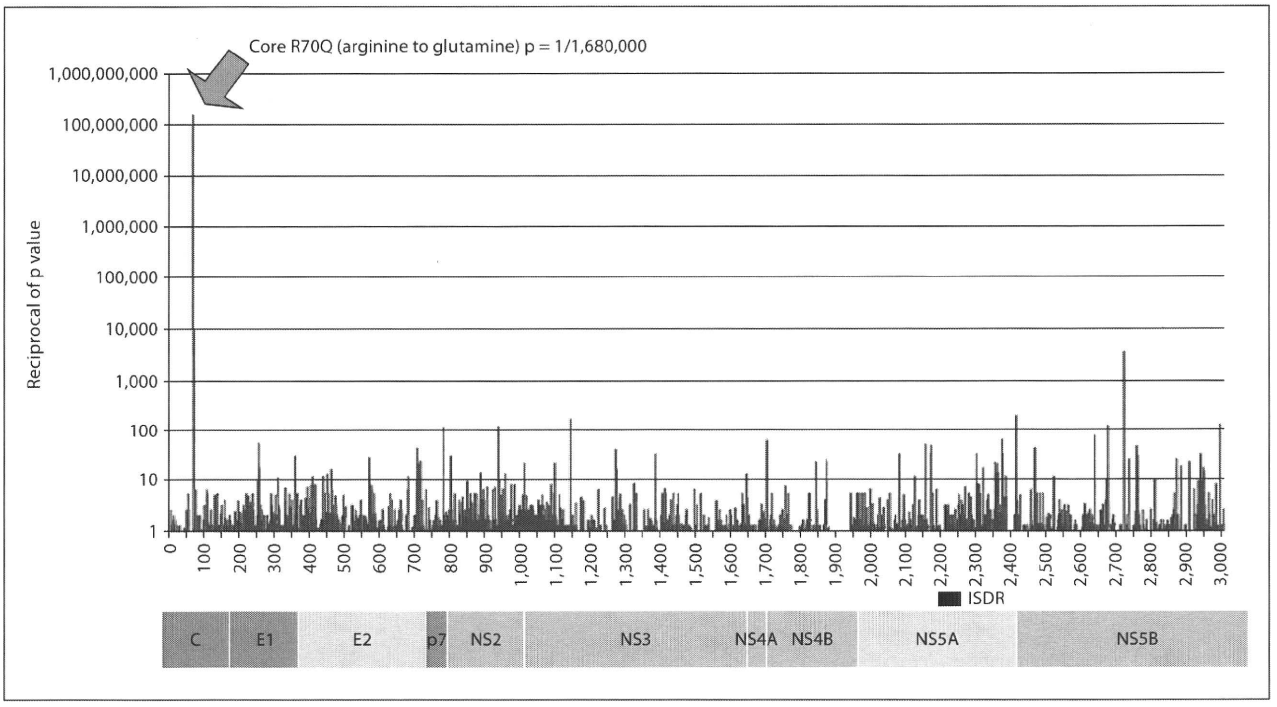

Fig. 2. Reciprocal of p value for single amino acid difference along the whole HCV sequence for non-EVR versus others.

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

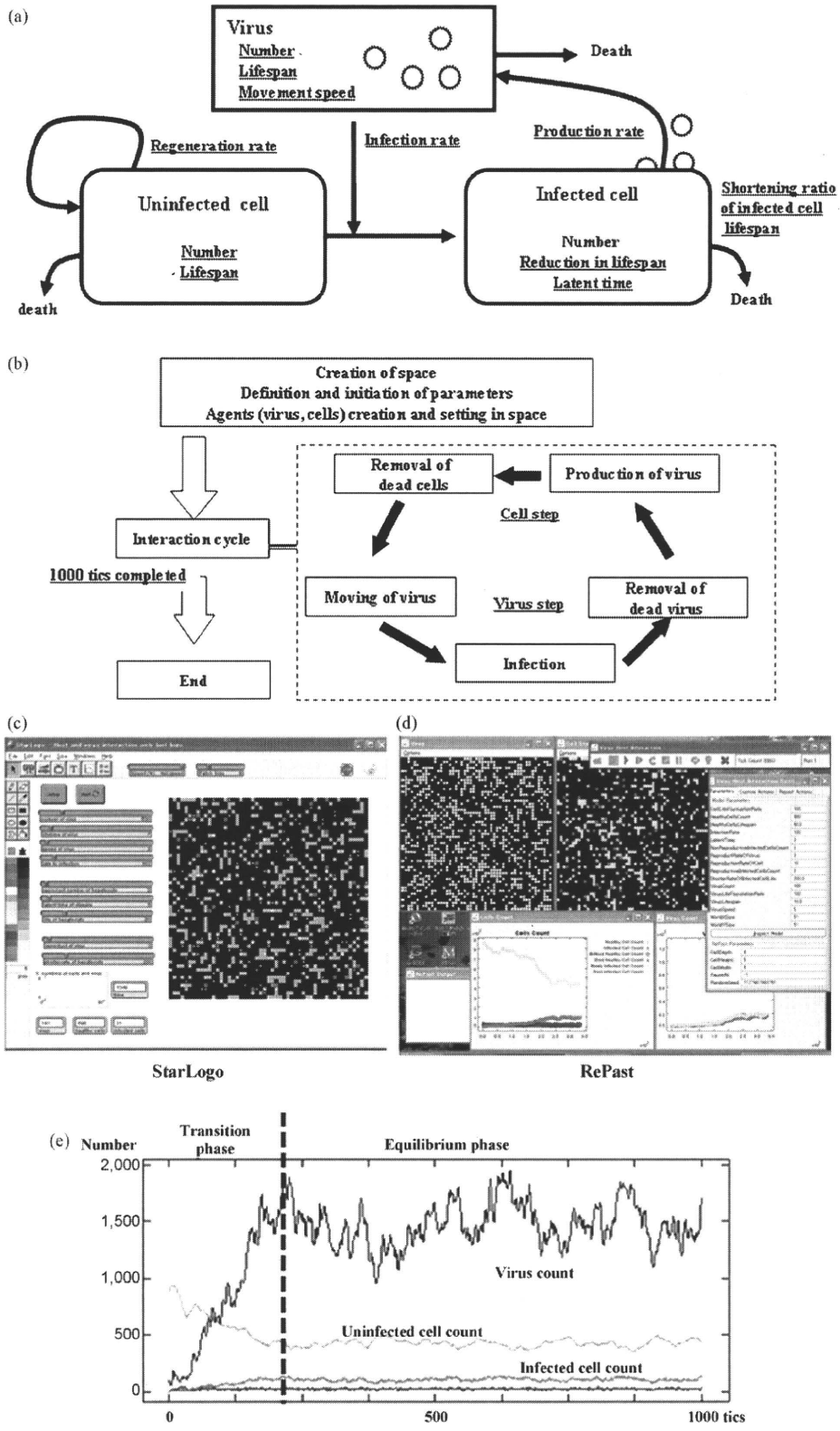

Fig. 1. Simulation design and an example of simulation results. (a) Model concept. Viruses, uninfected cells, and infected cells were treated as agents, and parameters were set for each of these and for interactions between agents (underlined). (b) Flowchart of the program. After preparing the simulation, 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.

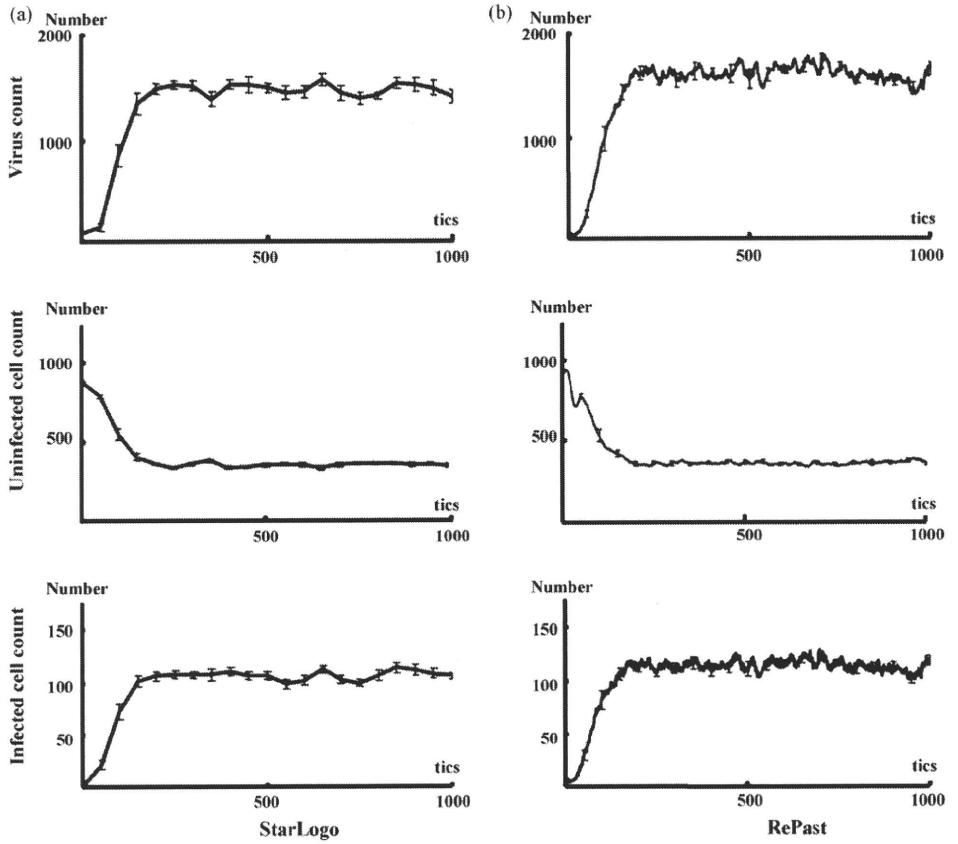

Fig. 2. Comparison of simulation results in (a) StarLogo and (b) RePast. The results were consistent when the parameters were made consistent. (Virus count [average ± SD]: StarLogo 1458.03 ± 173.1, RePast 1462.71 ± 178.8, $p = 0.94$. Uninfected cell count: 364.24 ± 30.4, 368.11 ± 33.4, $p = 0.83$. Infected 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.

