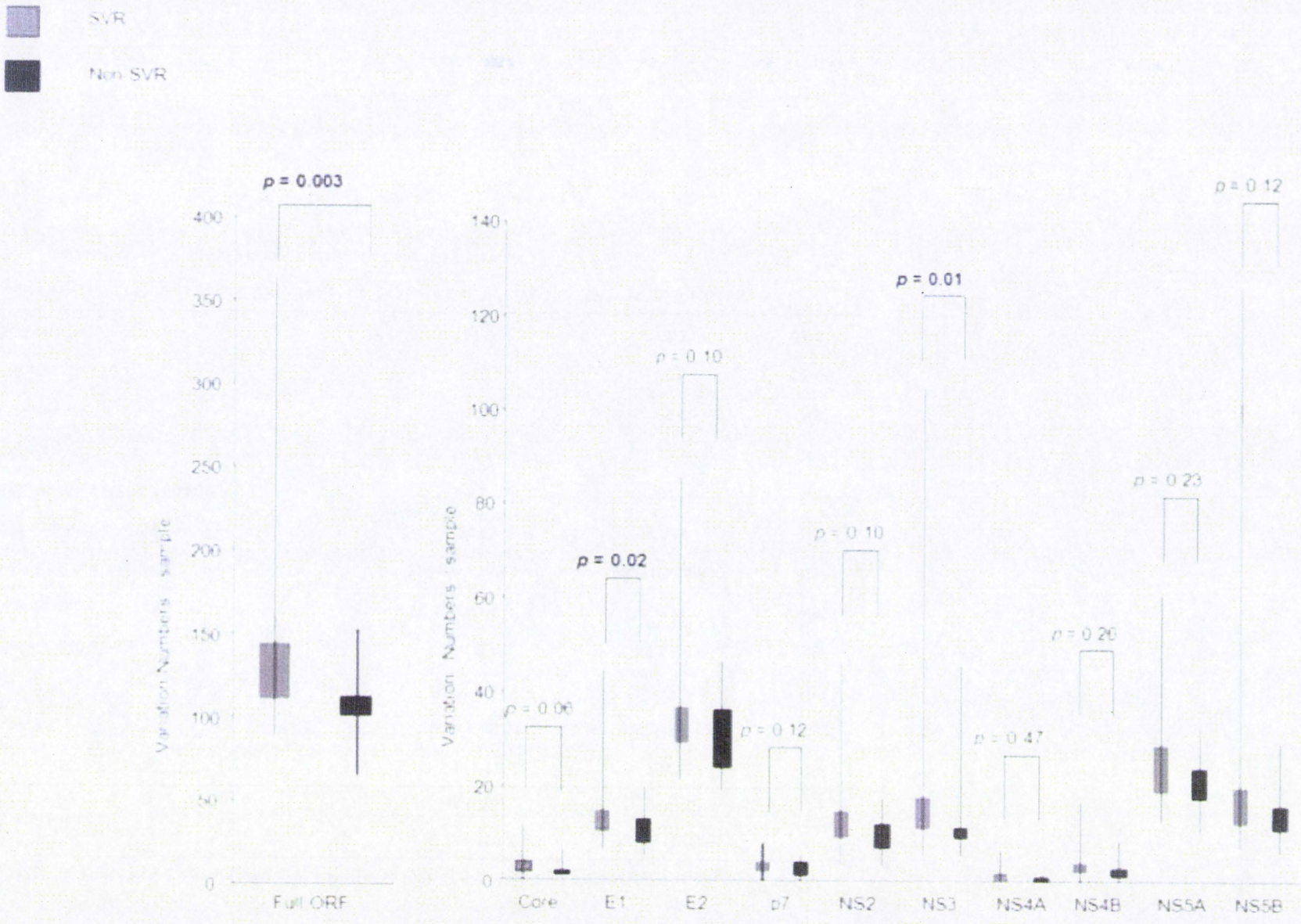


**Figure 2a**
Click here to download high resolution image



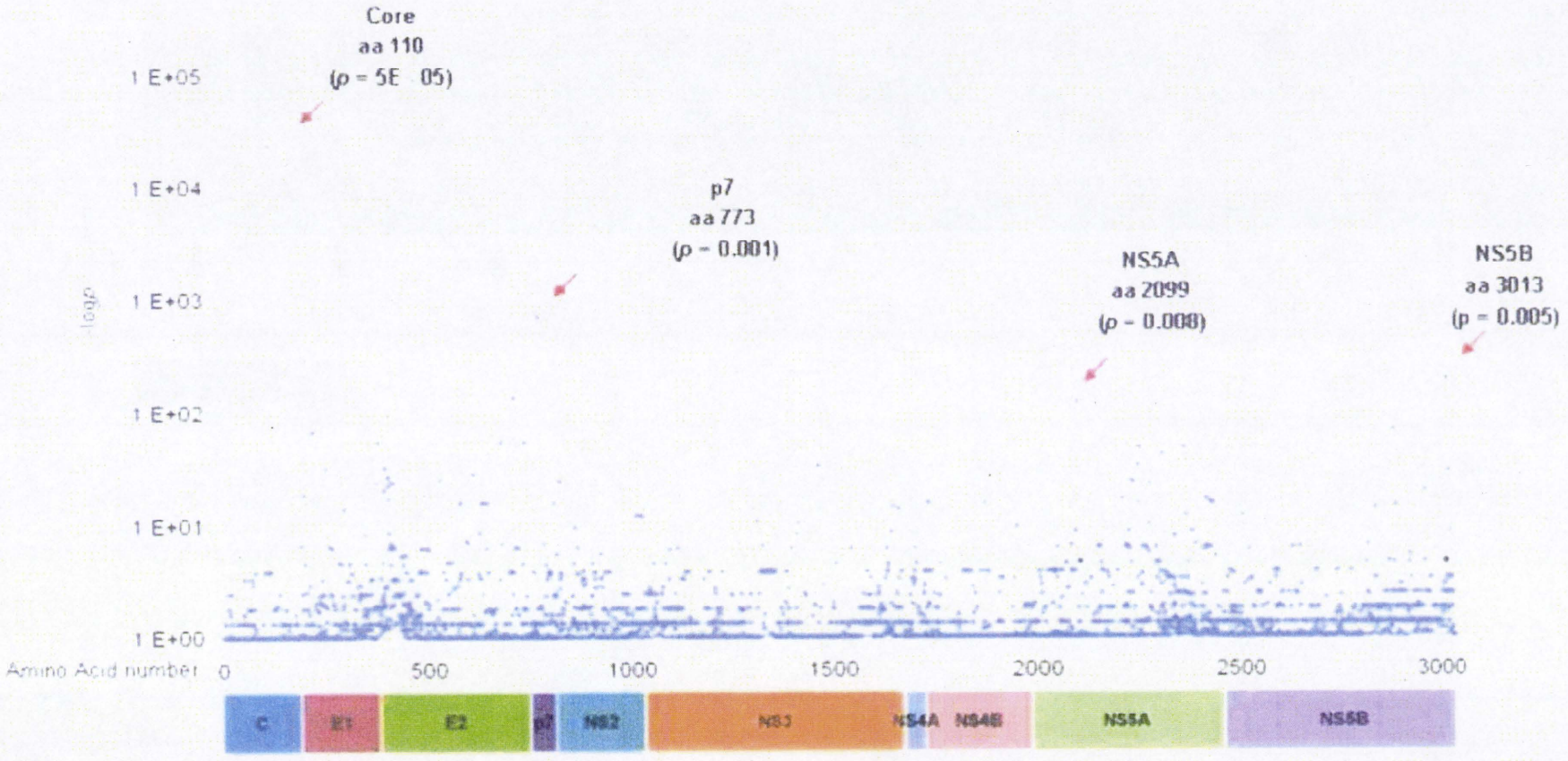


**Figure 2c**
Click here to download high resolution image



