A Cyclic Model of the Universe

Paul J. Steinhardt[1]* and Neil Turok[2]

We propose a cosmological model in which the universe undergoes an endless sequence of cosmic epochs that begin with a "bang" and end in a "crunch." Temperature and density at the transition remain finite. Instead of having an inflationary epoch, each cycle includes a period of slow accelerated expansion (as recently observed) followed by contraction that produces the homogeneity, flatness, and energy needed to begin the next cycle.

The current standard model of cosmology (*1–4*) combines the original big bang model and the inflationary scenario. Inflation, a brief period (10^{-30} s) of very rapid cosmic acceleration shortly after the big bang, can explain the homogeneity and isotropy of the universe on large scales (>100 Mpc), its spatial flatness, the distribution of galaxies, and the spatial fluctuations in the cosmic microwave background. However, the standard model has some cracks. The recent discoveries of cosmic acceleration and gravitationally self-repulsive dark energy (*5–8*) were not predicted and have no particular role in the standard model (*1–3*). Furthermore, the standard model does not explain the "beginning of time," the initial conditions of the universe, or what will happen in the long-term future.

Here, we present a cosmological model with an endless sequence of cycles of expansion and contraction. By definition, there is neither a beginning nor end of time, nor is there a need to define initial conditions. In addition, we explain the role of dark matter and generate the homogeneity, flatness, and density fluctuations without invoking inflation.

The cyclic aspect of the model is reminiscent of oscillatory models introduced in the 1930s based on a closed universe that undergoes a sequence of expansions, contractions, and bounces. The oscillatory models were constrained by having to pass through a singularity in which the energy and temperature diverge. Furthermore, as pointed out by Tolman (*9*, *10*), entropy produced during one cycle would add to the entropy produced in the next, causing each cycle to be longer than the one before it. Extrapolating backward in time, the universe would have to have originated at some finite time in the past so that the problem of explaining the "beginning of time" remains. Furthermore, today we know that the universe is flat, rather than closed, based on measurements of the cosmic microwave background anisotropy and large-scale structure.

In our cyclic model, the universe is infinite and flat, rather than finite and closed as in the oscillatory models. We introduce a negative potential energy rather than spatial curvature to cause the reversal from expansion to contraction. Before the reversal, though, the universe undergoes the usual period of radiation and matter domination, followed by a long period of accelerated expansion [presumably the acceleration that has been recently detected (*5–8*)]. The accelerated expansion, caused by dark energy, is necessary to dilute the entropy, black holes, and other debris produced in the previous cycle so that the universe is returned to its original pristine vacuum state before it begins to contract, bounce, and begin a cycle anew.

Essential Ingredients

As in inflationary cosmology, the cyclic scenario can be described in terms of the evolution of a scalar field ϕ along a potential $V(\phi)$ in a four-dimensional (4D) quantum field theory. The essential differences are the form of the potential and the couplings between the scalar field, matter, and radiation.

The analysis of the cyclic model follows from the action S that describes gravity, the scalar field ϕ, and the matter and radiation fluids:

$$S = \int d^4x \sqrt{-g} \left(\frac{1}{16\pi G} \mathcal{R} - \frac{1}{2} (\partial\phi)^2 - V(\phi) + \beta^4(\phi)(\rho_M + \rho_R) \right) \quad (1)$$

where g is the determinant of the metric $g_{\mu\nu}$, G is Newton's constant, and $\mathcal{R}$ is the Ricci scalar. The coupling $\beta(\phi)$ between ϕ and the matter (ρ_M) and radiation (ρ_R) densities is crucial because it causes the densities to remain finite at the big crunch/big bang transition.

The line element for a flat, homogeneous universe is $-dt^2 + a^2 d\mathbf{x}^2$, where a is the Robertson-Walker scale factor. The equations of motion following from Eq. 1 are

$$H^2 = \frac{8\pi G}{3} \left(\frac{1}{2}\dot{\phi}^2 + V + \beta^4 \rho_R + \beta^4 \rho_M \right) \quad (2)$$

$$\frac{\ddot{a}}{a} = -\frac{8\pi G}{3} \left(\dot{\phi}^2 - V + \beta^4 \rho_R + \frac{1}{2}\beta^4 \rho_M \right) \quad (3)$$

where a dot denotes a derivative with respect to t and $H \equiv /a$ is the Hubble parameter. The equation of motion for ϕ is

