Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 4232–4235, April 1999

Perspective

Observing the epoch of galaxy formation

*Charles C. Steidel**

Palomar Observatory, California Institute of Technology, Caltech 105-24, Pasadena, CA 91125

Significant observational progress in addressing the question of the origin and early evolution of galaxies has been made in the past few years, allowing for direct comparison of the epoch when most of the stars in the universe were forming to prevailing theoretical models. There is currently broad consistency between theoretical expectations and the observations, but rapid improvement in the data will provide much more critical tests of theory in the coming years.

Recovering the entire history of galaxies, from their first gravitational collapse in the early universe to the rich variety and detail seen in the present-day universe, has long been a major goal of both theoretical and observational astronomy. There are essentially two approaches to this problem. The first is the "classical" approach, in which observations of the chemical content, kinematics, and spatial distribution of stars in our galaxy and in nearby galaxies are used to infer the sequence of events that produced what we see today. The second is to capitalize on the finite speed of light to use samples of galaxies at progressively larger distances as snapshots of the increasingly remote past. The latter approach has the advantage that it may be possible to observe directly the evolution of the galaxy "populations" and, with observations at the largest distances, actually catch incipient galaxies in the process of formation. As straightforward as this second option sounds in principle, there have been a number of practical barriers making the observations difficult, and a variety of subtleties that make the interpretation of observations nontrivial. Many of these observational barriers have been overcome in just the last few years, and it is on this progress that we focus; it is now possible to observe galaxies over >90% of the age of the universe, thanks to new telescopes, instruments, and techniques. A separate question is whether galaxy formation is understood; despite the frantic pace of observational progress, galaxy formation is perhaps one of the most complicated questions in astrophysics, and many aspects of it are likely to remain unsolved well into the next millennium.

Monolithic Collapse

There was a great deal of work on galaxy formation throughout the 1960s and 1970s. The emerging view was that large galaxies formed through a rapid gravitational collapse that would have resulted in huge bursts of star formation at early epochs (1, 2), leaving behind galaxies that had rapidly exhausted their fuel to age quiescently until the present; these galaxies appear now as large collections of old, red stars, such as elliptical galaxies and the central bulges of spiral galaxies. This formation scenario is often referred to as monolithic collapse. It inspired a great deal of work aimed at detecting the predicted "protogalaxies" during the past period when they were forming new stars at the rate of several hundred to several thousand solar masses per year. (The formation rate of stars in the Milky Way at the present time is only a few solar masses per year, a number typical of present-day spiral galaxies.) The searches for protogalaxies were remarkably unsuccessful, in general; either the cosmic epoch to which the observations were sensitive was not the "epoch of galaxy formation," or somehow the star-forming galaxies had escaped detection. A galaxy forming stars at the expected rate would have been easily detectable even in the days when photographic plates were the best available detectors, because the prodigious UV emission produced by massive young stars would be redshifted into the optical waveband where the observational techniques were most sensitive. However, in addition to stars, one of the products of huge bursts of star formation is dust, and dust is capable of obscuring UV photons from view. The energy that is absorbed by the dust is reradiated thermally in the far-IR part of the spectrum, and until very recently, the sensitivity for searching at these wavelengths was confined to the relatively nearby universe.

Cold Dark Matter

The modern theoretical framework for understanding galaxy formation arguably dates back only about 20 years (3); it is fundamentally different from the classical approach, which attempts to work backwards from the present time by using our understanding of stellar evolution and stellar dynamics. Instead, the modern picture works forward from prescribed initial conditions by using the physics of structure formation and attempts to understand galaxy formation as a natural consequence of the growth of mass fluctuations by gravitational instability. The sea change in the thinking about galaxy formation benefited from a growing appreciation that the universe is dominated by dark matter whose presence is inferred only through its gravitational interaction with visible matter and from improved understanding of initial conditions in the early universe, which produced the gravitational seeds from which structure would grow. Much of this insight came through improved limits on fluctuations in the cosmic microwave background. The hypothesis that gravity is responsible for producing structure on all scales made it especially straightforward to follow the evolution of what would become galaxies in an *ab initio* sense. Theoretical arguments and numerous observations suggest that initial conditions imposed in the early universe, together with gravitational instability, will result in a universe in which the smallest mass fluctuations collapse first and merge into progressively larger structures as time goes on.

The theoretical models that are favored currently are all variations on this general framework, which usually is referred to as "hierarchical structure formation" (see the article by Evrard in this series; ref. 4). The basic premise is that dark matter dominates the overall mass density and interacts only gravitationally with normal matter but ultimately dictates the underlying structure of all matter, determining where normal baryonic matter will end up and thus where visible galaxies will form. In hierarchical galaxy formation, the early history of the universe begins with a spectrum of fluctuations imposed as initial conditions by the physics of the very early universe. Gravity acts over time to alter this spectrum and to increase the amplitude of the fluctuations; large gravitationally bound density enhancements such as galaxies and

PNAS is available online at www.pnas.org.