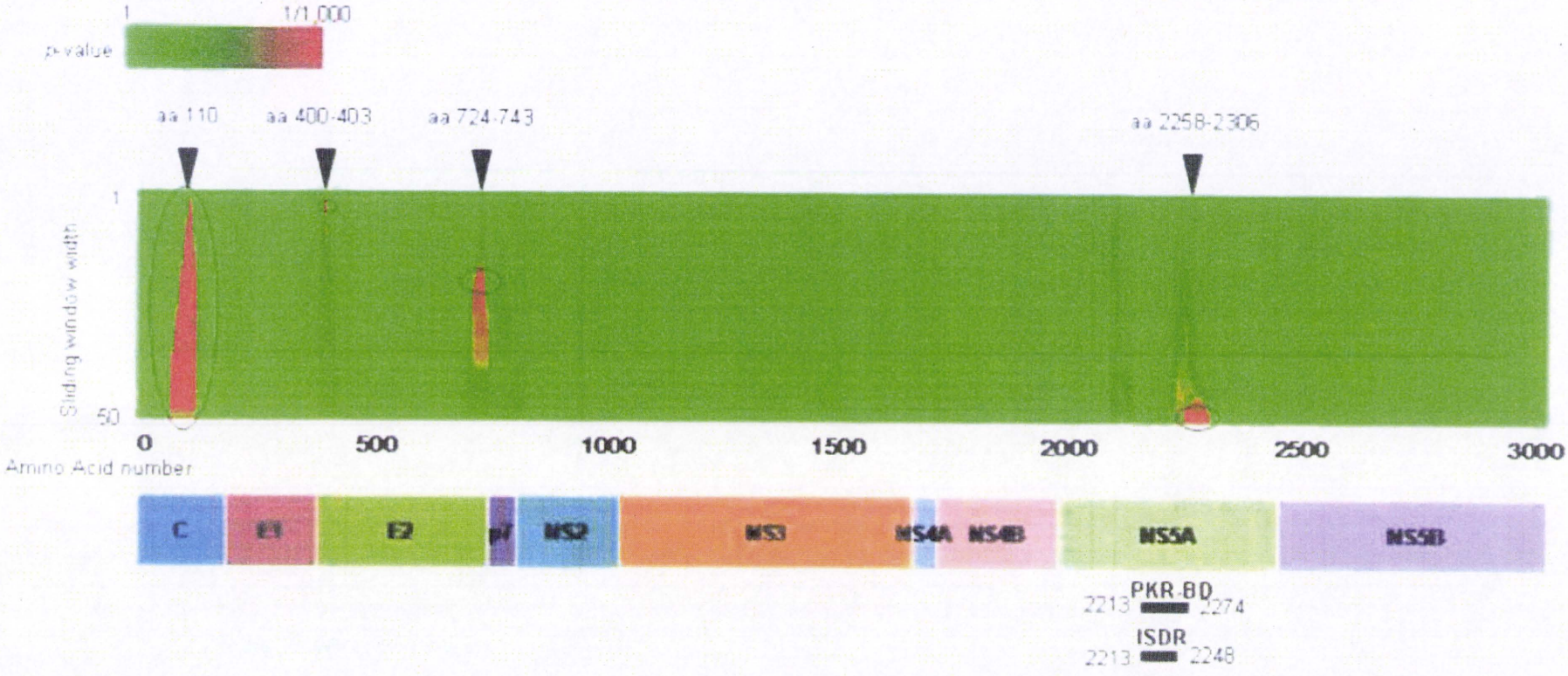


**Figure 2d**

Click here to download high resolution image

**Figure 2b**
Click here to download high resolution image

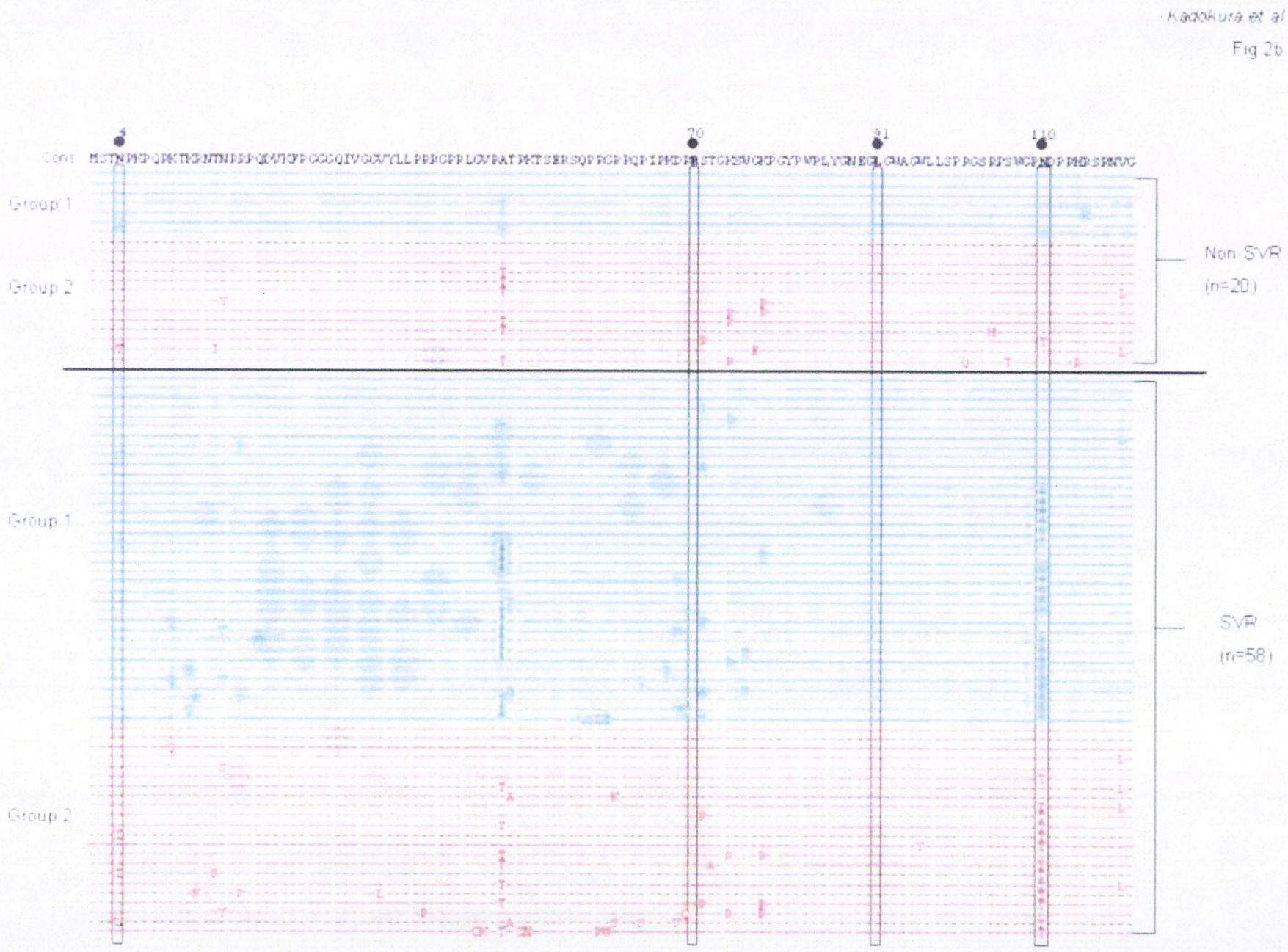
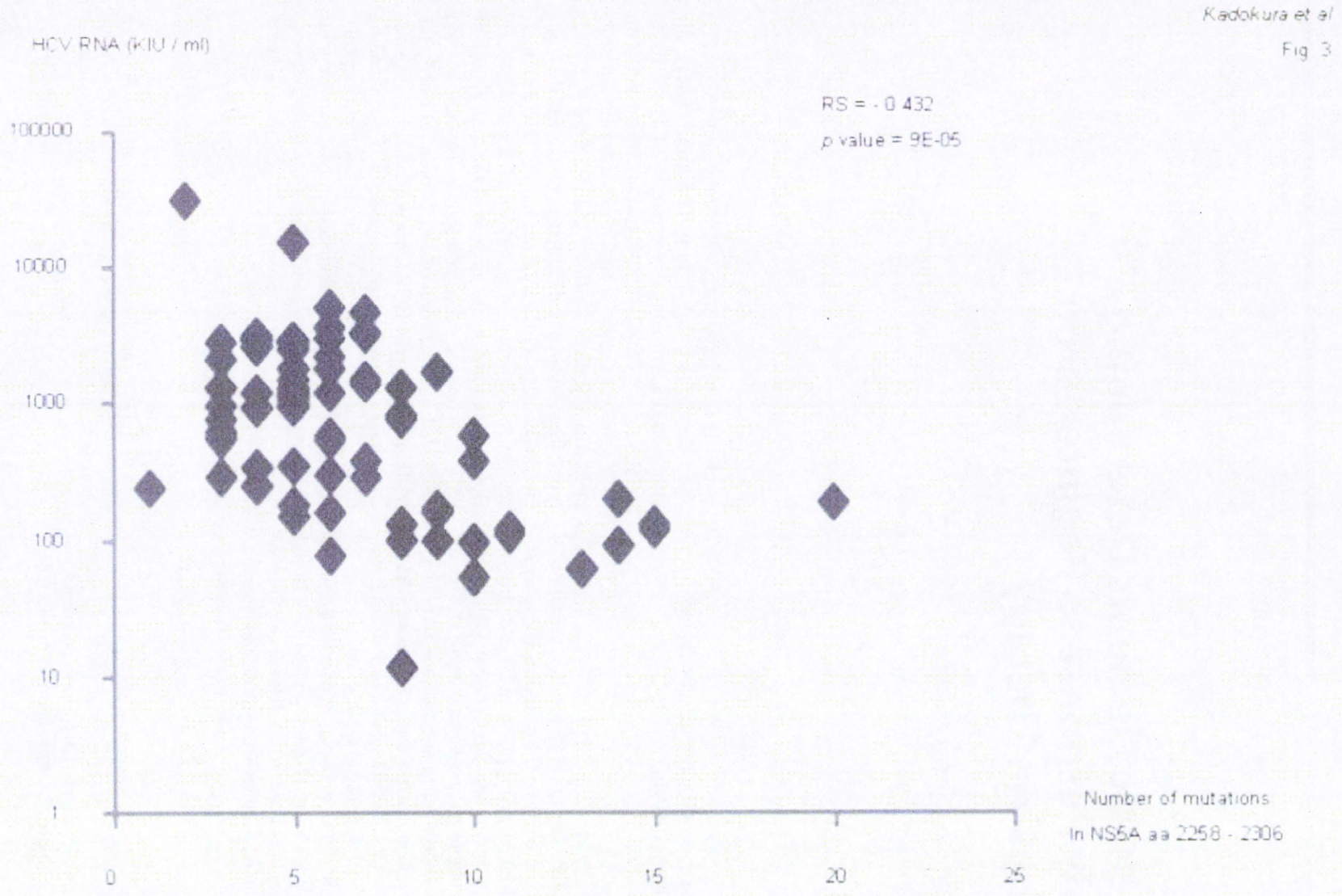