$$\ddot{\phi} + 3H\dot{\phi} = -V_{,\phi} - \beta_{,\phi}\beta^3 \rho_M \quad (4)$$

and the fluid equation of motion for matter (M) or radiation (R) is

$$\hat{a}\frac{d\rho_i}{d\hat{a}} = a\frac{\partial\rho_i}{\partial a} + \frac{\beta}{\beta_{,\phi}}\frac{\partial\rho_i}{\partial\phi_i} = -3(\rho_i + p_i), \qquad i = \mathrm{M,R} \quad (5)$$

where $\hat{a} \equiv a\beta(\phi)$, p_i is the pressure of the fluid component with energy density ρ_i, and ",ϕ" is $d/d\phi$. The implicit assumption is that matter and radiation couple to $\beta^2(\phi)g_{\mu\nu}$ (with scale factor $\hat{a}$) rather than the Einstein metric $g_{\mu\nu}$ alone (or the scale factor a). Note that the radiation term in Eq. 1 is actually independent of ϕ (because $\rho_R \propto \hat{a}^{-4}$) so only ρ_M enters the ϕ equation of motion.

We assume the potential $V(\phi)$ has the following features, as illustrated in Fig. 1: (i) $V(\phi)$ must approach zero rapidly as $\phi \to -\infty$; (ii) the potential must be negative for intermediate ϕ; and (iii) as ϕ increases, the potential must rise to a shallow plateau with a positive value V_0. An example of a potential with these properties is

$$V(\phi) = V_0 (1 - e^{-c\phi}) F(\phi) \quad (6)$$

where from this point onward we adopt units in which $8\pi G = 1$. $F(\phi)$ is a function we introduce to ensure that $V(\phi) \to 0$ as $\phi \to -\infty$. Without loss of generality, we take $F(\phi)$ to be nearly unity for ϕ to the right of potential minimum. The detailed manner in which it tends to zero at smaller ϕ is not crucial for the main predictions of the cyclic model. A quantitative analysis of this model potential (*11*) shows that a realistic cosmology can be obtained by choosing $c \geq 10$ and V_0 equal to today's dark-energy density ($\sim 6 \times 10^{-30}$ g/cm^3) in Eq. 6.

We have already mentioned that the coupling $\beta(\phi)$ is chosen so that $\hat{a}$ and, thus, the matter and radiation density are finite at $a = 0$. For example, we consider $\beta(\phi) \sim e^{-\phi/\sqrt{6}}$ as $\phi \to -\infty$. The presence of $\beta(\phi)$ and the consequent coupling of ϕ to nonrelativistic matter represent a modification of Einstein's theory of general relativity. Because the scalar field ϕ evolves by an exponentially small

[1]Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544, USA. [2]Department of Applied Mathematics and Theoretical Physics, Center for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK.

*To whom correspondence should be addressed. E-mail: steinh@princeton.edu

amount between nucleosynthesis ($t \sim 1$ s) and today ($t \sim 10^{17}$ s), the deviations from standard general relativity are small enough to easily satisfy all current cosmological constraints (*11*). However, the coupling of matter to ϕ produces other potentially measurable effects, including a fifth force that violates the equivalence principle. Provided $(\ln \beta)_{,\phi} \ll 10^{-3}$, for today's value of ϕ, these violations are too small to be detected (*11–13*). We shall assume this to be the case. Hence, the deviations from general relativity are negligible today.

The final crucial ingredient in the cyclic model is a matching rule that determines how to pass from the big crunch to the big bang. The transition occurs as $\phi \to -\infty$ and then rebounds toward positive ϕ. Motivated again by string theory (see below), we propose that some small fraction of the ϕ-field kinetic energy is converted to matter and radiation. The matching rule amounts to

$$\dot{\phi} e^{\sqrt{3/2}\phi} \rightarrow -(1 + \chi)\dot{\phi} e^{\sqrt{3/2}\phi} \quad (7)$$

where χ is a parameter measuring the efficiency of production of radiation at the bounce. Both sides of this relation are finite at the bounce.

Stringy Motivation

From the perspective of 4D quantum field theory, the introduction of a scalar field, a potential, and the couplings to matter in the cyclic model is no more arbitrary or tuned than the requirements for the inflationary models. However, the ingredients for the cyclic model are also strongly motivated by string theory and *M*-theory. This connection ties our scenario into the leading approach to fundamental physics and quantum gravity. The connection should not be overemphasized. String theory is not yet proven and quantum gravity effects are unimportant for describing cosmology at wavelengths much longer than a Planck length (10^{-33} cm). If the reader prefers, the connection to string theory can be ignored. On the other hand, we find the connection useful because it provides a natural geometric interpretation for the scenario. Hence, we briefly describe the relation.