*To whom reprint requests should be addressed. e-mail: ccs@astro.caltech.edu.

clusters of galaxies are made over a protracted period of time through successive merging or accretion of smaller lumps of matter. Within this framework, it is much easier to build small galaxies than large ones, and the characteristic mass of the largest gravitationally bound dark-matter "halos" is an increasing function of time, with the time scale to produce a given mass set by the details of the cosmological model, most notably the mean density of matter. The original version of the so called cold-dark-matter picture (5, 6), which assumed particular values of the cosmological density parameter Ω and the Hubble constant H_0, predicted rather late formation of large galaxies, a conclusion that did not seem to be supported by some observational studies of galaxy and galaxy-cluster evolution (e.g., refs. 7–10) and was met with a considerable degree of skepticism by much of the observational community. However, the hierarchical framework has become much more general over the last 10 years, and versions using the same notion of cold dark matter but more fashionable cosmologies (see the article by Kirshner in this series; ref. 11) allow for earlier formation of massive galaxies and galaxy clusters. These different "flavors" of cold dark-matter models begin to blur the perceived distinction between the classical picture of galaxy formation and a hierarchical universe; massive objects such as those that might be the antecedents of present-day bright galaxies are capable of forming at quite early epochs in some hierarchical models, and it is possible that galaxies formed and completed the bulk of their star formation on a relatively rapid time scale, a long time ago. The difference now is that this kind of formation scenario can make sense within a broad, physically motivated context that is supported by observations of the cosmic microwave background, by observations of large-scale structure in the distribution of galaxies, by inferences concerning dark matter on all scales, *and* by what we know from the stellar fossil record.

Thus, a modern theory for galaxy formation needs to specify the distribution of dark matter on all scales and the abundance of gravitationally collapsed halos as a function of mass and of cosmic time. These specifications now can be made routinely by using N body simulations (given a prescribed set of initial conditions) on supercomputers (see the article by Evrard in this series; ref. 4). In addition to treatment of the dark matter, a successful model needs to understand the physics of normal baryonic matter (which produces all of the luminous material that can be detected directly) within dark matter halos—how does it cool; how do stars form; what are the effects of the first generations of stars on subsequent star formation; and how do all of the processes relate to the "parent" dark-matter distribution? These are all exceedingly complicated processes that are not particularly well understood; the range of scales over which important physical processes are acting also makes a cosmological volume essentially impossible to simulate in detail. The root of the problem is that the stuff that can be modeled most easily is the dark matter, but what one can observe is the luminous matter. The details of how one maps to the other—how the luminous objects in the universe trace the underlying dark-matter distribution—is the key to understanding galaxy formation.

Observational Windows on Early Galaxy Formation

Recent observational progress has largely to do with making the early stages of the universe less mysterious by filling in some of the glaring gaps in the empirical information on the earliest observable galaxies. Broadly speaking, what has emerged over the last ≈3 years is an outline of the overall energy production by star formation for nearly the entire history of the universe (though not without ongoing controversy), the ability to discern the detailed morphologies (i.e., shapes and sizes) of distant galaxies with the refurbished Hubble Space Telescope, and the demonstration of efficient observational techniques that make feasible large-scale surveys of galaxies in the distant universe.

Star-Formation History of the Universe

The current understanding of the history of the integrated star-formation rate per unit volume of the universe is summarized in Fig. 1. Broadly speaking, although we do not yet know exactly when star formation began in galaxies (clearly, some time in the first 10% of the age of the universe or the first ≈1.4 billion years for the currently favored ≈14-billion-year-old universe), we know that we are living in a period when the universal star-formation rate is significantly smaller than in the past and that about 80% of the stars we see in the universe today probably formed in a broad time interval encompassing the first 40% of the age of the universe. At redshifts $z < 1$ or a lookback time of roughly 60% of the age of the universe ago, the data come from traditional "redshift surveys," in which galaxy evolution is studied based on spectroscopic samples of galaxies from a small region of the sky, down to a particular limiting flux. Such surveys become inefficient at larger lookback times, because the number of truly distant galaxies among the sample is dwarfed by less intrinsically luminous foreground objects and because, for practical reasons, it becomes difficult to measure spectroscopic redshifts because of a dearth of prominent spectral features that appear in the ground-based optical window of 0.3–0.9 μm. The points at earlier cosmic epochs ($z > 1$) come from application of a culling process that