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

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

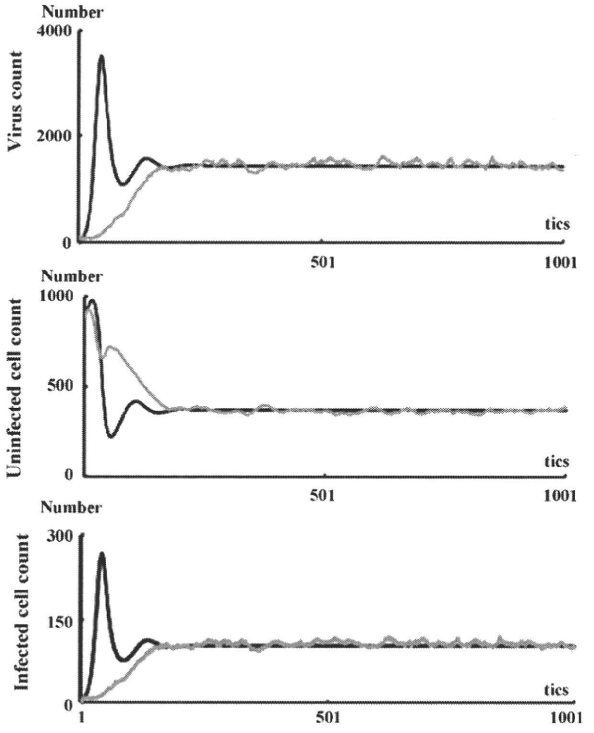

Fig. 3. Comparison of results of agent-based simulation and mathematical simulation. Both sets of results were consistent for the equilibrium phase, but differed in the shift in transition phase. Black line: mathematical model; grey line: results of simulation in RePast. Parameter values were set as 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 δ, and the death rate of viruses is c. The infection rate is indicated by β. 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

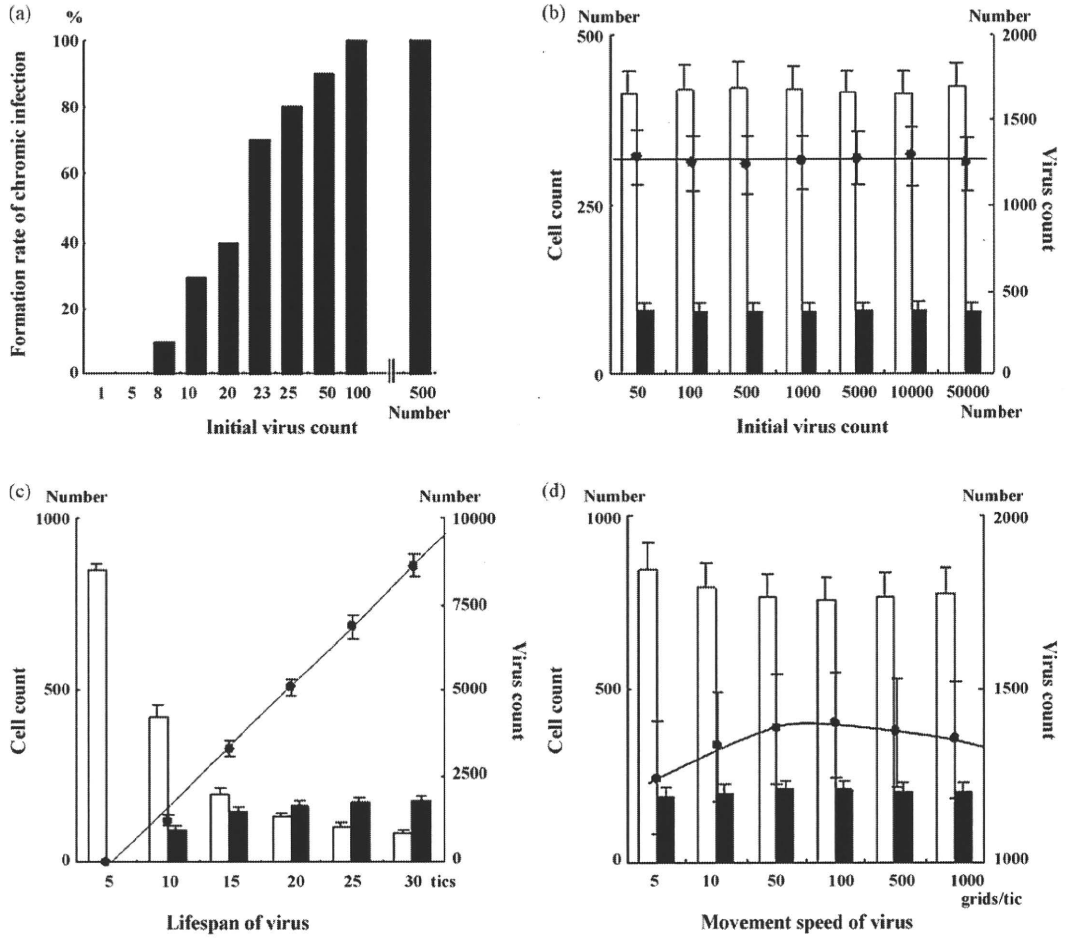

Fig. 4. Effects of changes in viral parameters. (a) The higher the initial virus count, the greater is the increase in the rate of formation of chronic infection, but (b) there was no effect on the conditions in the equilibrium phase. (c) Extending the virus lifespan increased the virus count. (d) Increasing 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

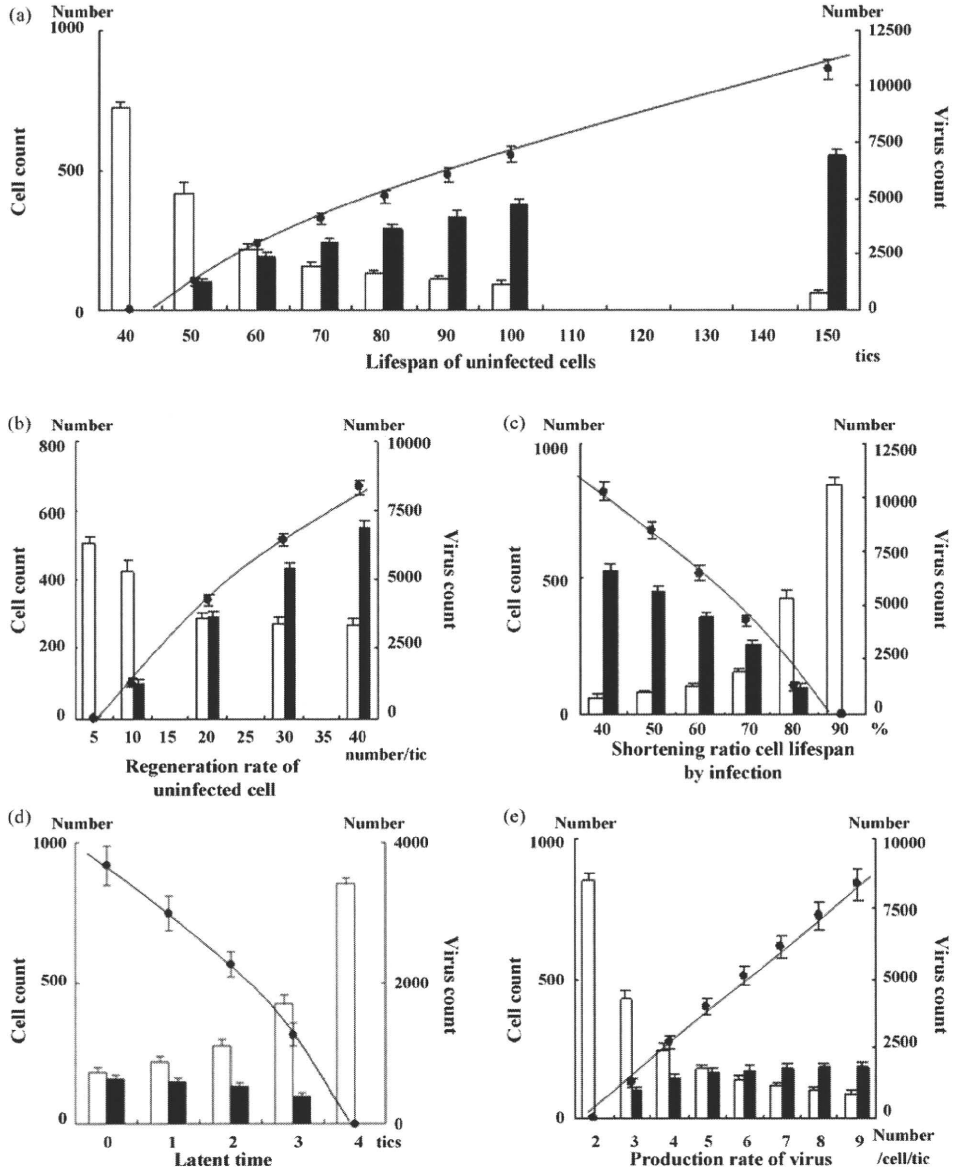

Fig. 5. Effects of changes in cell parameters. (a) Extending the uninfected cell lifespan and (b) increasing the uninfected cell regeneration rate increased the virus count. (c) Raising the lifespan-shortening ratio as a result of infection shortened the lifespan of infected cells, thereby decreasing the virus count. (d) Extending the latent period shortened the period of virus production from infected cells, thereby decreasing the virus count. (e) Increasing the virus production count resulted in a linear increase in equilibrium-phase virus count. Black circles: virus count; line: virus count approximation curve; white bars: uninfected cell count; black bars: infected cell count.

phase itself (Fig. 4b). Extending the lifespan of viruses resulted in a linear increase in equilibrium-phase virus count (Fig. 4c). Although the infected cell count increased, the rate of increase gradually declined. Changing the speed of viral movement resulted in the equilibrium-phase virus count to eventually decline after 100 grids/tic was reached, allowing movement over an area twice the size of the simulation space (Fig. 4d).

3.5. Uninfected Cell Parameters

Extending the lifespan of uninfected cells led to an increased virus count during the equilibrium phase (Fig. 5a). Increasing the uninfected cell regeneration rate also contributed to increased equilibrium-phase virus count (Fig. 5b). In both the cases, the increases in virus count and infected cell count were not linear, but showed a tendency for the rate of increase to decline gradually.

3.6. Infected Cell Parameters

We carried out an investigation of the effects of variation in the lifespan-shortening ratio on the virus count on the assumption that cell lifespan is shortened by infection. When this ratio was increased, the virus count decreased (Fig. 5c). An extended latent period was also related to a decreased virus count (Fig. 5d). However, the virus production from infected cells led to a linear increase in the virus count (Fig. 5e).