According to *M*-theory, the universe consists of a 4D "bulk" space bounded by two 3D domain walls, known as "branes" (short for membranes), one with positive and the other with negative tension (*14–17*). The branes are free to move along the extra spatial dimension, so that they may approach and collide. The fundamental theory is formulated in 10 spatial dimensions, but 6 of the dimensions are compactified on a Calabi-Yau manifold, which for our purposes can be treated as fixed, and therefore ignored. Gravity acts throughout the 5D spacetime, but particles of our visible universe are constrained to move along one of the branes, sometimes called the visible brane. Particles on the other brane interact only through gravity with matter on the visible brane and hence behave like dark matter.

The scalar field ϕ is naturally identified with the field that determines the distance between branes. The potential $V(\phi)$ is the interbrane potential caused by nonperturbative virtual exchange of membranes between the boundaries. The interbrane force is what causes the branes to repeatedly collide and bounce. At large separation (corresponding to large ϕ), the force between the branes should become small, consistent with the flat plateau shown in Fig. 1. Collision corresponds to $\phi \to -\infty$. But the string coupling $g_s \propto e^{\gamma\phi}$, with $\gamma > 0$, so g_s approaches zero in this limit (*18*). Nonperturbative effects vanish faster than any power of g_s, for example, as or e^{-1/g_s^2}, or e^{-1/g_s}, accounting for the prefactor $F(\phi)$ in Eq. 6.

The coupling $\beta(\phi)$ also has a natural interpretation in the brane picture. Particles reside on the branes, which are embedded in an extra dimension whose size and warp are determined by β. The effective scale factor on the branes is $\hat{a} = a\beta(\phi)$, not a, and $\hat{a}$ is finite at the big crunch or big bang. The function $\beta(\phi)$ is in general different for the two branes (due to the warp factor) and for different reductions of M-theory. However, the standard Kaluza-Klein behavior $\beta(\phi) \sim e^{-\phi/\sqrt{6}}$ as $\phi \to -\infty$ is universal, because the warp factor becomes irrelevant as the branes approach one another (*11*, *18*).

Most importantly, the brane-world provides a natural resolution of the cosmic singularity (*11*, *18*). One might say that the big crunch is an illusion, because the scale factors on the branes ($\to \hat{a}$) are perfectly finite there. That is why the matter and radiation densities, and the Riemannian curvature on the branes, are finite. The only respect in which the big crunch is singular is that the one extra dimension separating the two branes momentarily disappears. Our scenario is built on the hypothesis (*19*) that the branes separate after collision, so the extra dimension immediately reappears. This process cannot be completely smooth, because the disappearance of the extra dimension is nonadiabatic and leads to particle production; that is, the brane collision is partially inelastic. Preliminary calculations of this effect are encouraging, because they indicate that a finite density of particles is produced (*20*). The matching condition, Eq. 7, parameterizes this effect. Ultimately, a well-controlled string-theoretic calculation (*11*, *18*, *20*) should determine the value of χ.

Dark Energy and the Cyclic Model

The role of dark energy in the cyclic scenario is novel. In the standard big bang and inflationary models, the recently discovered dark energy and cosmic acceleration (*5–8*) are an unexpected surprise with no obvious explanation. In the cyclic scenario, however, not only is the source of dark energy explained, but the dark energy and its associated cosmic acceleration are actually crucial to the consistency of the model; namely, the associated exponential expansion suppresses density perturbations and dilutes entropy, matter, and black holes to negligible levels. By periodically restoring the universe to an empty, smooth state, the acceleration causes the cyclic solution to be a stable attractor.