**Figure 3**
Click here to download high resolution image



Original Article

Intervirology

Intervirology 2010;53:66–69
DOI: 10.1159/000252787

Published online: January 5, 2010

# HCV Genetic Elements Determining the Early Response to Peginterferon and Ribavirin Therapy

Nobuyuki Enomoto Shinya Maekawa

First Department of Medicine, Faculty of Medicine, University of Yamanashi, Chuo, Japan

**Key Words**
Full open reading frame analysis · Hepatitis C virus · Peginterferon/ribavirin therapy

**Abstract**
The aim of this study was to search hepatitis C virus (HCV) genetic elements determining the early response to peginterferon/ribavirin therapy using HCV genome-wide analysis. From a total of 88 chronic hepatitis C patients with HCV-1b treated with peginterferon/ribavirin, the whole HCV amino acid sequence was determined and analyzed according to the viral response during the treatment. Mutations in NS5A-ISDR (interferon sensitivity-determining region) are associated with rapid viral response at week 4, and the core arginine70glutamine (R70Q) mutation is associated with no early viral response at week 12, revealing that core 70 and NS5A are the most important factors determining the virological kinetics during peginterferon and ribavirin therapy.



## Introduction

Hepatitis C virus (HCV) is a major cause of chronic liver diseases, and worldwide 170 million people are infected with HCV. With the introduction of the recent combination therapy of pegylated-interferon (PEG-IFN) and ribavirin (RBV), half of patients can eradicate the virus (sustained virological response, SVR). The SVR rate of HCV to the PEG-IFN/RBV therapy is dependent on HCV genotypes, and the viral kinetics during the treatment strongly affect the final viral clearance [1, 2]. It is generally considered that HCV structures affect the treatment response since the SVR rate to PEG-IFN/RBV therapy depends upon viral genotypes as described above. However, comprehensive analysis of the contribution of HCV structures to different responses has not yet been conducted. In the present study, in order to clarify the relationship between HCV sequences and viral responses, we have determined the complete HCV open reading frame sequences obtained from pretreatment patients' serum, and investigated their response by searching for HCV genetic elements determining the early response to PEG-IFN/RBV therapy using HCV genome-wide analysis.

## Methods

A total of 88 chronic hepatitis C patients with HCV-1b treated with PEG-IFN/RBV were studied. From pretreatment sera, the whole HCV deduced amino acid sequence (3,010 amino acids) was determined in each patient by direct RT-PCR.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com


0300–5526/10/0531–0066$26.00/0

Accessible online at:
www.karger.com/int

Nobuyuki Enomoto, MD
First Department of Medicine, Faculty of Medicine
University of Yamanashi
1110 Shimokato, Chuo, Yamanashi 409-3898 (Japan)
Tel. +81 55 273 9584, Fax +81 55 273 6748, E-Mail enomoto@yamanashi.ac.jp