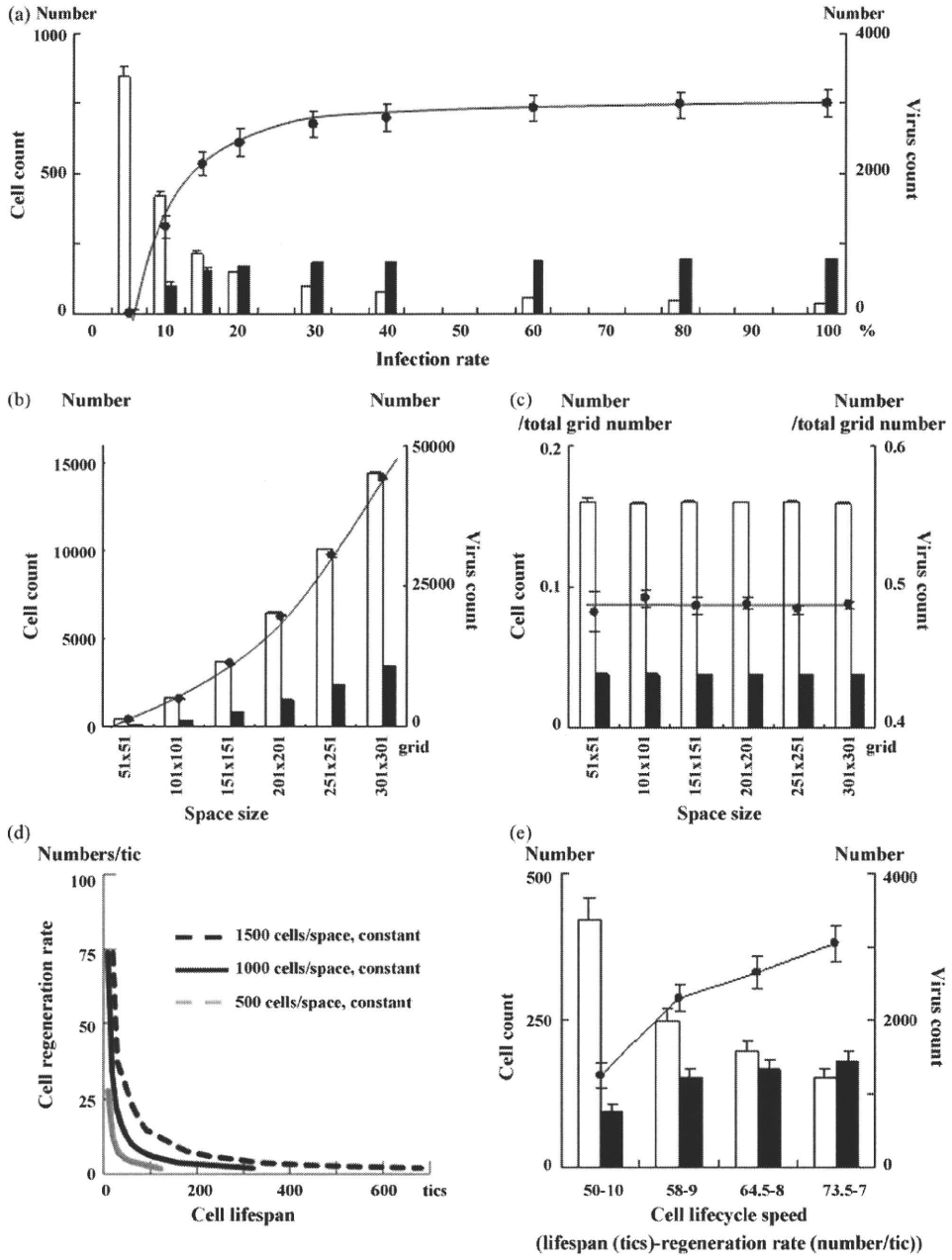

Fig. 6. (a) Increasing the infection rate increased the virus count in equilibrium periods, but the virus count did not change at infection rates of 30% or more. (b) The size of the simulation space increased not only virus count but also the cell count; however, (c) when virus and cell counts were divided by the total number of grids in the space, they were constant for all space sizes. (d) Changing the lifespan and regeneration rate of uninfected cells in opposite directions at the same time makes it possible to change only the cell cycle speed without altering the uninfected cell count. (e) When the cell cycle speed was reduced, the virus count increased toward the right of the graph. This may be because the effect of extending the lifespan of cells exceeds that of reducing their regeneration rate. (a–c and e) Black circles: virus count; line: virus count approximation curve; white bars: uninfected cell count; black bars: infected cell count.

3.7. Infection Rate and Space Size

Increasing the infection rate caused an increase in the virus count, but the change was minimal at an infection rate of 30% or more. The same results were seen for infected cell count, but a decrease in uninfected cell count resulted in a tendency for the infection rate to decrease by up to 60% (Fig. 6a).

The larger the space, higher the increase in both virus and cell counts (Fig. 6b). This increase was proportional to space size, however, when virus and cell counts were divided by the total number of grids in the space they were all constant (Fig. 6c).

3.8. Cell Cycle Speeds

Running a simulation with the initial virus count set to zero enables only the equilibrium condition for uninfected cells to be simulated. Changing the lifespan and regeneration rate of uninfected cells in opposite directions at the same time makes it possible

to change the cell cycle speed without altering the uninfected cell count (Fig. 6d). We used this technique to investigate how changing the cell cycle speed affected the equilibrium phase. Fig. 6e shows the results. Cell lifespan increases while the cell cycle speed declines. The equilibrium virus count increased in accordance with slower cell cycle speeds.

4. Discussion

In this study, we investigated the models using two agent-based simulation methods to program a simple virus–host chronic infection model. The same model written in two different programming language systems displayed the same results. The transient phase was unlike that seen in a mathematical simulation with no overshoot in virus count, but rather a smooth transition to the equilibrium phase. The virus count at the start of the simulation only had effect on the rate of infection development. Increases in virus lifespan, uninfected cell lifespan, uninfected cell regeneration rate, virus production count from infected cells, and infection rate all led to increased equilibrium-phase virus count. Rises in the infected cell lifespan-shortening ratio, latent period, and cell cycle speed decreased the equilibrium-phase virus count. The size of the space itself had no innate effect on the equilibrium phase, but a speed of movement of the virus that was twice the size of the space produced the maximum virus count.

Reproducibility is the basis for all scientific study, but there are many problems to prove it in computer simulations, such as programming bugs. As agent-based simulation deals with numerous agents individually, it requires vast amounts of calculations. Accumulation of very small change of values leads to large differences of results. In this study, we investigated two programs based on two programming languages to confirm the reproducibility of our simulation results in different programming languages. The results of two simulations were consistent, but in StarLogo, the lifespan parameters had a tendency to be lower than when they were set while simulations were actually in progress. This may be because the number of digits used in calculations was different between the two programs. RePast performs calculations to at least eight decimal places. In StarLogo, the library settings only enable settings to be made up to five decimal places. It is probable that these small differences accumulate during repeated calculations and are reflected in the simulation. Ultimately, we confirmed that the differences in results obtained by using different libraries and programming languages were not innate and by making the parameters consistent during simulation, consistent results were obtained.

Mathematical models using formulae for HIV therapy was published in 1994, the method has since been applied to HBV and HCV (Ho et al., 1995; Nowak et al., 1996; Neumann et al., 1998), and they were thought to be good reflections of the reality. In the mathematical model, viruses and cells are conceived as individuals in the concept itself, but both of them are perceived *en masse* when calculations are performed. However a feature of the agent-based simulation is that it deals with individual viruses and cells as separate agents. By moving each agent individually, it probes the factors influencing overall shifts from the micro viewpoint. When the space is viewed as a whole, it is possible to watch on the screen the collective movement of groups of agents. Recently, models that provide a visual representation of Epstein-Barr virus and HIV infection have been reported, both of which are useful for an instinctive and intuitive understanding (Duca et al., 2007; Shapiro et al., 2008; Castiglione et al., 2007).

In agent-based simulation model, virus count transit smoothly to the equilibrium phase. On the other hand, virus counts overshoot during transient phase in mathematical model. We think this difference is derived from technicality of different modeling. The difference in concepts between mathematical models and agent-based models is the space. The mathematical model has no space in concept, but agents move across the space in the agent-based model. In agent-based models, the densities of virus and cells change overtime especially in the transition phase because of the limited space. These changes of the densities of virus and cells lead to the dynamic change of the encounter rate of viruses and cells. The mathematical model does not make such concept of the density; the encounter rate is constant. This may be the reason for the difference between two models in the transition phase. Since no overshoot of virus counts in transient phase had been reported from in vivo studies of hepatitis C virus and simian immunodeficiency virus (Dahari et al., 2005; Nowak et al., 1997), agent-based model correlates with actual biology in vivo at least for these viruses. The increase of initial virus count at the start of simulation correlates with higher encounter rate of viruses and cells which make the linear increasing of infection forming rate. Mathematical model can only express the infection formation rate as "infected or not".

The importance of viral passing speed in the agent-based model is also explained by the "space". Although the virus actually moves through the blood stream in our body and virus could not decide their moving speeds by themselves, there is most appropriate speed for virus to meet the cells on the simulation space by the highest probability. The effect of cell cycle speed should be mentioned by another affection of the space. A fast cell cycle speed means that the lifespan of uninfected cells is short. Then fast cell cycle speed leads to the short lifespan of infected cells. A higher regeneration rate for uninfected cells results in a higher rate of infection among uninfected cells by viruses, but in situations where viruses and cells are dispersed around the space this is ineffective in increasing the infection rate, as the latter depends on the probability that they will encounter one another. As a result, the infected cell count decreases during the equilibrium phase, as does the virus count.

In this study, we confirmed the reproducibility and usability of agent-based models in expressing the interaction between viruses and cells. A feature of this simulation system is that it uses the concept of space as actual space, which means that the existence of the space becomes an additional controlling factor on the simulation results. This is a concept that is absent from mathematical models. The reality is that we have a spatial existence, and an advantage of the agent-based simulation system is the fact that it accounts for the space. Another feature of the simulation system is that it enables the condition to be perceived in visual terms, making it easy to understand. However it may be affected by computer performance and by the limitations of programming languages or the program itself, this system may offer a powerful tool for the future analysis of real virus–host interaction disease.

Conflict of interest

No conflicts of interest exist for all authors.

References

Castiglione, F., Pappalardo, F., Bernaschi, M., Motta, S., 2007. Optimization of HAART with genetic algorithms and agent-based models of HIV infection. Bioinformatics 23, 3350–3355, doi:10.1093/bioinformatics/btm408.

Dahari, H., Major, M., Zhang, X., Mihalik, K., Rice, C.M., Perelson, A.S., Feinstone, S.M., Neumann, A.U., 2005. Mathematical modeling of primary hepatitis c infection: noncytolytic clearance and early blockage of virion production. Gastroenterology 128, 1056–1066, doi:10.1053/j.gastro.2005.01.049.

Duca, K.A., Shapiro, M., Delgado-Eckert, E., Hadinoto, V., Jarrah, A.S., Laubenbacher, R., Lee, K., Luzuriaga, K., Polys, N.F., Thorley-Lawson, D.A., 2007. A virtual look at Epstein-Barr virus infection: biological interpretations. PLoS Pathog. 3, 1388–1400, doi:10.1371/journal.ppat.0030137.

Gilbert, N., Bankes, S., 2002. Platforms and methods for agent-based modelling. Proc. Natl. Acad. Sci. U.S.A. 99 (Suppl. 3), 7197–7198.

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M., 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373, 123–126, doi:10.1038/373123a0.

Naniche, D., 2009. Human immunology of measles virus infection. Curr. Top. Microbiol. Immunol. 330, 151–171.

Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S., 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 282, 103–107, doi:10.1126/science.282.5386.103.

Nowak, M.A., Bonhoeffer, S., Hill, A.M., Boehme, R., Thomas, H.C., McDade, H., 1996. Viral dynamics in hepatitis B virus infection. Proc. Natl. Acad. Sci. U.S.A. 93, 4398–4402.

Nowak, M.A., Lloyd, A.L., Vasquez, G.M., Wiltrout, T.A., Wahl, L.M., Biscofberger, N., Williams, J., Kinter, A., Fauci, A.S., Hirsch, V.M., Lifson, J.D., 1997. Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. J. Virol. 71, 7518–7525.

Shapiro, M., Duca, K.A., Lee, K., Delgado-Eckert, E., Hawkins, J., Jarrah, A.S., Laubenbacher, R., Polys, N.F., Hadinoto, V., Thorley-Lawson, D.A., 2008. A virtual look at Epstein-Barr virus infection: simulation mechanism. J. Theor. Biol. 252, 633–648, doi:10.1016/j.jtbi.2008.01.032.