Right after a big bang, the scalar field ϕ is increasing rapidly. However, its motion is damped by the expansion of the universe, and ϕ essentially comes to rest in the radiation-dominated phase (stage 1 in Fig. 1). Thereafter it remains nearly fixed until the dark energy begins to dominate and cosmic acceleration commences. The positive potential energy density at the current value of ϕ acts as a form of quintessence (*21*, *22*), a time-varying energy component with negative pressure that causes the present-day accelerated expansion. This choice entails tuning V_0, but it is the same degree of tuning required in any cosmological model (including inflation)

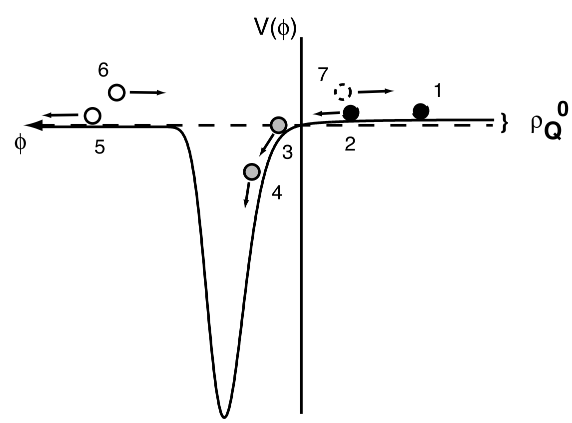

Fig 1. Schematic plot of the potential $V(\phi)$ as a function of the field ϕ. In M theory, ϕ determines the distance between branes, where $\phi \to -\infty$ as the branes collide. We define ϕ to be zero where $V(\phi)$ crosses zero and, therefore, ϕ is positive when the branes are at their maximal separation. Far to the right, the potential asymptotes to ρ_Q^0, the current value of the quintessence (dark energy) density. The solid circles represent the dark energy–dominated stage; the gray circles represent the contracting phase during which density fluctuations are generated; the open circles represent the phase when the scalar kinetic energy dominates; and the broken circle represents the stage when the universe is radiation dominated. Further details of the sequence of stages are described in the text.

to explain the recent observations of cosmic acceleration (*5–8*). In this case, because the dark energy serves several purposes, the single tuning resolves several problems at once.

The cosmic acceleration is nearly 100 orders of magnitude smaller than considered in inflationary cosmology. Nevertheless, if sustained for hundreds of e-folds (trillions of years) or more, the cosmic acceleration can flatten the universe and dilute the entropy, black holes, and other debris (neutron stars, neutrinos, and so forth) created over the preceding cycle, overcoming the obstacle that has blocked previous attempts at a cyclic universe. In this picture, we are presently about 14 billion years into the current cycle, and have just begun the trillions years of cosmic acceleration. After this amount of accelerated expansion, the number of particles in the universe may be suppressed to less than one per Hubble volume before the cosmic acceleration ends. Ultimately, the scalar field begins to roll back toward $-\infty$, driving the potential to zero. The scalar field ϕ is thus the source of the currently observed acceleration, the reason why the universe is homogeneous, isotropic, and flat before the big crunch, and the root cause for the universe reversing from expansion to contraction.

A Brief Tour of the Cyclic Universe

Putting together the various concepts that have been introduced, we can now present the sequence of events in each cycle beginning from the present epoch, stage 1 in Fig. 1. The universe has completed radiation- and matter-dominated epochs during which ϕ is nearly fixed. We are presently at the time when its potential energy begins to dominate, ushering in a period of slow cosmic acceleration lasting trillions of years or more, in which the matter, radiation, and black holes are diluted away and a smooth, empty, flat universe results. Very slowly the slope in the potential causes ϕ to roll in the negative direction, as indicated in stage 2. Cosmic acceleration continues until the field nears the point of zero potential energy, stage 3. The universe is dominated by the kinetic energy of ϕ, but expansion causes this to be damped. Eventually, the total energy (kinetic plus negative potential) reaches zero. From Eq. 2, the Hubble parameter is zero and the universe is momentarily static. From Eq. 3, $\ddot{a} < 0$, so that a begins to contract. So long as a is nearly static, the universe satisfies the ekpyrotic conditions for creating a scale-invariant spectrum of density perturbations (*19*, *23*). As the field continues to roll toward $-\infty$, a contracts and the kinetic energy of the scalar field grows; that is, gravitational energy is converted to scalar field kinetic energy during this part of the cycle. Hence, the field races past the minimum of the potential and off to $-\infty$, with kinetic energy becoming increasingly dominant as the bounce nears (stage 5). The scalar field diverges as a tends to zero. After the bounce, radiation is generated and the universe is expanding. At first, scalar kinetic energy density ($\propto 1/a^6$) dominates over the radiation ($\propto 1/a^4$) (stage 6). Soon after, however, the universe becomes radiation dominated (stage 7). The motion of ϕ is rapidly damped away, so that it remains close to its maximal value for the rest of the standard big bang evolution (the next 15 billion years). Then, the scalar field potential energy begins to dominate, and the field rolls towards $-\infty$, where the next big crunch occurs and the cycle begins anew.