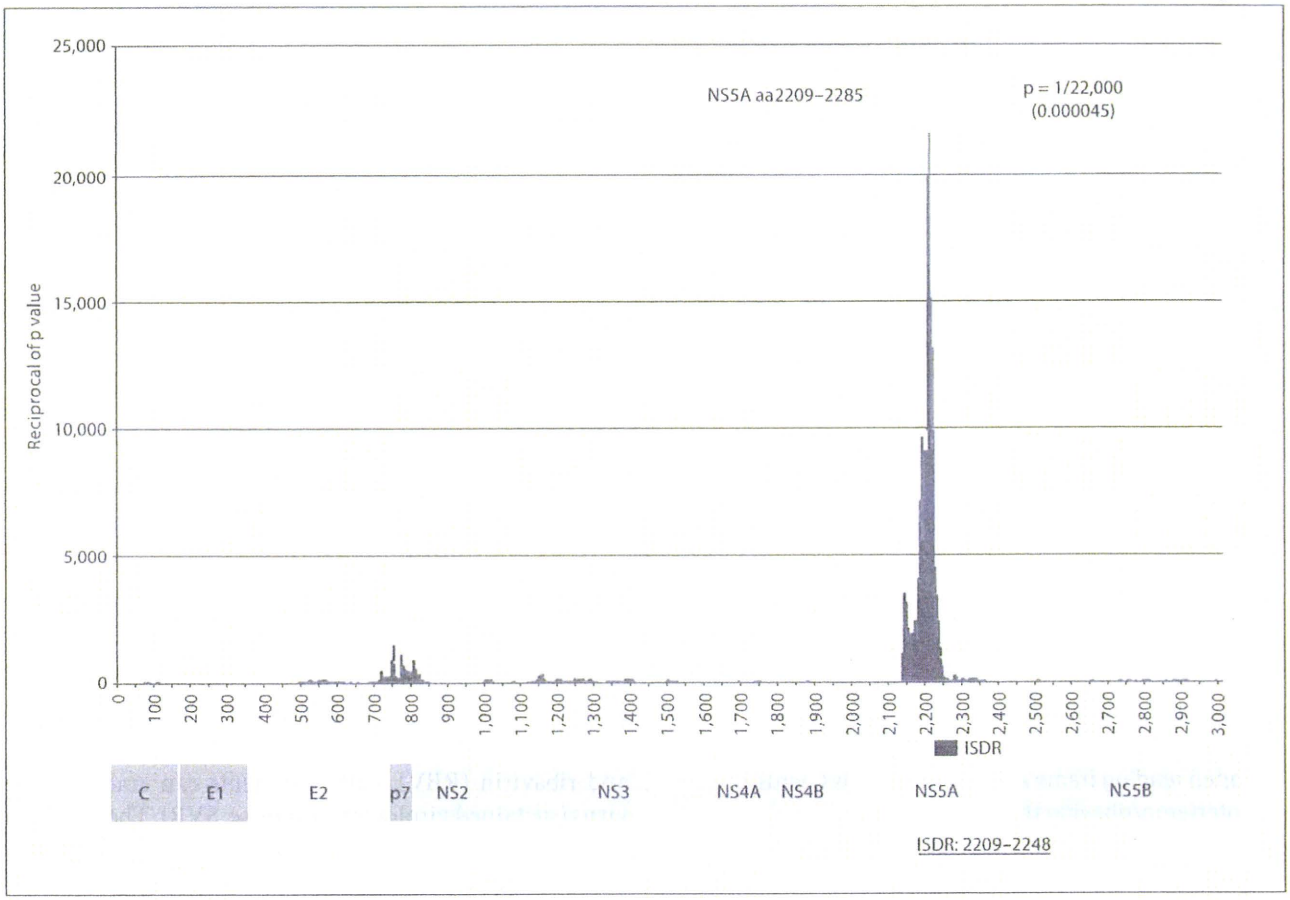



**Fig. 1.** Reciprocal of p value for sliding window analysis with 77 amino acid width for RVR versus others.

Amino acid usage of each of the 3,010 positions was compared according to the different virological response in order to identify the single amino acid differences determining the virological response. In addition, sliding window analyses were carried out in order to identify the amino acid region associated with the virological response. The number of the amino acid changes in the fixed stretch of the sequence (window: 2–100 amino acids) were compared according to the virological response, scanning the whole HCV amino acid sequence by sliding this window one by one.

## Results

Of 88 patients studied, 9 showed rapid viral response (RVR; HCV-RNA undetectable at week 4) and 71 showed early viral response (EVR; over 2-log drop of HCV-RNA at week 12). The other 17 patients showed no EVR, indicating these patients are highly resistant to the treatment.

Mutations in the region overlapping NS5A-ISDR (interferon sensitivity-determining region, aa2209–2248) are associated with the good response to PEG-IFN/RBV therapy as shown in sliding window analysis comparing RVR patients at week 4 and others (fig. 1). In contrast, the core R(arginine)70Q (glutamine) mutation is associated with a poor response resulting in no EVR at week 12 by single amino difference analysis comparing non-EVR patients and the others (fig. 2).

## Discussion

In the present study, using a sliding window analysis comparing all HCV amino acids, the amino acid region located in ISDR was extracted as the most significant region discriminating the RVR and non-RVR patients. By

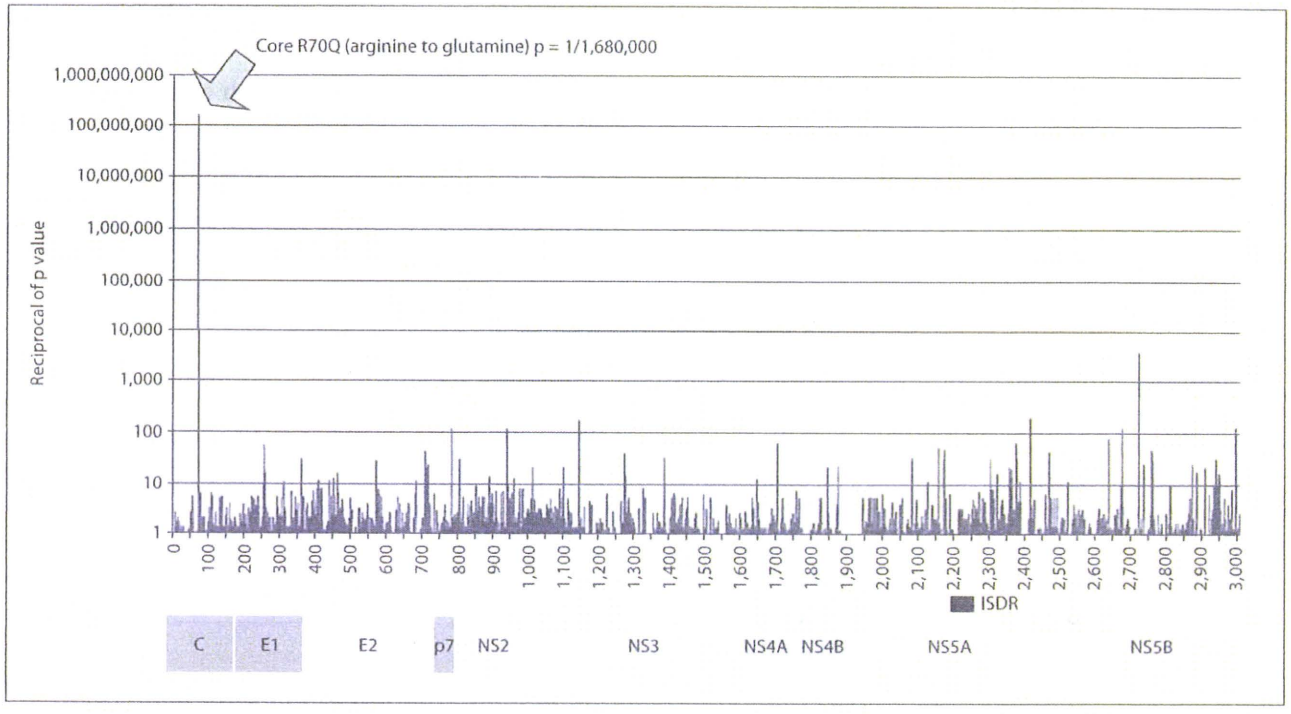



**Fig. 2.** Reciprocal of p value for single amino acid difference along the whole HCV sequence for non-EVR versus others.

comparing amino acids between the non-EVR patients and the others, remarkable differences were clustered in a single amino acid polymorphism in the core 70. Recent studies have proven that the initial viral response at week 4 and week 12 of the PEG-IFN/RBV therapy could be a useful predictor of the final outcome, indicating that the present findings are important for predicting treatment outcome and individualizing the treatment regimen for each patient as well as understanding the mechanism of diverse response to PEG-IFN/RBV therapy.

ISDR was first identified as the region significantly related to SVR in the era of IFN monotherapy in Japanese patients [3, 4]. 'Mutant type', meaning 4 or more mutations in the region, was associated with high SVR rate, while the rate was low in the 'intermediate type' (1–3 mutations) and wild type (no mutation). Though there were controversies as to the predictive value of ISDR, since studies in Europe and in North America did not necessarily reproduce evident correlation between ISDR and SVR, a recent meta-analysis proved its value by demonstrating a clear relationship all over the world, even in Western countries [5]. The present study reproduced the significance of ISDR in PEG-IFN/RBV therapy. Mutations in ISDR make HCV highly sensitive to IFN, leading to RVR. Current guidelines indicate that RVR patients with low viral load before treatment can be treated with 24 weeks instead of the standard 48 weeks of therapy. Since most ISDR mutant patients show low viral loads, these easy-to-treat patients in genotype 1b should be mainly infected with HCV with ISDR mutations, suggesting ISDR genotyping would identify the patients treatable with the abbreviated regimen.