See, H., Wark, P., 2008. Innate immune response to viral infection of the lungs. Paediatr. Respir. Rev. 9, 243–250, doi:10.1016/j.prrv.2008.04.001.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 38, pp. 29270–29278, September 17, 2010

Induction of Pluripotent Stem Cells from Human Third Molar Mesenchymal Stromal Cells*[S]♦

Received for publication, August 16, 2009, and in revised form, June 26, 2010 Published, JBC Papers in Press, July 1, 2010, DOI 10.1074/jbc.M109.055889

Yasuaki Oda[‡], Yasuhide Yoshimura[‡1], Hiroe Ohnishi[‡1], Mika Tadokoro[‡], Yoshihiro Katsube[‡], Mari Sasao[‡], Yoko Kubo[‡], Koji Hattori[‡], Shigeru Saito[§], Katsuhisa Horimoto[§], Shunsuke Yuba[‡], and Hajime Ohgushi[‡2]

From the [‡]Tissue Engineering Research Group, Health Research Institute, National Institute of Advanced Industrial Science and Technology, Hyogo 661-0974, Japan, and the [§]Biological Network Team, Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

The expression of four transcription factors (*OCT3/4*, *SOX2*, *KLF4*, and *MYC*) can reprogram mouse as well as human somatic cells to induced pluripotent stem (iPS) cells. We generated iPS cells from mesenchymal stromal cells (MSCs) derived from human third molars (wisdom teeth) by retroviral transduction of *OCT3/4*, *SOX2*, and *KLF4* without *MYC*, which is considered as oncogene. Interestingly, some of the clonally expanded MSCs could be used for iPS cell generation with 30–100-fold higher efficiency when compared with that of other clonally expanded MSCs and human dermal fibroblasts. Global gene expression profiles demonstrated some up-regulated genes regarding DNA repair/histone conformational change in the efficient clones, suggesting that the processes of chromatin remodeling have important roles in the cascade of iPS cells generation. The generated iPS cells resembled human embryonic stem (ES) cells in many aspects, including morphology, ES marker expression, global gene expression, epigenetic states, and the ability to differentiate into the three germ layers *in vitro* and *in vivo*. Because human third molars are discarded as clinical waste, our data indicate that clonally expanded MSCs derived from human third molars are a valuable cell source for the generation of iPS cells.

In 2006, groundbreaking research demonstrated that mouse somatic cells could be reprogrammed to a pluripotent state by transduction of four transcription factors (*Oct3/4*, *Sox2*, *Klf4*, and *Myc*) (1). Furthermore, the researchers reported that the same four factors are also effective in reprogramming human somatic cells (2). Another group reported that a different set of four factors (*OCT3/4*, *SOX2*, *NANOG*, and *LIN28*) could reprogram human cells (3). The reprogrammed cells, named induced pluripotent stem (iPS)[3] cells, closely resembled human embryonic stem (ES) cells in many aspects including morphology, gene expression, surface markers expression, epigenetic states, and the ability to differentiate into the three germ layers (endoderm, mesoderm, and ectoderm) *in vitro* and *in vivo* (1–12). Generation of iPS cells opened up a new avenue for the generation of patient-specific pluripotent stem cells, which are useful for drug screening, understanding the mechanisms of pathogenesis, and cell transplantation therapies (13–16).

However, three major concerns exist in the current reprogramming strategies for clinical applications: (i) the low reprogramming efficiency of human somatic cells makes it difficult to generate patient-specific iPS cells, especially using a small amount of the tissue of the patient; (ii) genomic integration of retro- or lentiviral fragments might cause carcinogenesis (17–19); and (iii) reactivation of *MYC* might also cause malignant tumor formation (9). Although iPS cells can be generated by three transcription factors (*OCT3/4*, *SOX2*, and *KLF4*) without *MYC*, reprogramming efficiency was significantly reduced (20).

Recent studies have demonstrated human iPS cell generation from various cell origins, such as neonatal or adult dermal fibroblasts (2, 4, 5, 7, 20), foreskin fibroblasts (2–4, 7, 8), lung fibroblasts (3, 4), fibroblast-like synoviocytes (2), and keratinocytes (6). Human dermal fibroblasts (HDFs) are easily obtained by biopsy. However, cellular senescence and/or low proliferation capability have caused failure of iPS generation (4). One ideal cell source for iPS cell generation is discarded tissue, which contains cells with high proliferation capability with less incidence of cellular senescence. We and other groups previously reported that mesenchymal stromal cells (MSCs) from teeth or third molars (wisdom teeth), which were usually discarded as clinical waste, showed high proliferation when compared with MSCs from bone marrow (21–23).

Quite recently, another group reported that mesenchymal-like stem/progenitor cells from dental tissue could be reprogrammed by four factors transduction of *OCT4/SOX2/KLF4/MYC* or *OCT4/SOX2/NANOG/LIN28* (24). However, the *MYC* is an oncogene, and overexpression of *LIN28* might promote carcinogenesis via repression of *let-7* and following derepressing the targets genes involved in *MYC* (25). Here, we show that various MSCs from human third molars could be reprogrammed to a pluripotent state by retroviral transduction with *OCT3/4*, *SOX2*, and *KLF4* without *MYC* or *LIN28*. In particular, some clonally expanded MSCs showed extremely higher reprogramming efficiency than that of HDFs. We also

* This work was supported by a grant from the project for realization of regenerative medicine of the Ministry of Education, Culture, Sports, Science and Technology (MEXT).
♦ This article was selected as a Paper of the Week.
[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4 and Tables S1–S3.
[1] Both authors contributed equally to this work.
[2] To whom correspondence should be addressed: 3-11-46 Nakouji, Amagasaki, Hyogo 661-0974, Japan. Tel.: 81-6-6494-7806; Fax: 81-6-6494-7861; E-mail: hajime-ohgushi@aist.go.jp.
[3] The abbreviations used are: iPS, induced pluripotent stem; MSC, mesenchymal stromal cell; HDF, human dermal fibroblast; EB, embryoid body; ALP, alkaline phosphatase; PARP, poly(ADP-ribose)polymerase; Tg, transgene.

addressed the reprogramming efficiency from global gene expression profiles.

EXPERIMENTAL PROCEDURES

Plasmid Construction—Reading frame cassette A (Invitrogen) was introduced into the EcoRI site of the pMXs retroviral vector (26). The open reading frames of human *OCT3/4* (*POU5f1* isoform-1), *SOX2*, and *KLF4* were amplified by reverse transcription-polymerase chain reaction (RT-PCR) and cloned into pENTR-D/TOPO (Invitrogen). To evaluate the viral infection efficiency, the open reading frame of DsRed-Express from pIRES2-DsRed-Express (Clontech) was amplified by PCR and cloned into pENTR-D/TOPO (Invitrogen). All of the genes were transferred to the pMXs retroviral vector (kindly donated by Dr. Kitamura) by Gateway Technology (Invitrogen), according to the manufacturer's instructions.

Cell Culture—This study was approved by the ethics committee of the National Institute of Advanced Industrial Science and Technology. Isolation of third molars and culture expansion of MSCs from the molars were carried out from three donors (10, 16, and 13 years old) after informed consent was obtained. The cultured method of these MSCs was described in our previous report (23). The frozen stocked MSCs were thawed and used for the generation of iPS cells. HDFs were purchased from Cell Applications. Platinum-A (Plat-A) cells were purchased from Cell Biolabs (27). SNL76/7 feeder cells were purchased from the European Collection of Cell Cultures. MSCs were maintained in minimum essential medium α (Invitrogen) containing 15% fetal bovine serum (FBS; Invitrogen), 100 units/ml penicillin, and 100 μg/ml streptomycin (Invitrogen). HDF, Plat-A, and SNL feeder cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) containing 10% FBS, 100 units/ml penicillin, and 100 μg/ml streptomycin. The iPS cells were generated and maintained in human ES cell medium (DMEM/F-12 with GlutaMAX-I (Invitrogen), supplemented 20% knock-out serum replacement (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 0.1 mM 2-mercaptoethanol (Invitrogen), 100 units/ml penicillin, and 100 μg/ml streptomycin) supplemented with 5 ng/ml recombinant human basic fibroblast growth factor (basic FGF; WAKO). For passaging, MSCs, HDFs, Plat-A, and SNL76/7 feeder cells were trypsinized with 0.05% trypsin/0.53 mM EDTA (Invitrogen). The iPS cells were passaged every 5–7 days using dissociation solution (0.25% trypsin (Invitrogen), 0.1 mg/ml collagenase type IV (Invitrogen), 10 mM $CaCl_2$ (WAKO), and 20% knock-out serum replacement in distilled water). NC3T3-G2/PA6 (PA6) cells (RIKEN Bioresource Center, Tsukuba, Japan) were maintained in minimum essential medium α containing 10% FBS, 100 units/ml penicillin, and 100 μg/ml streptomycin.

Retroviral Production—Plat-A packaging cells were seeded at 8×10^6 cells/100 mm dish and cultured overnight. The next day, pMXs retroviral vectors containing the open reading frames of *OCT3/4*, *SOX2*, *KLF4*, and *DsRed-Express* were transfected into Plat-A cells with FuGENE HD Transfection Reagent (Roche Diagnostics). Viral supernatants were collected 48 h and 72 h after transfection, then filtered through a 0.45 μm pore size filter (Sartorius) and supplemented with 4 mg/ml Polybrene (Sigma). The target cells were transduced with *OCT3/4*:*SOX2*:*KLF4* = 1:1:1 mixture of viral supernatant. To determine the viral transduction efficiency, the retroviral supernatant containing *DsRes-Express* was transduced to MSCs and HDF. Medium was changed every other day and cultured for 5 days. The cells were trypsinized and analyzed by a FACSCalibur (BD Biosciences).

Generation of iPS Cells—MSCs and HDF cells were seeded at 5×10^5 cells/100-mm dish and cultured overnight. The next day, the cells were infected with viral supernatant for 24 h and then replaced with fresh viral supernatant. After 3 days after infection, the infected cells were seeded at 5×10^4 cells/100-mm dish on SNL feeder cells. The next day, the medium was replaced with human ES cell medium supplemented with 5 ng/ml basic FGF. The medium was changed every other day. Around 25–30 days after infection, iPS colonies were picked based on human ES cell-like colony morphology, and medium was changed every day. The picked colonies were subsequently expanded and maintained on SNL feeder cells in human ES cell medium. Reprogramming efficiency was determined as the number of total ES-like colonies per the total number of infected cells.

Proliferation Rate of Parental Cells—Each MSCs and HDF was seeded at 1×10^4 cells/6-well plate and cultured for 5 days. The cultured/proliferated cells were trypsinized and counted.

RNA Isolation and Reverse Transcription—Total RNA was isolated using the RNeasy mini kit (Qiagen) and treated with TURBO DNA-free (Applied Biosystems), according to the manufacturer's instructions. One microgram of total RNA was used for cDNA synthesis using ReverTra Ace-α (TOYOBO) and oligo$(dT)_{20}$ primers, according to the manufacturer's instructions. PCR was performed using ExTaq HS (TAKARA BIO). Primer sequences are shown in supplemental Table S1.

Alkaline Phosphatase Staining and Immunocytochemistry—Alkaline phosphatase staining was performed using the leukocyte alkaline phosphatase kit (Sigma) according to the manufacturer's instructions. For immunocytochemistry, cells were fixed with PBS containing 4% paraformaldehyde for 10 min at room temperature. After washing with PBS, the cells were treated with PBS containing 0.1% Triton X-100 for 10 min and then 1% bovine serum albumin (BSA, A2153, Sigma) for 10 min at room temperature. Primary antibodies included stage-specific embryonic antigen (SSEA)-3 (1:100, MAB4303, Millipore), SSEA-4 (1:100, MAB4304, Millipore), tumor-related antigen (TRA)-1–60 (1:100, ab16288-200, Abcam), TRA-1–81 (1:100, ab16289-200, Abcam), OCT4 (1:100, ab19857-100, Abcam), NANOG (1:50, ab21624, Abcam), SOX17 (1:200, AF1924, R&D Systems), α-fetoprotein (1:200, MAB1368, R&D Systems), desmin (1:200, RB-9014, Lab Vision), α-smooth muscle actin (prediluted, N1584, Dako), βIII-tubulin (1:200, CBL412, Millipore), and tyrosine hydroxylase (1:200, AB152, Millipore). Secondary antibodies used were Alexa Fluor 488 donkey anti-mouse IgG (1:200, A21202), Alexa Fluor 568-conjugated goat anti-mouse IgG (1:200, A11031), Alexa Fluor 568 goat anti-mouse IgM (1:200, A21043), Alexa Fluor 488 goat anti-mouse IgM (1:200, A21042), Alexa Fluor 568 goat anti-rabbit IgG (1:200, A11011), Alexa Fluor 488 goat anti-rabbit IgG (1:200, A11034), Alexa Fluor 488 donkey anti-goat IgG (1:200, A11055),

and Alexa Fluor 488 goat anti-rat IgM (1:200, A21212) (all from Molecular Probes). Nuclei were stained with 0.2 μg/ml Hoechst 33342 (Molecular Probes).