Obtaining Scale-Invariant Perturbations

One of the most compelling successes of inflationary theory was to obtain a nearly scale-invariant spectrum of density fluctuations that can seed large-scale structure (*4*, *24–26*). Here, the same feat is achieved using different physics during an ultraslow contraction phase (stage 3 in Fig. 1) (*19*, *23*). In inflation, the density fluctuations are created by very rapid expansion, causing fluctuations on microscopic scales to be stretched to macroscopic scales (*4, 24–26*). In the cyclic model, the fluctuations are generated during a quasistatic, contracting universe where gravity plays no significant role (*19*). Simply because the potential $V(\phi)$ is decreasing more and more rapidly, quantum fluctuations in ϕ are amplified as the field evolves downhill (*19*, *27*, *28*). Instabilities in long-wavelength modes occur sooner than those in short-wavelength modes, thereby amplifying long-wavelength power and, curiously, nearly exactly mimicking the inflationary effect. The nearly scale-invariant spectrum of fluctuations in ϕ created during the contracting phase transform into a nearly scale-invariant spectrum of density fluctuations in the expanding phase (*23*). Current observations of large-scale structure and fluctuations of the cosmic microwave background cannot distinguish between inflation and the cyclic model because both predict a nearly scale-invariant spectrum of adiabatic, gaussian density perturbations.

Future measurements of gravitational waves may be able to distinguish between the two pictures (*19*). In inflation, where gravity is paramount, quantum fluctuations in all light degrees of freedom are subject to the same gravitational effect described above. Hence, not only is there a nearly scale-invariant spectrum of energy density perturbations, but there is also a scale-invariant spectrum of gravitational waves. In the cyclic and ekpyrotic models, where the potential, rather than gravity, is the cause of the fluctuations, the only field that obtains a nearly scale-invariant spectrum is the one rolling down the potential, namely ϕ, which only produces energy density fluctuations. The direct search for gravitational waves or the search for their indirect effect on the polarization of the cosmic microwave background (*29*, *30*) is the crucial test for distinguishing inflation from the cyclic model.

Cyclic Solution as Cosmic Attractor

Not only do cyclic solutions exist for a range of potentials and parameters, but they are also attractors for a range of initial conditions. The cosmic acceleration caused by the positive-potential plateau plays the critical role here. For example, suppose the scalar field is jostled and stops at a slightly different maximal value on the plateau compared with the exactly cyclic solution. The same sequence of stages ensues. The scalar field is critically damped during the exponentially expanding phase. So by the time the field reaches stage 3, where $V = 0$, it is rolling almost at the same rate as if it had started at $\phi = 0$, and memory of its initial position has been lost (*11*). The argument suggests that it is natural to expect dark energy and cosmic acceleration following matter domination in a cyclic universe, in accordance with what has been recently observed.

Comparing Cyclic and Inflationary Models

The cyclic and inflationary models have numerous conceptual differences in addition to those already described. Inflation requires two periods of cosmic acceleration: a hypothetical period of rapid expansion in the early universe and the observed current acceleration. The cyclic model requires only one period of acceleration per cycle.

In the inflationary picture, most of the volume of the universe is completely unlike what we see. Even when inflation ends in one region, such as our own, it continues in others. Because of the superluminal expansion rate of the remaining inflating regions, they occupy most of the physical volume of the universe. Regions that have stopped inflating, such as our region of the universe, represent an infinitesimal fraction. By contrast, the cyclic model is one in which the local universe is typical of the universe as a whole. All or almost all regions of the universe are undergoing the same sequence of cosmic events, and most of the time is spent in the radiation-, matter-, and dark energy–dominated phases.

In the production of perturbations, the inflationary mechanism relies on stretching modes whose wavelength is initially exponentially sub-Planckian to macroscopic scales. Quantum gravity effects in the initial state are highly uncertain, and inflationary predictions may therefore be highly sensitive to sub-Planckian physics. In contrast, perturbations in the cyclic model are generated when the modes have wavelengths of thousands of kilometers, using macroscopic physics insensitive to quantum gravity effects.