On the other hand, in the present study, the polymorphism of core 70 was extracted as the most significant position to determine poor virological response in 12 weeks (non-EVR). The contribution of core region amino acid polymorphism in resistance to (PEG-)IFN/RBV therapy was previously reported by Akuta et al. [6], who first found that the polymorphisms in a combination of core 70 and 91 were closely related to the final outcome. The importance of core 70 polymorphism alone, however, was considered rather weak in their study for its smaller p value. Their end point was the final outcome of the treatment, which could be influenced by a variety of factors other than viral genetics, such as host factors (age, sex, fibrosis, body weight, etc.) and treatment (dose of

PEG-IFN/RBV). Further studies are needed to clarify the significance of the core mutations for final outcome of the treatment in the context of the HCV genome-wide analysis.

Different viral responses by polymorphisms in core 70 were also recently suggested in North American patients by Donlin et al. [7]. However, it was reported that the association with core 70 was weaker in their study. Very recently, the IL28B (interferon-lambda-3) gene polymorphism has been found to be closely associated with treatment response in patients in the United States, European Union and Japan by human genome-wide analysis [8–10]. The favorable IL28B genotype is found most frequently in Asian patients, second in European-Americans, and least in African-Americans, indicating that a well-known racial difference in treatment efficacy can be explained by the IL28B polymorphism. The interaction between viral and human genome polymorphisms should be studied further with regard to the treatment response.

## Conclusion

HCV genome-wide analysis with a large number of patients successfully revealed that core 70 and NS5A are the most important factors determining the virological kinetics during PEG-IFN/RBV therapy. Viral genome-wide analysis is a promising tool for elucidating the unknown viral factors for different pathological pictures, such as disease progression.

## Disclosure Statement

Supported in part by a Grant-in-Aid for Research on Hepatitis (grant ID H19-002) from the Ministry of Health, Labor, and Welfare, Japan. The authors report receiving grant support from Schering-Plough and Roche.

## References

1 Neumann AU, Pianko S, Zeuzem S, Yoshida EM, Benhamou Y, Mishan M, et al: Positive and negative prediction of sustained virologic response at weeks 2 and 4 of treatment with albinterferon alfa-2b or peginterferon alfa-2a in treatment-naive patients with genotype 1, chronic hepatitis C. J Hepatol 2009; 51:21–28.

2 Lee SS, Ferenci P: Optimizing outcomes in patients with hepatitis C virus genotype 1 or 4. Antivir Ther 2008;13(suppl 1):9–16.

3 Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al: Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b: sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J Clin Invest 1995;96:224–230.

4 Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al: Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 1996;334:77–81.

5 Pascu M, Martus P, Hohne M, Wiedenmann B, Hopf U, Schreier E, et al: Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. Gut 2004;53:1345–1351.

6 Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al: Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 2005;48:372–380.

7 Donlin MJ, Cannon NA, Yao E, Li J, Wahed A, Taylor MW, et al: Pretreatment sequence diversity differences in the full-length hepatitis C virus open reading frame correlate with early response to therapy. J Virol 2007; 81:8211–8224.

8 Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al: Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461:399–401.

9 Suppiah V, Moldovan M, Ahlenstiel G: *IL28B* is associated with response to chronic hepatitis C interferon-α and ribavirin therapy. Nat Genet 2009;41:1100–1104.

10 Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al: Genome-wide association of *IL28B* with response to pegylated interferon-α and ribavirin therapy for chronic hepatitis C. Nat Genet 2009; 41:1105–1109.

BioSystems 99 (2010) 70–78

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

# Reproducibility and usability of chronic virus infection model using agent-based simulation; comparing with a mathematical model

Jun Itakura[a,*], Masayuki Kurosaki[a], Yoshie Itakura[a], Sinya Maekawa[b], Yasuhiro Asahina[a], Namiki Izumi[a], Nobuyuki Enomoto[b]

[a] Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan
[b] First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, 1110, Shimogatou, Chuou-shi, Yamanashi 409-3898, Japan

ARTICLE INFO

*Article history:*
Received 30 June 2009
Received in revised form 27 August 2009
Accepted 6 September 2009

*Keywords:*
Agent-based model
Virus infectious disease

ABSTRACT

We created agent-based models that visually simulate conditions of chronic viral infections using two software. The results from two models were consistent, when they have same parameters during the actual simulation. The simulation results comprise a transient phase and an equilibrium phase, and unlike the mathematical model, virus count transit smoothly to the equilibrium phase without overshooting which correlates with actual biology in vivo of certain viruses. We investigated the effects caused by varying all the parameters included in concept; increasing virus lifespan, uninfected cell lifespan, uninfected cell regeneration rate, virus production count from infected cells, and infection rate had positive effects to the virus count during the equilibrium period, whereas increasing the latent period, the lifespan-shortening ratio for infected cells, and the cell cycle speed had negative effects. Virus count at the start did not influence the equilibrium conditions, but it influenced the infection development rate. The space size had no intrinsic effect on the equilibrium period, but virus count maximized when the virus moving speed was twice the space size. These agent-based simulation models reproducibly provide a visual representation of the disease, and enable a simulation that encompasses parameters those are difficult to account for in a mathematical model.



## 1. Introduction

All viruses need hosts as a basis for their life. When a virus enters the host body, it invades cells and uses both its own enzymes and those of the host cells to replicate. Host cells infected by viruses launch a self-defense system known as the innate immune system (See and Wark, 2008; Naniche, 2009), which inhibits viral replication and uses the human leukocyte antigen system and cytokines to elicit an immune response. Immune cells that have received signals from host cells activate other immune cells, neutralize viruses in the serum by means of antibodies, and prevent the virus from replicating and proliferating by destroying or curing host cells. Viral infection is a disorder based on the interactions between viruses and cells.

The power relationship between these agents changes along with the progression of the disease. In the very early stages of infection, as the host defense mechanisms are immature, the virus has the ability to overwhelm the host cells, actively replicate, and proliferate. Subsequently, as the capacity of the immune system improves, the speed of viral proliferation drops and the virus count reaches a peak. Infected host cells begin to be disrupted by the immune system or virus particles, and symptoms appear as a result. If the immune system is stronger than the virus, then the viral counts decline, and, in transient viral disorders, the virus is finally eliminated and the host recovers. In chronic viral disorders, however, the power relationship between the virus and host cells reaches equilibrium, and a long-term power balance is maintained with the virus count reaching a plateau.

Mathematical models have been proposed to study the dynamics of such viral disorders, and are regarded as being of value in understanding this phenomenon (Ho et al., 1995; Nowak et al., 1996; Neumann et al., 1998). However, these models are difficult to understand for clinicians, and their applicability is somewhat limited in everyday practice. In clinical research, measurements of viral dynamics in patients for short duration have been made for human

*Abbreviations:* HIV, human immunodeficiency virus; HBV, hepatitis B virus; HCV, hepatitis C virus.

* Corresponding author. Tel.: +81 422 32 3111; fax: +81 422 32 9551.
*E-mail address:* jitakura@musashino.jrc.or.jp (J. Itakura).