Bisulfite Sequencing—One microgram of genomic DNA was treated with the EpiTect bisulfite kit (Qiagen), according to the manufacturer's recommendations. The promoter regions of the human *OCT3/4* and *NANOG* genes were amplified by PCR. The PCR products were subcloned into pCR2.1-TOPO (Invitrogen). Ten clones of each sample were verified by sequencing with the M13 universal primer. Primer sequences used for PCR amplification are shown in supplemental Table S1.

Detection of Telomerase Activity—Telomerase activity was verified with a TRAPeze telomerase detection kit (Millipore) according to the manufacturer's instructions. Each sample was separated by Tris-borate-EDTA-based 10% polyacrylamide gel electrophoresis. The gel was stained with ethidium bromide.

In Vitro Differentiation—For embryoid body (EB) formation, human iPS cells were harvested by treatment with dissociation solution and transferred to a low attachment culture dish (PrimeSurface; Sumitomo Bakelite) in human ES cell medium without basic FGF. The medium was changed every other day. Eight days after floating culture, the EBs were transferred onto gelatin-coated plates and cultured in the same medium for another 8 days. PA6 cells were used for differentiation into dopaminergic neurons. The PA6 cells were seeded on gelatin-coated plates and cultured for 4 days. Small clumps of iPS cells were seeded on the PA6 cells in Glasgow minimum essential medium (Invitrogen) containing 10% knock-out serum replacement, 0.1 mM non-essential amino acids, 0.1 mM 2-mercaptoethanol, 100 units/ml penicillin, and 100 μg/ml streptomycin.

Teratoma Formation—One well of a 6-well plate of iPS cells cultured on SNL feeder cells was harvested by dissociation solution treatment and washed twice with PBS. Cell clumps of iPS cells were suspended in 50 μl of human ES cell medium. Twenty-five μl of cell clump suspension were injected into each testis of a severe combined immunodeficient (SCID) mouse. Nine weeks after injection, tumors were collected and fixed with 4% paraformaldehyde/PBS. The paraffin-embedded tumors were sectioned and stained with hematoxylin and eosin.

Karyotype Analysis—Chromosomal G-band analyses and multicolor FISH were performed at the Nihon Gene Research Laboratories (Sendai, Japan).

DNA Microarray—The microarray study was carried out using GeneChip Human Genome U133 Plus 2.0 gene expression arrays (Affymetrix). The analyses were performed according to Affymetrix technical protocols. Data from these experiments and the GEO data base were analyzed with GeneSpring GX10 software. For the cluster analyses, our data from the GeneChip Human Genome U133 Plus 2.0 gene expression arrays (archived at GEO as accession number GSE16963) were compared with DNA microarray data for the human ES cell lines BG01 and H9 retrieved from GEO datasets (GSM367061, GSM367062) as being representative of human ES cells for comparison purposes.

Statistical Analysis—Differences between the two data sets from DNA microarray were evaluated by the Student's *t* test for the expression of each gene. Then, the false discovery rate was estimated using the Benjamini-Hochberg procedure (52). Differentially expressed genes were selected if they satisfied false discovery rate <0.05 in the average values for each set being compared.

Real-time PCR—PCR reactions were carried out in a StepOnePlus real-time PCR system (Applied Biosystems) using TaqMan gene expression master mix (Applied Biosystems) according to the manufacturer's instructions. The following TaqMan primers and probes (TaqMan gene expression assays; Applied Biosystems) were used: β-actin (Hs99999903_m1), *OCT3/4* (Hs03005111_g1), *SOX2* (Hs01053049_s1), *NANOG* (Hs02387400_g1), *KLF4* (Hs00358836_m1), and *P53* (Hs99999147_m1). The expression of genes of interest was normalized to that of β-actin in all samples. Relative expression of genes of interest was estimated using the relative standard curve method.

RESULTS

Generation of iPS Cells from Culture-expanded MSCs—We have previously reported that culture-expanded MSCs from human third molars show high proliferation as well as differentiation capability into mesenchymal lineages (28). The MSCs from three donors (10F, 10-year-old female; 16F, 16-year-old female; and 13M, 13-year-old male) were seeded at a cell density of 5×10^5 cells/100-mm dish the day before retrovirus infection (day 0). We introduced the pMXs retrovirus vectors containing human *OCT3/4*, *SOX2*, and *KLF4* into PLAT-A packaging cells (26, 27). Viral supernatants were collected at 48 and 72 h after transduction and infected into MSCs (day 1 and day 2). At day 4, viral-infected cells were trypsinized and seeded onto SNL76/7 feeder cells (5×10^4 viral-infected cells/100-mm dish) and then cultured in human ES cell medium supplemented with basic FGF. At ~25 days, we detected several colonies displaying human ES cell-like morphology (supplemental Fig. S1). At day 30, the reprogramming efficiencies of MSCs from 10F, 16F, and 13M were 0.0302, 0.0042, and 0.0026%, respectively ($n = 10$). Thus, the 10F cells showed higher reprogramming efficiency when compared with the other donors (Fig. 1*A*).

Generation of iPS Cells from Clonally Expanded MSCs—One of the possibilities was that the 10F cells might contain tissue-specific stem or progenitor cells, such as dental pulp stem cells (21) or stem cells from human exfoliated deciduous teeth (22). We also reported that tooth germ progenitor cells, which were clonally expanded mesenchymal cells from human third molars, differentiated into osteoblasts, neural cells, and hepatocytes (23). Consequently, we examined whether the high reprogramming efficiency of the 10F cells was due to the presence of stem/progenitor cells. For this purpose, we clonally expanded the MSCs from the 10F cells and selected five clones, named 10F-5, 10F-8, 10F-15, 10F-101, and 10F-107. The 10F-5, 10F-8, and 10F-15 clones could differentiate into osteoblasts under osteogenic culture conditions; however, the 10F-101 and 10F-107 clones could not (data not shown). We found that 10F-5, 10F-8, and 10F-15 exhibited rapid proliferation when compared with 10F-101, 10F-107, or HDF (Fig. 1*B*).

Using the methods described above, we seeded these five clonal cells and HDFs and then transduced human *OCT3/4*, *SOX2*, and *KLF4* with retrovirus. At day 30, we counted the

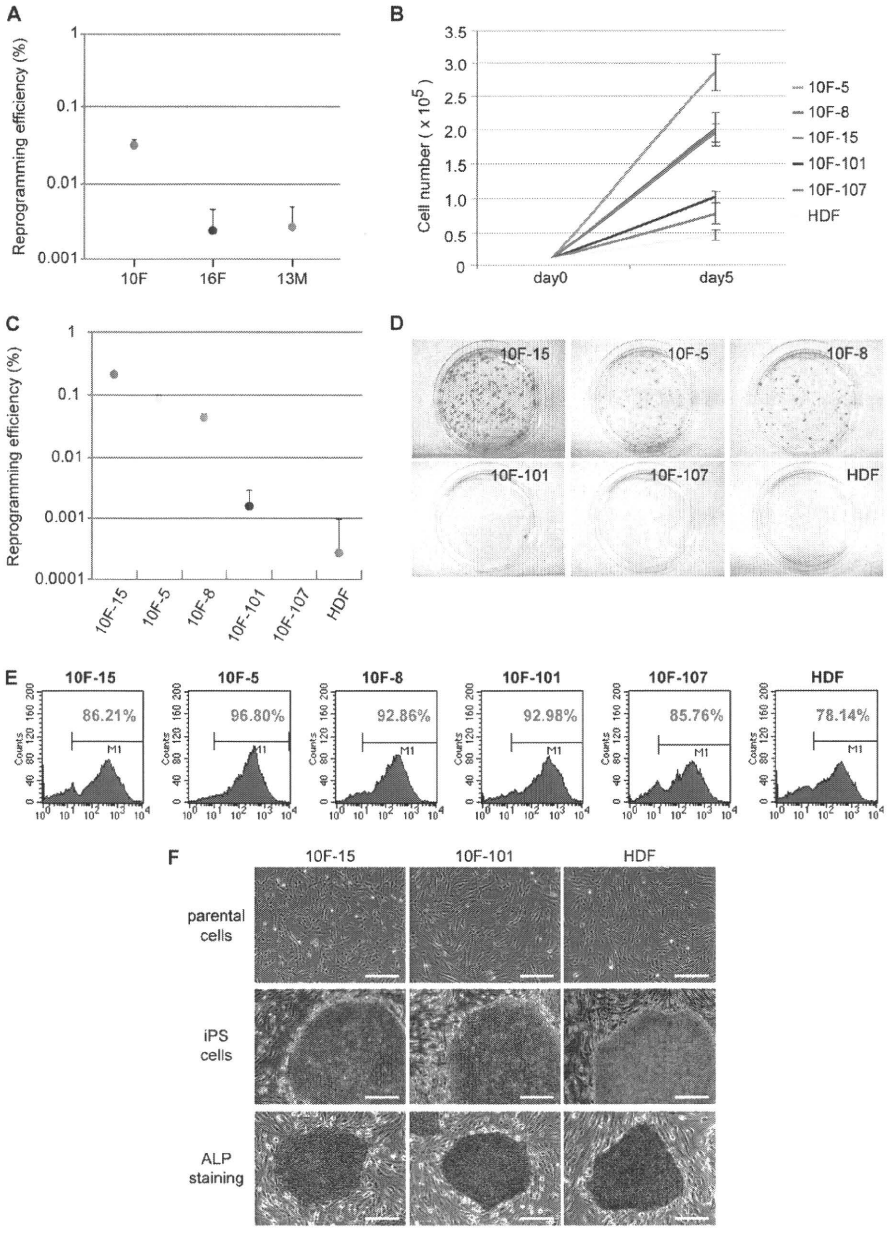

FIGURE 1. **Generation of iPS cells from MSCs.** *A*, reprogramming efficiency of non-cloned parental MSCs from the third molars of 10-, 16-, and 13-year-old donors. The *bars* represent the S.D. ($n = 10$). *B*, proliferation rate of clonally expanded MSCs from 10-year-old donor and HDF at day 5. The *bars* indicate S.D. ($n = 3$). *C*, Reprogramming efficiency of clonally expanded MSCs and HDF cells. The *closed circles* represent average of two independent experiments, and the *bars* indicate S.D. ($n = 8$). *D*, ALP staining of three factors transduced cells at day 30. *E*, viral transduction efficiency of the pMXs retroviral vector. The percentages indicate the ratio of DsRed-Express-positive cells. *F*, parental cells of clonally expanded MSC and HDF cells (*upper panel*), iPS cells from the parental cells (*middle panel*), and ALP staining of each iPS cell (*lower panel*). *Scale bars* = 100 μm.

number of colonies that resembled the ES cell-like morphology and performed alkaline phosphatase (ALP) staining. Numerous ALP-positive colonies were seen in 10F-5, 10F-8, and 10F-15; however, a few colonies were seen in 10F-101 and HDF. The colonies were not seen in 10F-107. Therefore, the rapid proliferation clones were efficiently reprogrammed (Fig. 1, *B*, *C*, and *D*). To check the viral transduction efficiency, we introduced pMXs retrovirus vector containing *DsRed-Express* into each parental cell. There was no significant difference of viral transduction efficiency among these cells (Fig. 1*E*). These colonies were picked and expanded for several passages and then checked for ALP activity. All the colonies showed high ALP activity (Fig. 1*F*). The data in the following sections were obtained from the putative iPS colonies derived from the clonally expanded MSCs. Particularly, clones of 10F-15 and 10F-101 were used. Corresponding parental cells and HDF were also used as controls.