The cyclic model deals directly with the cosmic singularity, explaining it as a transition from a contracting to an expanding phase. Although inflation does not address the cosmic singularity problem directly, it does rely implicitly on the opposite assumption: that the big bang is the beginning of time and that the universe emerges in a rapidly expanding state. Inflating regions with high potential energy expand more rapidly and dominate the universe. If there is a preexisting contracting phase, then the high–potential energy regions collapse and disappear before the expansion phase begins. String theory or, more generally, quantum gravity can play an important role in settling the nature of the singularity and, thereby, distinguishing between the two assumptions.

The cyclic model is a complete model of cosmic history, whereas inflation is only a theory of cosmic history following an assumed initial creation event. Hence, the cyclic model has more explanatory and predictive power. For example, we have already emphasized how the cyclic model leads naturally to the prediction of quintessence and cosmic acceleration, explaining them as essential elements of an eternally repeating universe. The cyclic model is also inflexible with regard to its prediction of no long-wavelength gravitational waves.

Inflationary cosmology offers no information about the cosmological constant problem. The cyclic model provides a fascinating new twist. Historically, the problem is assumed to mean that one must explain why the vacuum energy of the ground state is zero. In the cyclic model, the vacuum energy of the true ground state is not zero. It is negative and its magnitude is large, as is apparent from Fig. 1. However, we have shown that the universe does not reach the true ground state. Instead, it hovers above the ground state from cycle to cycle, bouncing from one side of the potential well to the other but spending most time on the positive-energy side.

Reviewing the overall scenario and its implications, what is most remarkable is that the cyclic model can differ so much from the standard picture in terms of the origin of space and time and the sequence of cosmic events that lead to our current universe. Yet, the model requires no more assumptions or tunings (and by some measures less) to match the current observations. It appears that we now have two disparate possibilities: a universe with a definite beginning and a universe that is made and remade forever. The ultimate arbiter will be nature. Measurements of gravitational waves and the properties of dark energy (*11*) can provide decisive ways to discriminate between the two pictures observationally.

References and Notes

1. A. H. Guth, *Phys. Rev. D* **23**, 347 (1981).
2. A. D. Linde, *Phys. Lett. B* **108**, 389 (1982).
3. A. Albrecht, P. J. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982).
4. J. Bardeen, P. J. Steinhardt, M. S. Turner, *Phys. Rev. D* **28**, 679 (1983).
5. S. Perlmutter *et al.*, *Astrophys. J.* **517**, 565 (1999).
6. A. G. Riess *et al.*, *Astron. J.* **116**, 1009 (1998).
7. P. M. Garnavich *et al.*, *Astrophys. J.* **509**, 74 (1998).
8. N. Bahcall, J. P. Ostriker, S. Perlmutter, P. J. Steinhardt, *Science* **284**, 1481 (1999).
9. R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford Univ. Press, Clarendon Press, Oxford, UK, 1934).
10. See, for example, R. H. Dicke, J. P. E. Peebles, in *General Relativity: An Einstein Centenary Survey*, S. H. Hawking, W. Israel, Eds. (Cambridge Univ. Press, Cambridge, UK, 1979).
11. P. J. Steinhardt, N. Turok, *Phys. Rev. D*, in press (preprint available at http://arXiv.org/abs/hep-th/0111098).
12. T. Damour, A. Polyakov, *Nucl. Phys. B* **423**, 532 (1994).
13. T. Damour, *C. R. Acad. Sci.*, in press (preprint available at http://arXiv.org/abs/gr-qc/0109063).
14. J. Dai, R. G. Leigh, J. Polchinski, *Mod. Phys. Lett. A* **4**, 2073 (1989).
15. P. Horava, E. Witten, *Nucl. Phys. B* **460**, 506 (1996).
16. ———, *Nucl. Phys. B* **475**, 94 (1996).
17. J. Polchinski, *String Theory* (Cambridge Univ. Press, Cambridge, UK, 1998), vols. 1 and 2.
18. J. Khoury, B. A. Ovrut, N. Seiberg, P. J. Steinhardt, N. Turok, *Phys. Rev. D* **65**, 086007 (2002).
19. J. Khoury, B. A. Ovrut, P. J. Steinhardt, N. Turok, *Phys. Rev. D* **64**, 123522 (2001).
20. H. Liu, G. Moore, N. Seiberg, *High Energy Phys. Theory*, in press (preprint available at http://arXiv.org/abs/hep-th/0204168); A J. Tolley, N. Turok, *High Energy Phys. Theory*, in press (preprint available at http://arXiv.org/abs/hep-th/0204091).
21. R. R. Caldwell, R. Dave, P. J. Steinhardt, *Phys. Rev. Lett.* **80**, 1582 (1998); see also (*21*).
22. J. Ostriker, P. J. Steinhardt, *Sci. Am.* **284** (January 2001), p. 46.
23. J. Khoury, B. A. Ovrut, P. J. Steinhardt, N. Turok, *Phys. Rev. D* , in press (preprint available at http://arXiv.org/abs/hep-th/0109050).
24. A. H. Guth, S.-Y. Pi, *Phys. Rev. Lett.* **49**, 1110 (1982).
25. A. A. Starobinskii, *Phys. Lett. B* **117**, 175 (1982).
26. S. W. Hawking, *Phys. Lett. B* **115**, 295 (1982).
27. R. Kallosh, L. Kofman, A. Linde, *Phys. Rev. D* **64**, 123523 (2001).
28. J. Khoury, B. A. Ovrut, P. J. Steinhardt, N. Turok, *High Energy Phys. Theory*, in press (preprint available at http://arXiv.org/abs/hep-th/0105212).
29. R. Caldwell, M. Kamionkowski, *Sci. Am.* (January 2001), p. 46; for a recent discussion see (*30*).
30. A. Lewis, A. Challinor, N. Turok, *Phys. Rev. D* **65**, 023505 (2002).
31. We thank M. Bucher, S. Gratton, J. Khoury, B. A. Ovrut, J. Ostriker, P. J. E. Peebles, A. Polyakov, M. Rees, N. Seiberg, D. Spergel, A. Tolley, T. Wiseman, and E. Witten for useful conversations. We thank L. Rocher for pointing out historical references. This work was supported in part by U.S. Department of Energy grant DE-FG02-91ER40671 (P.J.S.) and by PPARC-UK (N.T.).