0303-2647/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2009.09.001

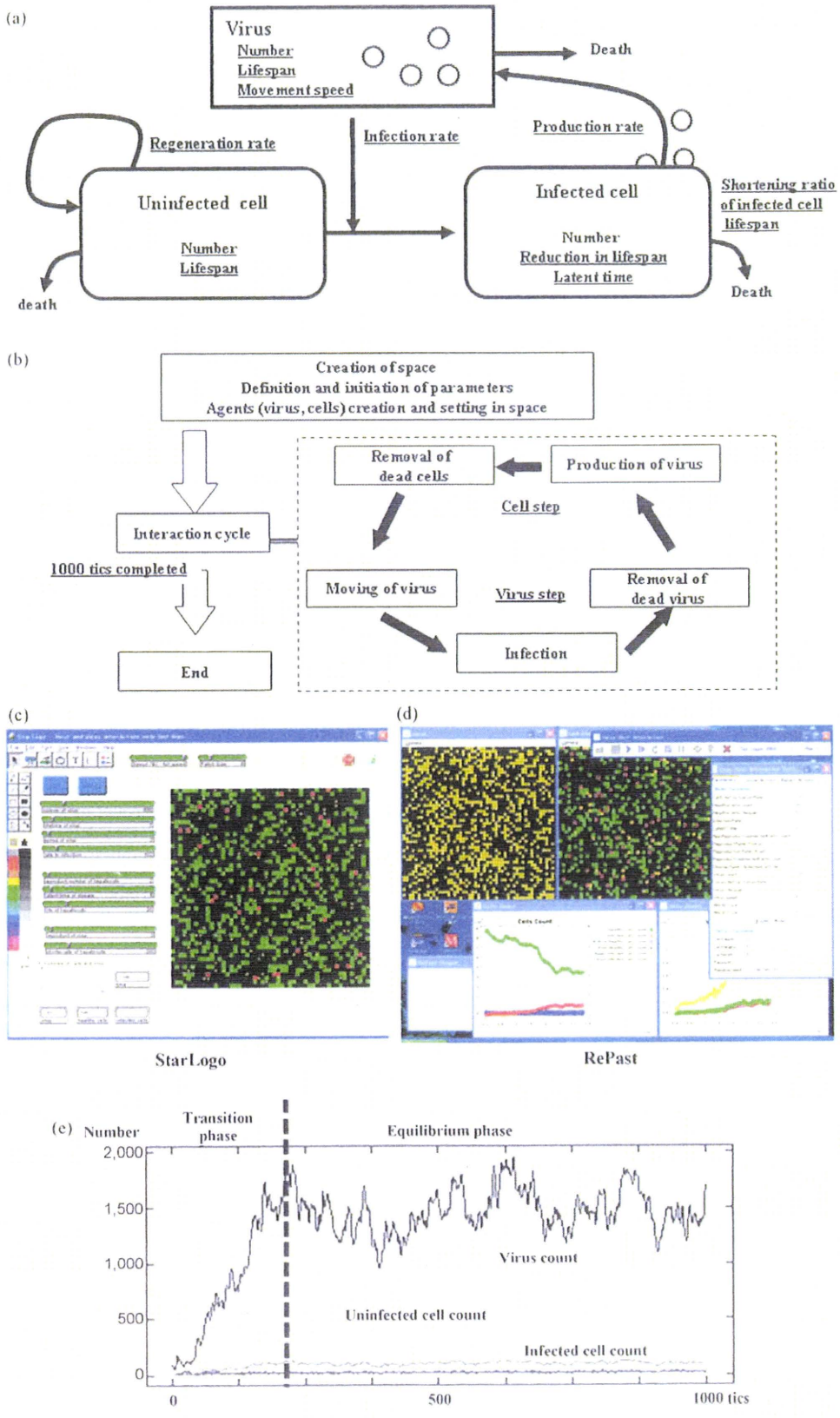



**Fig. 1.** Simulation design and an example of simulation results. (a) Model concept. Viruses, uninfected cells, and infected cells were treated as agents, and parameters were set for each of these and for interactions between agents (underlined). (b) Flowchart of the program. After preparing the simulation, we entered the interaction cycle, in which virus steps (such as movement) and cell steps were repeated. One cycle was counted as 1 tic, and the simulation concluded after 1000 tics. (c and d) Simulation screen using (c) StarLogo and (d) RePast. Yellow circles are viruses, green squares are uninfected cells, and orange and red indicate infected cells, with orange indicating the latent period. In StarLogo, all the agents are shown on the same screen, but in RePast, viruses and cells are shown in separate windows. (e) Example of a simulation chart in StarLogo. After the start of simulation the virus count and infected cell count increase while the uninfected cell count decreases, with equilibrium state reached after a certain number of tics.

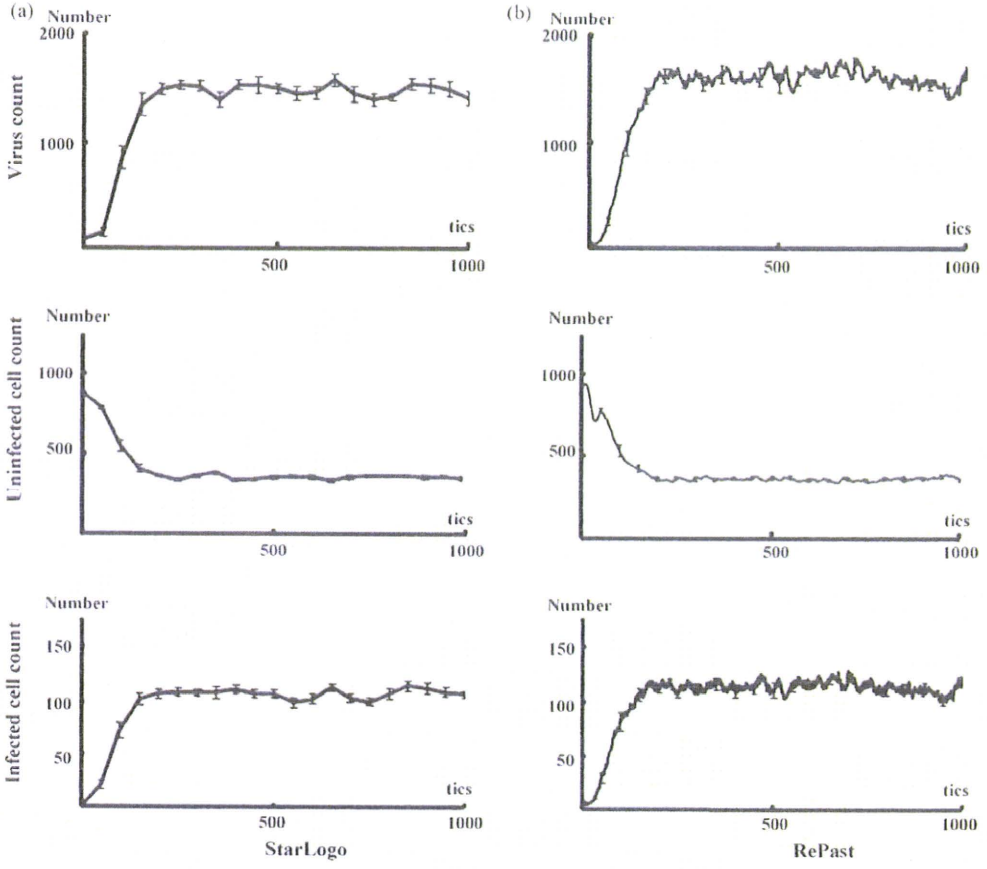



**Fig. 2.** Comparison of simulation results in (a) StarLogo and (b) RePast. The results were consistent when the parameters were made consistent. (Virus count [average ± SD]: StarLogo 1458.03 ± 173.1, RePast 1462.71 ± 178.8, $p = 0.94$. Uninfected cell count: 364.24 ± 30.4, 368.11 ± 33.4, $p = 0.83$. Infected cell count: 105.73 ± 13.0, 107.74 ± 13.0, $p = 0.24$. Unpaired Student's $t$-test.) Parameter values were set as follows: initial virus count, 100; uninfected cell count, 880; infected cell count, 0; virus speed of movement, 5 grids/tic; infection rate, 10%; uninfected cell regeneration rate, 1%; latent period, 3 tics; and virus reproduction rate, 5/cells/tic. The following parameter settings were taken from actual measurements: virus lifespan, 4.5 tics; uninfected cell lifespan, 49.8 tics; and infected cell lifespan, 6.7 tics.

immunodeficiency virus (HIV) (Ho et al., 1995), hepatitis B virus (HBV) (Nowak et al., 1996) and hepatitis C virus (HCV) (Neumann et al., 1998), and research is also underway on a range of models based on animal experiments and cell culture systems. As chronic viral disorders persist over long periods of time complete follow-up of viral dynamics is difficult. Furthermore, limitations of items that can be measured, such as the difficulty of measuring whole numbers of host cells, make it extremely difficult to investigate their consistency in mathematical models.