Characterization of iPS Cells from Clonally Expanded MSCs—To confirm that the cells from the putative colonies were iPS cells, we examined the human ES cell markers by immunocytochemistry, DNA methylation states, RT-PCR, and telomerase activity. Immunocytochemistry showed that the iPS cells expressed human ES cell-specific surface antigens, such as SSEA-3, SSEA-4, TRA-1–60, and TRA-1–81. These cells also expressed the ES cell-specific transcription factors, OCT3/4, and NANOG (Fig. 2*A* and supplemental Fig. S2).

The DNA methylation states of CpG dinucleotides in the *OCT3/4* and *NANOG* promoter regions were evaluated by bisulfite genomic sequencing analysis. All The parental cells were highly (*OCT3/4*) or partially (*NANOG*) methylated, whereas all of the iPS cells were highly unmethylated (Fig. 2*B*). These data suggested that these promoters were active in these iPS cells.

All of the iPS cells expressed undifferentiated ES cell maker genes, such as *OCT3/4*, *SOX2*, *NANOG*, reduced expression 1 (*REX1*), undifferentiated embryonic cell transcription factor 1 (*UTF1*), growth and differentiation factor 3 (*GDF3*), developmental pluripotency-associated 2 (*DPPA2*), *DPPA4*, and *DPPA5*. However, the cells did not express retroviral transgenes (Tgs), such as *Tg-OCT3/4*, *Tg-SOX2*, and *Tg-KLF4*, indicating retroviral silencing. The cells before retroviral transduction (parental cell) expressed *KLF4* and *MYC* but did not express other markers (Fig. 2*C*).

We also examined telomerase activity, which is known to be highly activated in cancer cells and ES cells. Each iPS cell showed high telomerase activity, whereas parental cells as well as heat-treated cells did not (Fig. 2*D*).

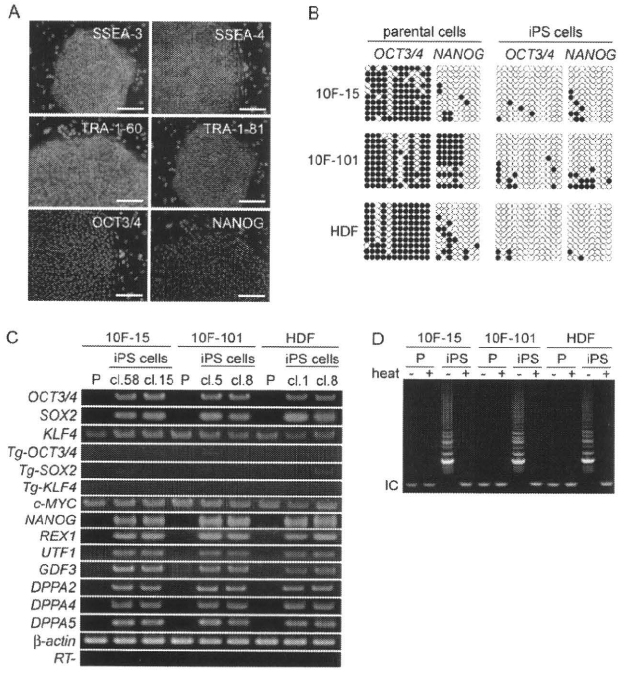

FIGURE 2. **Characterization of iPS cells from clonally expanded MSCs and HDF.** *A*, immunocytochemistry of SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, OCT3/4, and NANOG for iPS cells. *Scale bars* = 100 μm. *B*, DNA methylation states of the *OCT3/4* and *NANOG* promoter regions. *Open circles* indicate unmethylated CpGs, and *closed circles* indicate methylated CpGs. *C*, RT-PCR analysis of ES cell marker genes and retroviral Tgs for iPS cells from different colonies (*cl.58/cl.15* derived from 10F-15 clonal cells, *cl.5/cl.8* derived from 10F-101 clonal cells, and *cl.1/cl.8* derived from HDF). *D*, telomerase activity detected by TRAP assay. +, heat-treated; −, non-treated samples; *P*, parental cell. *IC*, internal control.

Embryoid Body Formation and in Vitro Differentiation—To examine the differentiation potential of iPS cells, we performed floating culture to demonstrate embryoid body formation. After 8 days of floating culture, iPS cells formed embryoid bodies seen as sphere structures (Fig. 3*A*). These embryoid bodies were transferred onto gelatin-coated plates for another 8 days to induce spontaneous differentiation. The cells attached to the culture dish showed various types of cell morphologies; immunocytochemistry demonstrated that the cells were positive for SOX17 (marker for endoderm, Fig. 3*B*), α-fetoprotein (AFP, endoderm; Fig. 3*C*, *green*), desmin (mesoderm; Fig. 3*C*, red), α-smooth muscle actin (α-SMA, mesoderm, Fig. 3*D*), and βIII-tubulin (ectoderm, Fig. 3*E*). For directed differentiation into neural lineages, iPS cells were seeded onto PA6 cells and cultured for 18 days in differentiation medium (29). Immunocytochemistry demonstrated the appearance of βIII-tubulin (*red*)- and tyrosine hydroxylase (*green*)-positive neural cells (Fig. 3*F*). Therefore, iPS cells had successfully differentiated into different types of cells.

RT-PCR analysis confirmed that these EB-derived various types of cells expressed *AFP*, *FOXA2* (markers for endoderm), brachyury, *TnTc* (mesoderm), *MAP2*, and *PAX6* (endoderm) (Fig. 3*G*). These data indicate that iPS cells derived from clonally expanded MSCs could differentiate into cell types of three embryonic germ layers.

Teratoma Formation—To test for pluripotency, we injected iPS cells into the testis of severe combined immunodeficient mice. Nine weeks after injection, all showed tumor formation. Histological examination of the tumor demonstrated representative tissues of three embryonic germ layers, such as gut-like epithelium (endoderm), cartilage (mesoderm), and neuroepithelial rosettes (ectoderm) (Fig. 4, *A–I*). These findings indicated that the tumor was teratoma. Thus, like ES cells, iPS cells derived from clonally expanded MSCs have *in vivo* capability of differentiation into the three embryonic germ layers.

Karyotype Analysis—We also performed karyotype investigations by chromosomal G-band analysis (supplemental Fig. S3) and multicolor FISH analysis (data not shown). Two lines of iPS cells from 10F-101 showed an additional change in chromosome 16. However, the corresponding parental cell also showed the addition at the same locus, suggesting that the chromosomal aberration was already present in the parental cell (data not shown). Other iPS cells (10F-15 and HDF) showed normal karyotypes (supplemental Fig. S3).

Global Gene Expression Profiles (Parental Cells versus iPS Cells)—We explored global gene expression profiles of the parental cells (10F-15, 10F-101, and HDF) and iPS cells. Previous studies reported that expression of reprogramming factors (*OCT3/4*, *SOX2*, *KLF4*, *MYC*, *NANOG*, *LIN28*, and *UTF1*) and inhibition of *p53* expression were effective for the generation of iPS cells (2, 3, 8). Microarray analysis showed that the mRNA levels of reprogramming factors, including *OCT3/4*, *SOX2*, *NANOG*, *LIN28*, and *UTF1*, were highly expressed in iPS cells when compared with those of the parental cells (Fig. 5*A*). The expression levels of *MYC* and *p53* were similar in each parental cell and iPS cells. However, *KLF4* expression was significantly lower (0.1-fold) in 10F-15 parental cells when compared with that of other parental cells and iPS cells. In addition, *KLF5* expression was slightly higher (3-fold) when compared with that of the other cells (Fig. 5*A*).

Scatter plot analysis showed that ES cell marker genes, such as *OCT3/4*, *NANOG*, *LIN28*, sal-like 4 (*SALL4*), and teratocarcinoma-derived growth factor 1 (*TDGF1*), were highly expressed in iPS cells (Fig. 5*B*). Heat map analysis showed that gene expression profiles of each iPS cell were similar but different from the corresponding parental cell. These profiles of iPS cells were also similar to those of BG01 and H9 human ES cells (Fig. 5*C*).

Global Gene Expression Profiles (High Reprogramming Cells versus Low Reprogramming Cells)—To explore the reprogramming mechanism(s) whose pathways and genes are significantly different between parental cells of high and low reprogramming cells, we examined global gene expression profiles of microarray data using z-score transformation (30). First, we analyzed previously reported reprogramming factors (2, 3, 8, 20, 31–35). Although expressions of *KLF4*, *KLF2*, *KLF5*, *MYC*, *TP53*, *TBX3*, and *CHD1* were detected in each parental cell, it is hard to show statistical differences of these gene expressions between the high and low reprogramming cells (supplemental Table S2). Expressions of other reprogramming factors (*OCT3/4*, *SOX* family, *KLF1*, *MYC* family, *NANOG*, *LIN28*, *UTF1*, *ESRRB*, *SALL4*, and *NR5A2*) were not detected in each parental cell (supplemental Table S2). Although significant difference was not found, the *KLF4* gene expressions

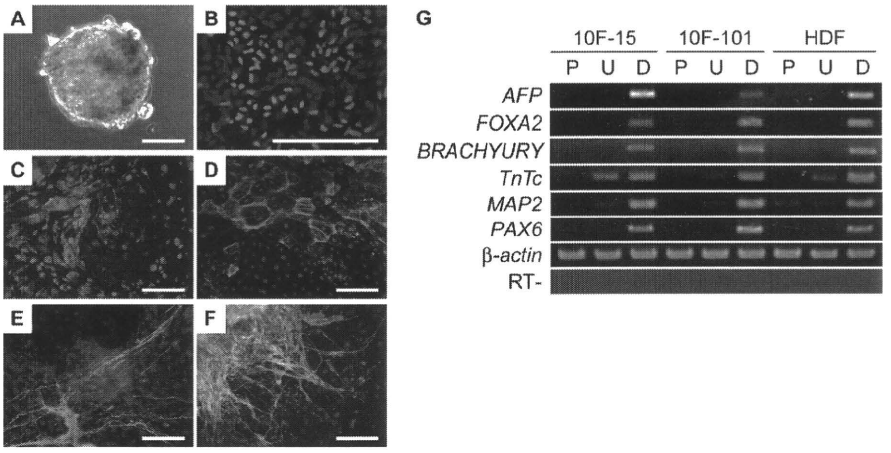

FIGURE 3. ***In vitro* differentiation of iPS cells from clonally expanded MSCs and HDF.** *A*, EB formation of iPS cells at day 8. *B–E*, immunocytochemistry for SOX17 (*B*), AFP (*green*) (*C*), desmin (*red*) (*D*), and α-smooth muscle actin βIII-tubulin (*E*). *F*, immunocytochemistry of differentiated iPS cells cultured on PA6 feeder cells at day 18. Tyrosine hydroxylase- and βIII-tubulin-positive cells are indicated as *green* and *red cells*, respectively. Nuclei were stained with Hoechst 33342 (*blue*). *Scale bars* = 100 μm. *G*, RT-PCR analysis of three germ layer marker genes. *P*, parental cell; *U*, undifferentiated iPS cells; *D*, differentiated iPS cells (EB).