1 February 2002; accepted 10 April 2002
Published online 2 April 2002;
10.1126/science.1070462
Include this information when citing this paper.

Evidence of HIV-1 Adaptation to HLA-Restricted Immune Responses at a Population Level

Corey B. Moore,[1] Mina John,[1,2] Ian R. James,[1]
Frank T. Christiansen,[2,3] Campbell S. Witt,[1,2] Simon A. Mallal[1,2]*

Antigen-specific T cell immunity is HLA-restricted. Human immunodeficiency virus–type 1 (HIV-1) mutations that allow escape from host immune responses may therefore be HLA allele–specific. We analyzed HIV-1 reverse transcriptase sequences from a large HLA-diverse population of HIV-1–infected individuals. Polymorphisms in HIV-1 were most evident at sites of least functional or structural constraint and frequently were associated with particular host HLA class I alleles. Absence of polymorphism was also HLA allele–specific. At a population level, the degree of HLA-associated selection in viral sequence was predictive of viral load. These results support a fundamental role for HLA-restricted immune responses in driving and shaping HIV-1 evolution in vivo.

Selection of viral mutations associated with loss of antiviral cytotoxic T lymphocyte (CTL) responses has been described in humans with acute and chronic HIV-1 infection (*1*), macaques infected with simian immunodeficiency virus (SIV) (*2*, *3*), and rhesus monkeys challenged with simian-human immunodeficiency virus (SHIV) after vaccination (*4*). However, the full extent and importance of CTL escape mutation to HIV-1 evolution remains to be established. CTL escape mutations occur at critical sites within HLA-restricted CTL epitopes where an amino acid substitution may abrogate epitope-HLA binding, reduce T cell receptor recognition, or generate antagonistic CTL responses (*1*). Mutations that affect proteosome cleavage sites flanking CTL epitopes may also disrupt cellular processing of the epitope (*5*). The capacity of the virus to mutate at any amino acid residue is constrained, however,

[1]Centre for Clinical Immunology and Biomedical Statistics, Royal Perth Hospital and Murdoch University, Level 2 North Block, Royal Perth Hospital, Wellington Street, WA 6000, Australia. [2]Department of Clinical Immunology and Biochemical Genetics, Royal Perth Hospital. [3]Department of Pathology, University of Western Australia, WA 6009, Australia.

*To whom correspondence should be addressed. E-mail: mallal@prodigal.murdoch.edu.au