The recent ascend of dynamic-models owes much to advances in computers. Computer performance has improved markedly in recent years, not only in terms of their calculating capacity but also with regard to image displays, and models that offer a visual representation of viral disorders are now being reported (Gilbert and Bankes, 2002; Duca et al., 2007; Shapiro et al., 2008; Castiglione et al., 2007). One advantage of such visual models is that by providing a visual representation, they make understanding the disease status easy. Another benefit is that they enable parameters to be identified that are hidden as background noise in mathematical models. However, these models have some problems; it is difficult to prove the reproducibility of the simulation results derived from different languages or libraries, difficult to prove the validity of the model's concepts, and difficult to prove that the simulation results accurately reflect the reality. In this study, we created agent-based computer models that visually simulate the conditions of chronic viral infections using two software. The reproducibility of two agent-based computer models and the differences between agent-based models and the mathematical model were analyzed. This agent-based model enabled us to investigate how each parameter included in the concept affects the conditions of chronic viral infections.

## 2. Methods

### 2.1. Selection of Software

In this study, we used two different types of softwares: StarLogo version 2.0 (http://education.mit.edu/starlogo/) supplied by MIT Media Laboratory and Recursive Porous Agent Simulation Toolkit (RePast-3.0, http://repast.sourceforge.net/) supplied by the Argonne National Laboratory. StarLogo uses Logo, one of the simplest programming languages, and has a fixed graphical user interface. RePast is a library that uses Java, another programming language, which also has a fixed graphical user interface.

Logo is an assembly language, and StarLogo carries out processing sequentially. Java is an object-oriented language, and RePast has a faster processing speed than StarLogo. In addition, StarLogo has a number of stipulations to simplify simulations, such as parameters can only be set up to five decimal places and the simulation space is also fixed as 51 × 51 square grids. RePast, on the other hand, has fewer such restrictions. Thus, it offers a higher degree of freedom in program settings than StarLogo. Taking simulation space as an example, in spite of the restrictions imposed by the underlying operating system's image display system, any number of grids can be set and a hexagonal grid could also be chosen rather than a square one. However, users must stipulate and set all parameters themselves. This means that they must first declare the shape of the grid and the number of grids they will use to fill the simulation space. Java is also more difficult to learn than Logo, and debugging and correcting the program is also more difficult. Thus, it is difficult to judge whether or not the results agree with the planned simulation.

In effect, these two different types of softwares are polar opposites. It is simple to start a simulation in StarLogo, but producing results takes time and it is difficult to carry out more complex simulations. In RePast it is difficult to compose the program and judge whether or not the planned study has actually been achieved, but the

simulation itself takes only a short time to complete and there are lesser restrictions in the construction of a simulation model.

### 2.2. Concept for Modeling

We applied the basic virus–host interaction mathematical model to the agent-based simulation system with slight modifications. The mathematical model was used to describe the dynamics of HIV (Ho et al., 1995), HBV (Nowak et al., 1996), and HCV (Neumann et al., 1998) and the only agents involved were host cells and viruses, without the inclusion of immune cells. The effects of the immune system are expressed by varying parameters such as lifespan of host cells and viruses.

Fig. 1a illustrates the study concept. Viruses have the ability to infect healthy host cells (uninfected cells) and the infected cells produce new viruses. Both cells and viruses have definite lifespans, and the lifespan of infected cells is usually shorter than that of uninfected cells. Uninfected cells automatically regenerate within the space, whereas infected cells only arise due to infection of uninfected cells. Viruses also lack the ability to regenerate themselves and are only produced from infected cells.

### 2.3. Parameter Settings

In the present study, as the StarLogo settings are circumscribed, we limited the simulation space to 51 × 51 square grids. However, we made an exception here while investigating the effects of size of space on the simulation results. The numbers of viruses, uninfected cells, and infected cells could only be set before the start of the simulation. As described in the later, our simulation ran in cycles, with 1 cycle defined as 1 tic.

In mathematical simulation models, the death rate is required as a parameter. However, in our program we set lifespans for viruses and uninfected cells. These lifespans were not uniform, but were set to have a deviation of about 10%. The lifespan of cells was shortened by infection with ratio decided beforehand.

The infection ratio was meaningful only when an infected cell and a virus coincidentally occupied the same grid, and this was used to calculate the probability of the infection occurring in that situation. The virus production rate was set as the number of viruses produced by an infected cell during 1 tic. Infected cells could be set as a parameter indicating the latent period between the time of virus infection and the time of virus replication.

In order to emulate the tissue repair capacity, we set uninfected cell regeneration rate such that grids without any cells had a specified probability of producing uninfected cells on top of themselves. As a result, the more the cell count declined within a space the more regenerated uninfected cells were produced, whereas the number of regenerated cells declined as cell count increased.

The number of grids through which a virus could move in 1 tic was set as the speed of movement, and the direction of movement was set within a range of 90° toward the top of the simulation space. The program used a circulatory method of movement; when a virus arrived at the top of the space, it was translocated to the bottom, and moved again toward the top. Cells were fixed on the grid.

### 2.4. Simulation Flowchart

Fig. 1b shows a flowchart of the program. First, the simulation space was produced, after which each parameter was defined and the initial settings were made. Next the agents – viruses and uninfected and infected cells – were produced. The simulation cycle was as follows. Viruses moved to a new grid, and if an uninfected cell was present, this was infected with a probability based on the infection rate. The lifespan of the virus decreased, and viruses that had completed their lifespan and those that had caused an infection were removed from the space. Infected cells produced new viruses, the lifespans of both uninfected and infected cells decreased. Then, cells that had completed their lifespan were eliminated and a new cycle began. The program was set such that the simulation ended after this cycle had repeated 1000 times. This meant that one simulation was complete after 1000 tics.

### 2.5. Data Collection

The RePast model was programmed such that data for each tic was saved automatically as a text file at the end of the simulation. This text file could be opened by a database software. The StarLogo model was programmed to stop the simulation and collect data after every 50 tics.

### 2.6. Mathematical Model

In order to compare the results of this agent-based simulation, we used a viral infection mathematical model, which we improved as follows.

$$\frac{dT}{dt} = s[2601 - (T + I)] - dT - bVT \quad (1)$$

$$\frac{dI}{dt} = bVT - dI \quad (2)$$

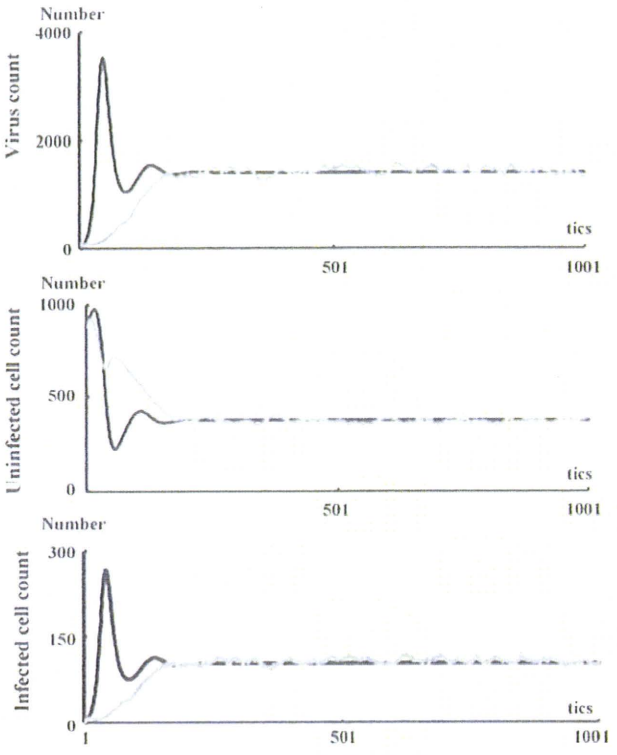



**Fig. 3.** Comparison of results of agent-based simulation and mathematical simulation. Both sets of results were consistent for the equilibrium phase, but differed in the shift in transition phase. Black line: mathematical model; grey line: results of simulation in RePast. Parameter values were set as follows: initial virus count, 100; uninfected cell count, 880; infected cell count, 0; virus speed of movement, 5 grids/tic; infection rate, 10%; uninfected cell regeneration rate, 1%; latent period, 3 tics; virus reproduction rate, 5/cells/tic; virus lifespan, 10 tics; uninfected cell lifespan, 50 tics; and cell lifespan-shortening ratio as a result of infection, 69%.

$$\frac{dV}{dt} = pI - cV \quad (3)$$

where, $T$ is the uninfected cell count, $I$ is the infected cell count, and $V$ is the virus count. Uninfected cells are supplied to the space with a probability $s[2601 - (T+I)]$, as the number of grids in this agent-based simulation model was 2601 (51 × 51). The death rate of uninfected cells is $d$, the death rate of infected cells is $\delta$, and the death rate of viruses is $c$. The infection rate is indicated by $\beta$. Viruses are released from infected cells at a probability $p$.