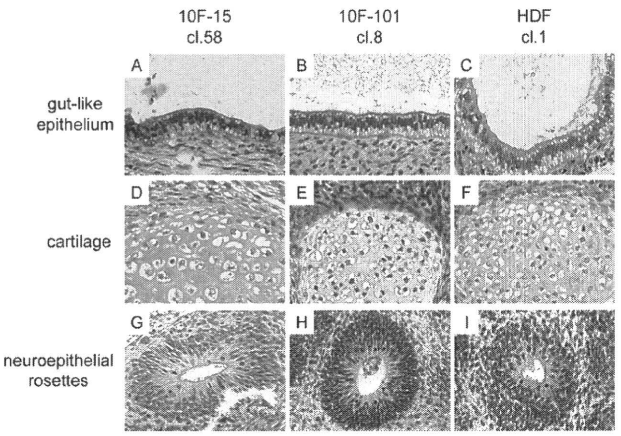

FIGURE 4. **Teratoma formation after *in* vivo implantation of iPS cells from clonally expanded MSCs and HDF.** *A–I*, hematoxylin and eosin staining of harvested tumors. Each tumor (teratoma) contained three embryonic germ layer tissues including gut-like epithelium (*A–C*: endoderm), cartilage (*D–F*: mesoderm), and neuroepithelial rosettes (*G–I*: ectoderm). *cl.58*, colony derived from 10F-15 clonal cells; *cl.8*, colony derived from 10F-101 clonal cells; *cl.1*, colony derived from HDF.

of high reprogramming cells (10F-15, 10F-5, and 10F-8) were lower than cells of low reprogramming cells (10F-101, 10F-107, and HDF). Furthermore, real-time PCR analysis showed the same tendency of the expressions of the *KLF4* gene (lower expressions in high reprogramming cells). Real-time PCR analysis also demonstrated the high expression of *TP53* in HDF cells but almost the same expression levels in other cells (supplemental Fig. S4).

We then sought unknown factor(s) that facilitate the reprogramming in high reprogramming cells. From comparison of microarray data that was normalized by z-score transformation, we picked out the genes that demonstrated statistically significant difference between the high reprogramming cells (10F-15, 10F-5, and 10F-8) and the low reprogramming cells (10F-101, 10F-107, and HDF). Up-regulated and down-regulated genes are listed in supplemental Table S3.

DISCUSSION

In our preliminary experiments, we examined whether or not MSCs from the human third molars could be reprogrammed by transduction of three transcription factors, *OCT3/4*, *SOX2*, and *KLF4*, using commercially available lentivirus. We generated iPS cells using the lentivirus; however, the efficiency was extremely low (data not shown). Therefore, we used pMXs retrovirus in this study. We transduced these three transcription factors using the retrovirus into culture-expanded but not clonally expanded MSCs from third molars of 10-, 13-, and 16-year-old donors. We detected the different efficiency of iPS generation. The most effective MSCs were from the 10-year-old donor (Fig. 1*A*). Because the MSCs were a heterogeneous cell population, we then clonally expanded the MSCs from the donor and selected five clones of 10F-5, 10F-8, 10F-15, 10F-101, and 10F-107. We found that proliferation ability was positively correlated with reprogramming efficiency (Fig. 1, *B*, *C*, and *D*), whereas viral transduction efficiency was similar (Fig. 1*E*). The clonally expanded MSCs of 10F-5, 10F-8, and 10F-15 could be efficiently reprogrammed at a rate that is ~30–100-fold higher than that of 10F-101, 10F-107, and HDF (Fig. 1*C*). In addition, 10F-15 reprogrammed 7-fold higher efficiency when compared with the non-cloned parental MSCs (Fig. 1, *A* and *C*). These results indicated that the parental MSCs from the 10-year-old donor contained more susceptible cells for iPS generation than other parental MSCs.

We tried to explore the difference among 10F-15, 10F-101, and HDF to determine the factor(s) affecting reprogramming promotion. One of the possibilities is that 10F-15 might have highly expressed reprogramming factor(s) in the same way as neural stem cells, which have highly expressed endogenous *Sox2*, can be easily reprogrammed (36). Furthermore, the family genes of *Sox2* (*Sox1*, *Sox3*, *Sox15*, and *Sox18*), *Klf4* (*Klf1*, *Klf2*, and *Klf5*), and *MYC* (*L-Myc* and *N-Myc*) can reprogram mouse embryonic fibroblasts to iPS cells (20). Recently, transduction of *UTF1* and/or introduction of *p53*-siRNA improved the efficiency of human iPS generation (8). Therefore, we analyzed these mRNA expression levels by DNA microarrays. The reprogramming factors, including *OCT3/4*, *SOX2*, *MYC*, *NANOG*, *LIN28*, *UTF1*, and *p53*, were not different in each parental cell (Fig. 5*A*). However, the expression level of *KLF4* was significantly lower (0.1-fold), and *KLF5* was slightly higher (3-fold) when compared with parental cells of 10F-101 and HDF (Fig. 5*A*). Other *SOX*, *KLF*, and *MYC* family genes did not differ in each parental cell (data not shown).

The data of low expression of *KLF4* in high reprogramming cells (10F-15, 10F-5, and 10F-8) (Fig. 5*A* and supplemental Fig. S4) was somehow confusing because the *KLF4* is a well

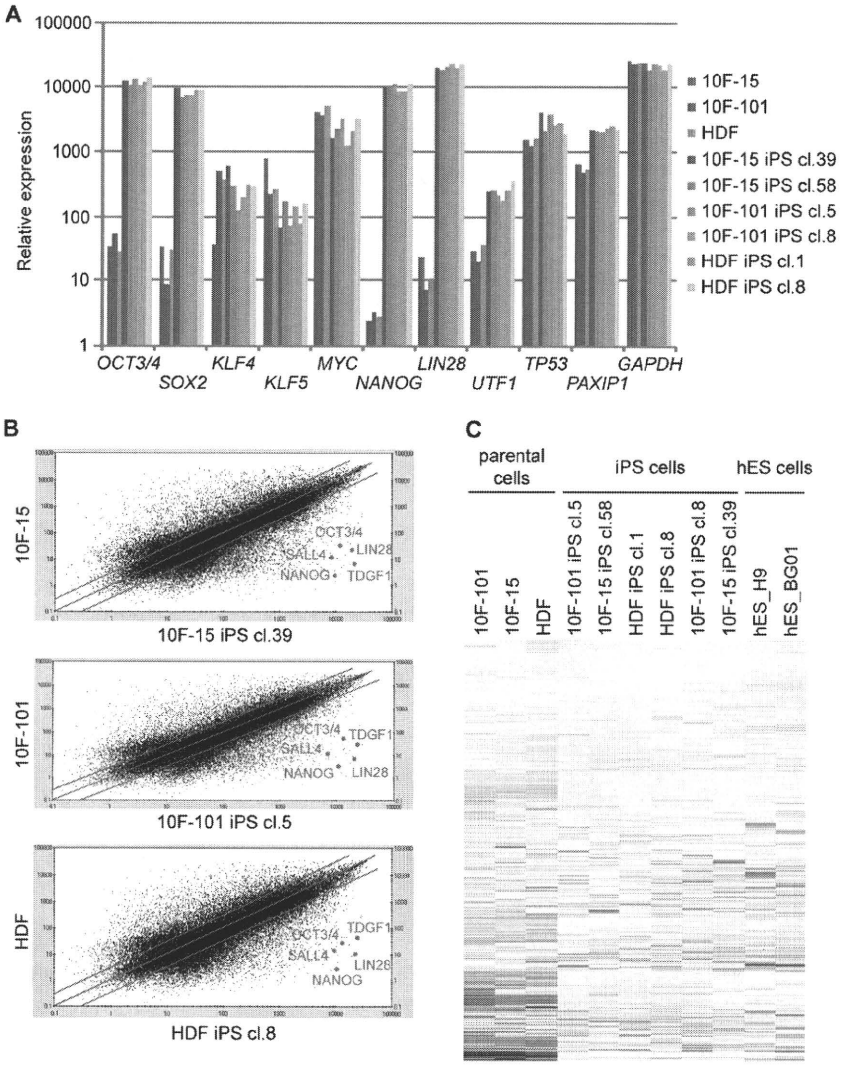

FIGURE 5. **Global gene expression analysis of iPS cells from clonally expanded MSCs and their parental cells.** *A*, expression levels of reprogramming genes in parental cells and iPS cells. *cl.39*, colony derived from 10F-15 clonal cells; *cl.58*, colony derived from 10F-15 clonal cells; *cl.5*, colony derived from 10F-101 clonal cells; *cl.8*, colony derived from 10F-101 clonal cells; *cl.1*, colony derived from HDF. *B*, global gene expression comparison between parental cell and iPS cell. Red circles indicate the expression levels of *OCT3/4*, *NANOG*, *SALL4*, *LIN28*, and *TDGF1*. *Red lines* indicate the equivalent and 3-fold differences between the two samples. *C*, comparison of gene expression profiles of parental cells, iPS cells, and human ES (*hES*) cells. *cl.8*, colony derived from HDF.

known factor to induce iPS cells. In this regard, it was reported that *KLF4* is highly expressed in growth-arrested cells and terminally differentiated cells (37, 38). 10F-5, 10F-8, and 10F-15 parental cells could differentiate into osteoblasts when cultured in osteogenic culture medium containing dexamethasone, whereas 10F-101 and 10F-107 parental cells could not. The data might indicate that the existence of abundant competent stem cells with low KLF4 expression in the third molar is feasible for iPS generation.

We attempted to identify the unknown factor(s) that facilitate the reprogramming in high reprogramming cells (10F-15, 10F-5, and 10F-8). As seen in supplemental Table S2, among these genes, we are interested in *PAXIP1* (also called *PTIP*) gene because *Ptip1*$^{-/-}$ mouse ES cells exhibited low expression of *Oct3/4*, whereas they exhibited high expression of trophectodermal marker *Hand1* (39). The *PAXIP1* gene was highly expressed in the high reprogramming cells detected by global gene expression profiles. PAXIP1 acts as component of a histone H3 lysine four (H3K4) methyltransferase complex (40, 41). It is known that histone acetylation and H3K4 methylation promote activation of transcription. One of the histone deacetylase inhibitors, valproic acid, accelerates reprogramming efficiency in mouse and human fibroblasts (7, 42). Recently, it was reported that H3K4 methylation is associated with efficient reprogramming of pluripotent gene (*Oct3/4*, *Sox2*, and *Sall4*) expression by mouse somatic cell nuclei transplantation into amphibian oocytes (43). These results might imply that the high expression of *PAXIP1* in the high reprogramming cells (10F-15, 10F-5, and 10F-8) enhanced the generation of iPS cells. Interestingly, PAXIP1 is also reported to have roles in DNA double-strand break repair (44). Similarly, INO80, SRCAP (also called SWR1), poly-(ADP-ribose)polymerase (PARP), and so on are known as other factors involved in both histone modification and DNA repair (45–47). Among them, we are particularly focusing on the PARP family (47), especially the *PARP-1* gene, because as seen in supplemental Table S2C, its expressions in all high reprogramming cells (10F-15, 10F-5, and 10F-8) were higher than those of low reprogramming cells.

As described, expressions of *PAXIP1* gene in high reprogramming cells were high and about 30% more than low reprogramming cells. Furthermore, we also detected 3–4 times more expression of *PAXIP1* in iPS cells when compared with each parental cell (Fig. 5*A*). This noticeable change (high expression after iPS cell induction) was also seen in *PARP-1* expression after iPS cell induction. This was confirmed at both the gene and the protein expression levels.[4] PARP-1 is the most abundant of PARP family members, accounting for >85% of nuclear PARP activity and modifying histone structure through NAD+-dependent "PARylation," *i.e.* poly(ADPribosyl)ation. We thus speculate that high iPS generation clones may be accessible for

[4] Y. Yoshimura, Y. Oda, H. Ohnishi, M. Tadokoro, Y. Katsube, M. Sasao, K. Hattori, S. Saito, K. Horimoto, S. Yuba, and H. Ohgushi, manuscript in preparation.

reprogramming factors due to feasible conformational change of chromatin by direct/indirect actions of chromatin modification proteins such as PAXIP1 and possibly PARP-1. Further extensive studies are needed to elucidate the mechanisms to promote reprogramming cascade.

In this study, we demonstrated that MSCs from human third molars can generate iPS cells by retroviral transduction of three transcription factors (*OCT3/4*, *SOX2*, and *KLF4*) without *MYC* or *LIN28*. The iPS cells were similar to human ES cells in many aspects, including morphology, surface markers and gene expression, DNA methylation states in the promoter regions, *in vitro* differentiation, and teratoma formation. Because human third molars are discarded as clinical waste and MSCs isolated from the molars could be frozen and stored for many years, MSCs are a useful cell source for the generation of iPS cells. Interestingly, even clonally expanded MSCs could be reprogrammed with high iPS generation efficiency. Consequently, the cells might be available for iPS generation by other methods using plasmids (48), chemicals, proteins (49, 50), and microRNAs (51) aiming for the clinical use of the iPS cells in regenerative medicine.

Acknowledgments—We warmly thank Drs. Shinya Yamanaka and Kazutoshi Takahashi for technical advice concerning the generation and maintenance of iPS cells and Dr. Toshio Kitamura for providing the retroviral system. We also thank members of our laboratory, especially Youichi Katoh, Hiroko Machida, and Dr. Takahiro Yagyuu, for technical assistance and valuable discussions.

REFERENCES

1. Takahashi, K., and Yamanaka, S. (2006) *Cell* **126,** 663–676
2. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007) *Cell* **131,** 861–872
3. Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I., and Thomson, J. A. (2007) *Science* **318,** 1917–1920
4. Park, I. H., Zhao, R., West, J. A., Yabuuchi, A., Huo, H., Ince, T. A., Lerou, P. H., Lensch, M. W., and Daley, G. Q. (2008) *Nature* **451,** 141–146
5. Lowry, W. E., Richter, L., Yachechko, R., Pyle, A. D., Tchieu, J., Sridharan, R., Clark, A. T., and Plath, K. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 2883–2888
6. Aasen, T., Raya, A., Barrero, M. J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., Tiscornia, G., Edel, M., Boué, S., and Izpisúa Belmonte, J. C. (2008) *Nat Biotechnol.* **26,** 1276–1284
7. Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D. A. (2008) *Nat. Biotechnol.* **26,** 1269–1275
8. Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z., Liu, Y., Yong, J., Zhang, P., Cai, J., Liu, M., Li, H., Li, Y., Qu, X., Cui, K., Zhang, W., Xiang, T., Wu, Y., Zhao, Y., Liu, C., Yu, C., Yuan, K., Lou, J., Ding, M., and Deng, H. (2008) *Cell Stem Cell* **3,** 475–479
9. Okita, K., Ichisaka, T., and Yamanaka, S. (2007) *Nature* **448,** 313–317
10. Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007) *Nature* **448,** 318–324
11. Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., Plath, K., and Hochedlinger, K. (2007) *Cell Stem Cell* **1,** 55–70
12. Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008) *Science* **321,** 699–702
13. Yamanaka, S. (2007) *Cell Stem Cell* **1,** 39–49
14. Dimos, J. T., Rodolfa, K. T., Niakan, K. K., Weisenthal, L. M., Mitsumoto, H., Chung, W., Croft, G. F., Saphier, G., Leibel, R., Goland, R., Wichterle, H., and Henderson, C. E., Eggan, K. (2008) *Science* **321,** 1218–1221
15. Park, I. H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M. W., Cowan, C., Hochedlinger, K., and Daley, G. Q. (2008) *Cell* **134,** 877–886
16. Ebert, A. D., Yu, J., Rose, F. F., Jr., Mattis, V. B., Lorson, C. L., Thomson, J. A., and Svendsen, C. N. (2009) *Nature* **457,** 277–280
17. Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J. I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M. (2003) *Science* **302,** 415–419
18. Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B. H., Wintergerst, U., Velez, M. C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F. D., Fischer, A., and Cavazzana-Calvo, M. (2008) *J. Clin. Invest.* **118,** 3132–3142
19. Cavazzana-Calvo, M., and Fischer, A. (2007) *J. Clin. Invest.* **117,** 1456–1465
20. Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008) *Nat. Biotechnol.* **26,** 101–106
21. Gronthos, S., Mankani, M., Brahim, J., Robey, P. G., and Shi, S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97,** 13625–13630
22. Miura, M., Gronthos, S., Zhao, M., Lu, B., Fisher, L. W., Robey, P. G., and Shi, S. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100,** 5807–5812
23. Ikeda, E., Yagi, K., Kojima, M., Yagyuu, T., Ohshima, A., Sobajima, S., Tadokoro, M., Katsube, Y., Isoda, K., Kondoh, M., Kawase, M., Go, M. J., Adachi, H., Yokota, Y., Kirita, T., and Ohgushi, H. (2008) *Differentiation* **76,** 495–505
24. Yan, X., Qin, H., Qu, C., Tuan, R. S., Shi, S., and Huang, G. T. (2010) *Stem Cells Dev.* **19,** 469–480
25. Viswanathan, S. R., Powers, J. T., Einhorn, W., Hoshida, Y., Ng, T. L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L. A., Lockhart, V. L., Shah, S. P., Tanwar, P. S., Mermel, C. H., Beroukhim, R., Azam, M., Teixeira, J., Meyerson, M., Hughes, T. P., Llovet, J. M., Radich, J., Mullighan, C. G., Golub, T. R., Sorensen, P. H., and Daley, G. Q. (2009) *Nat. Genet.* **41,** 843–848
26. Kitamura, T., Koshino, Y., Shibata, F., Oki, T., Nakajima, H., Nosaka, T., and Kumagai, H. (2003) *Exp. Hematol.* **31,** 1007–1014
27. Morita, S., Kojima, T., and Kitamura, T. (2000) *Gene Ther.* **7,** 1063–1066
28. Ikeda, E., Hirose, M., Kotobuki, N., Shimaoka, H., Tadokoro, M., Maeda, M., Hayashi, Y., Kirita, T., and Ohgushi, H. (2006) *Biochem. Biophys. Res. Commun.* **342,** 1257–1262
29. Kawasaki, H., Mizuseki, K., Nishikawa, S., Kaneko, S., Kuwana, Y., Nakanishi, S., Nishikawa, S. I., and Sasai, Y. (2000) *Neuron* **28,** 31–40
30. Cheadle, C., Vawter, M. P., Freed, W. J., and Becker, K. G. (2003) *J. Mol. Diagn.* **5,** 73–81
31. Feng, B., Jiang, J., Kraus, P., Ng, J. H., Heng, J. C., Chan, Y. S., Yaw, L. P., Zhang, W., Loh, Y. H., Han, J., Vega, V. B., Cacheux-Rataboul, V., Lim, B., Lufkin, T., and Ng, H. H. (2009) *Nat. Cell Biol.* **11,** 197–203
32. Tsubooka, N., Ichisaka, T., Okita, K., Takahashi, K., Nakagawa, M., and Yamanaka, S. (2009) *Genes Cells* **14,** 683–694
33. Heng, J. C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J. H., Orlov, Y. L., Huss, M., Yang, L., Lufkin, T., Lim, B., and Ng, H. H. (2010) *Cell Stem Cell* **6,** 167–174
34. Han, J., Yuan, P., Yang, H., Zhang, J., Soh, B. S., Li, P., Lim, S. L., Cao, S., Tay, J., Orlov, Y. L., Lufkin, T., Ng, H. H., Tam, W. L., and Lim, B. (2010) *Nature* **463,** 1096–1100
35. Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M. J., Heidersbach, A., Ramalho-Santos, J., McManus, M. T., Plath, K., Meshorer, E., and Ramalho-Santos, M. (2009) *Nature* **460,** 863–868
36. Shi, Y., Do, J. T., Desponts, C., Hahm, H. S., Schöler, H. R., and Ding, S. (2008) *Cell Stem Cell* **2,** 525–528
37. Shields, J. M., Christy, R. J., and Yang, V. W. (1996) *J. Biol. Chem.* **271,**

20009–20017
38. Conkright, M. D., Wani, M. A., Anderson, K. P., and Lingrel, J. B. (1999) *Nucleic Acids Res.* **27,** 1263–1270
39. Kim, D., Patel, S. R., Xiao, H., and Dressler, G. R. (2009) *Stem Cells* **27,** 1516–1523
40. Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M., and Ge, K. (2007) *J. Biol. Chem.* **282,** 20395–20406
41. Patel, S. R., Kim, D., Levitan, I., and Dressler, G. R. (2007) *Dev. Cell* **13,** 580–592
42. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A. E., and Melton, D. A. (2008) *Nat. Biotechnol.* **26,** 795–797
43. Murata, K., Kouzarides, T., Bannister, A. J., and Gurdon, J. B. (2010) *Epigenetics Chromatin* **3,** 4
44. van Attikum, H., and Gasser, S. M. (2009) *Trends Cell Biol.* **19,** 207–217
45. Pandita, T. K., and Richardson, C. (2009) *Nucleic Acids Res.* **37,** 1363–1377
46. Lilyestrom, W., van der Woerd, M. J., Clark, N., and Luger, K. (2010) *J. Mol. Biol.* **395,** 983–994
47. Rouleau, M., Patel, A., Hendzel, M. J., Kaufmann, S. H., and Poirier, G. G. (2010) *Nat. Rev. Cancer* **10,** 293–301
48. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008) *Science* **322,** 949–953
49. Zhou, H., Wu, S., Joo, J. Y., Zhu, S., Han, D. W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Schöler, H. R., Duan, L., and Ding, S. (2009) *Cell Stem Cell* **4,** 381–384
50. Kim, D., Kim, C. H., Moon, J. I., Chung, Y. G., Chang, M. Y., Han, B. S., Ko, S., Yang, E., Cha, K. Y., Lanza, R., and Kim, K. S. (2009) *Cell Stem Cell* **4,** 472–476
51. Judson, R. L., Babiarz, J. E., Venere, M., and Blelloch, R. (2009) *Nat. Biotechnol.* **27,** 459–461
52. Benjamini, Y., and Hochberg, Y. (1995) *J. Royal Stat. Soc. B* **57,** 289–300