### 2.7. Statistical Analysis

Statistical analyses were performed by statistical tests using the program StatView 5.0 (SAS Institute Inc.). All tests of significance were two-tailed, with $p$ values of <0.05 considered to be significant.

## 3. Results

### 3.1. Reproducibility of Chronic Viral Infection Disease Models Using Agent-based Simulation Methods

We constructed the chronic viral infection model with StarLogo library. Fig. 1c shows the simulation screen, and Fig. 1e shows one sample result. Immediately after the start of the simulation, the virus count temporarily dropped in accordance with the onset of an infection. Subsequently, the virus count started to increase with an increase in the infected cells and a decrease in the uninfected cells. After a certain number of tics (around 300 in this example), although the virus count, infected cell count, and uninfected cell count had some fluctuation, an equilibrium state was reached. We use the following descriptive terms in this paper: the transient phase is the period during which virus growth peaks, and the equilibrium phase is the period during which an equilibrium state is

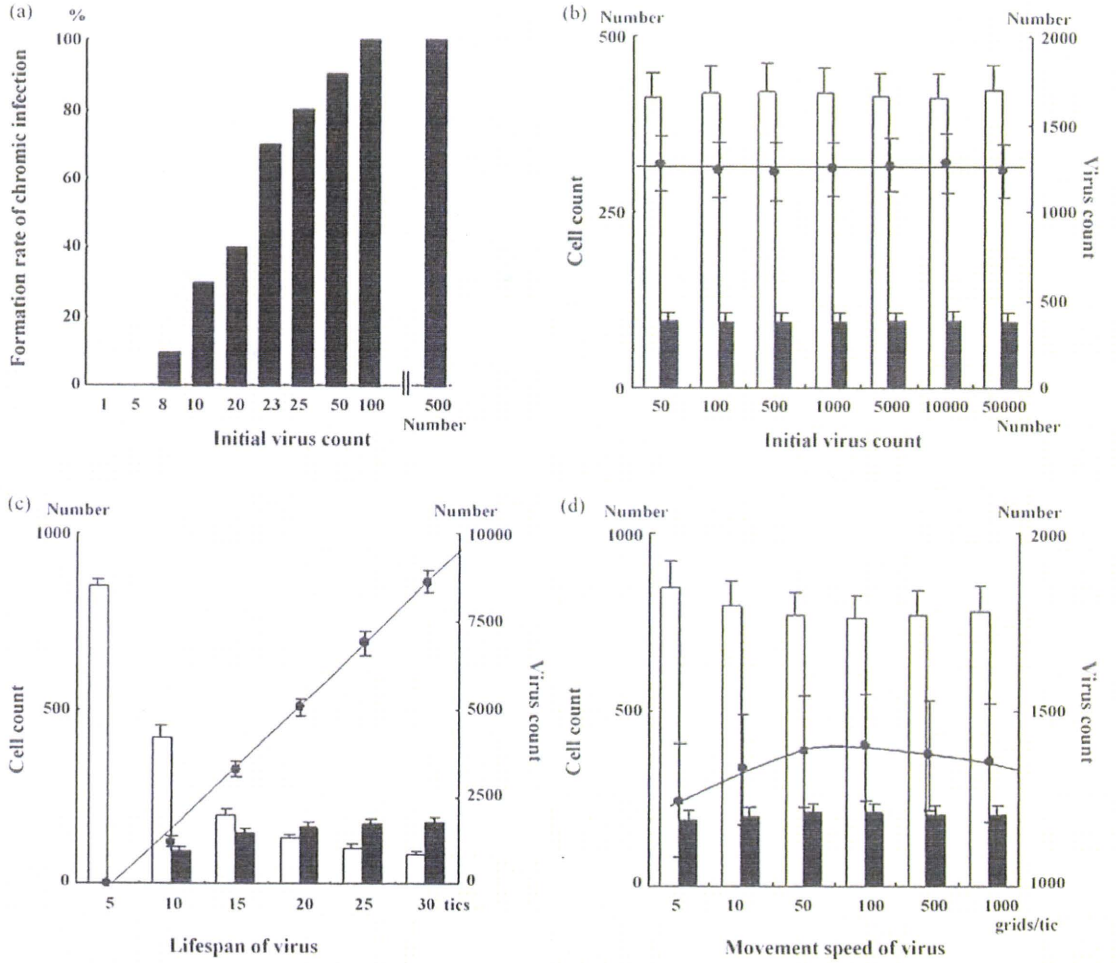



**Fig. 4.** Effects of changes in viral parameters. (a) The higher the initial virus count, the greater is the increase in the rate of formation of chronic infection, but (b) there was no effect on the conditions in the equilibrium phase. (c) Extending the virus lifespan increased the virus count. (d) Increasing the speed of virus movement to 100 grids/tic increased the virus count, but increasing it to 500 grids/tic had the opposite effect, with a slight declining trend. (a) Black bars: number of infections produced; (b–d) black circles: virus count; line: virus count approximation curve; white bars: uninfected cell count; black bars: infected cell count.

established. When the simulation was performed multiple times, the features described above were maintained, and the average values for virus, infected cell, and uninfected cell counts during the equilibrium state were all consistent.

Fig. 1d shows the simulation screen of the RePast. When we attempted setting all the initial parameters to the same values as those in the StarLogo, the results were not consistent. When we recalculated the parameters from the simulation results, in RePast, the parameters were largely maintained at the levels of the settings, but in StarLogo, the lifespans of both cell types became shorter than the settings while the simulation was in progress. We made the results of both simulations consistent by using the same parameters during the actual simulation (Fig. 2a and b).

### 3.2. Comparison Between Agent-based Simulation Models and Mathematical Simulation Model

We investigated whether the results of a chronic viral infection disease model produced by RePast would be consistent with the results of a mathematical model. For the mathematical model, we carried out an approximate integration using a four-dimensional Runge–Kutta method to ensure that the uninfected cell count and infected cell count would be in the same class. Parameters were always fixed as constant between simulations. The simulation results were consistent for the equilibrium phase, but transitions in virus count during the transient phase varied, with a shift to equilibrium state following two overshoots in the mathematical model, but a monotonic increase following a logistic curve in the agent-based model (Fig. 3). In the mathematical model, when the equilibrium condition was calculated with $dT/dt = dI/dt = dV/dt = 0$, the equilibrium-phase virus count, uninfected cell count, and infected cell count were very similar to those of the agent-based model (virus count: mathematical model 371.8/space, agent-based model 371.1 ± 32.4/space [average ± SD]; uninfected cell count: mathematical model 1605/space, agent-based model 1454 ± 194/space; infected cell count: mathematical model 115.9/space, agent-based model 108.3 ± 14.2/space).

### 3.3. Usability of the Models; Effect of Changing Parameters

We investigated the changes in the equilibrium phase brought about by changing each parameter. All the investigations below were carried out by using RePast, and we used the average values from ten simulations.

### 3.4. Viral Parameters

The lower the virus counts at the beginning of the simulation, the lower the probability of a chronic infection (Fig. 4a). However, the initial virus count did not have any effect on the equilibrium