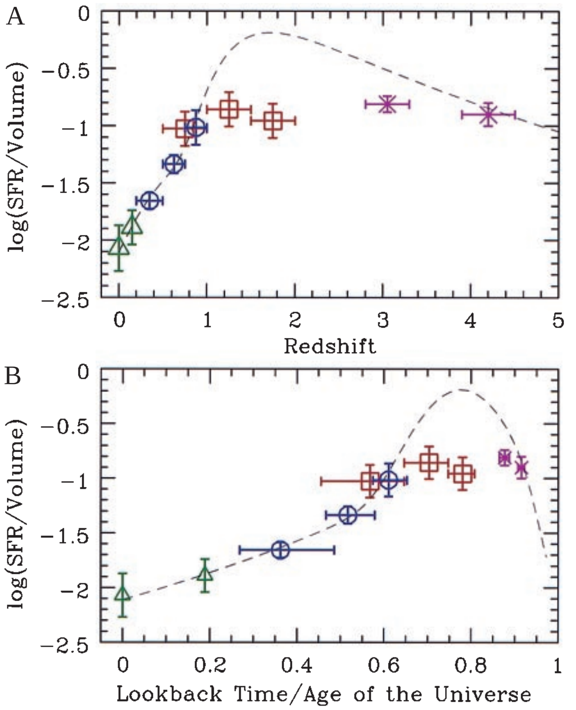

FIG. 1. A schematic diagram showing the current measurements of the star-formation history of the universe or, more properly, the energy emitted per unit volume, converted into a star formation rate (SFR; adapted from ref. 12). The same data are plotted in two different ways. (*A*) The SFR per unit volume is plotted against the redshift (the observable quantity). (*B*) The redshift is converted into a time axis and normalized to the age of the universe; the "lookback" time at the Big Bang is 1.0, such that observations now extend over ≈90% of the age of the universe. The data are from the references indicated: △ (13, 14), ○ (15), □ (16), and × (17). Essentially all of the data have been published in the last 3 years. The dashed curves represent one of many possible star-formation histories that is consistent with the current data on the far-IR backgrounds and with galaxies detected in the new submillimeter window (18). Note that the cosmic time interval beyond a redshift of $z \approx 4.5$ is so limited that it is unlikely that star formation at still higher redshifts can have a large effect on the integrated production of stars over the history of the universe. All of the points beyond $z = 1$ or a lookback time of ≈60% of the age of the universe are based on samples selected by using photometric methods, one of which is described in Fig. 2.

uses expected strong features in the spectra of galaxies (for the highest-redshift objects, the feature is the photoelectric absorption edge for hydrogen atoms, at a wavelength of 912 Å, which by virtue of the redshift can be observed from the ground at optical wavelengths for redshifts larger than 2.5; see Fig. 2) to isolate particular cosmic epochs photometrically—that is, based purely on images taken in broad-filter passbands, as illustrated in Fig. 2. After selecting the objects, they then can be followed up by confirming spectroscopic measurements, which provide much more accurate distances and more secure classification. Progress in the spectroscopy of these extremely faint galaxies (typically <10% of the brightness of the dark night sky) has been revolutionized thanks to the new generation of large (8- to 10-m aperture) ground-based optical/IR telescopes, of which the W. M. Keck telescopes in Hawaii are the first working examples.

The downside of this new color-selection technique is that it explicitly depends on the UV luminosity of galaxies, and as mentioned above, the UV output of a nascent galaxy can depend critically on how enshrouded in dust it might be. The points in Fig. 1 have been corrected for a moderate amount of extinction by dust, but they may not represent the whole story.

Obscured Star Formation at Early Epochs

Nearly in parallel with the high-redshift galaxy work focused on the rest-frame ultraviolet output of star-forming galaxies, there have been huge gains in understanding and appreciating the importance of obscured star formation at high redshift. One development has been the reliable measure of the integrated far-IR background through analysis of data obtained by the Cosmic Background Explorer satellite and by the Infrared Space Observatory (see ref. 19 and references therein), indicating that there is about as much energy emitted at these wavelengths as in the UV and optical when integrated over the entire history of the universe. Even more exciting have been the results of a new instrument that is capable of resolving the rest-frame far-IR emission from individual distant galaxies, called SCUBA (Submillimeter Common-User Bolometer Array; James Clerk Maxwell telescope, Mauna Kea, HI). This instrument, operating at an observed wavelength of 850 μm, has shown that there is a population of distant objects that are emitting the energy equivalent of star-formation rates of hundreds of solar masses per year, similar to that expected for classical protogalaxies. The problem at present is that there are few secure identifications and measured redshifts for submillimeter sources (the correct optical counterparts are usually somewhat ambiguous because of much lower spatial resolution achievable for the submillimeter observations), although it can be argued that most have $z > 1$. The full implications for the star-formation history of the universe await extensive spectroscopic identifications. It is not yet clear how much overlap there will be in objects selected by their far-IR emission and

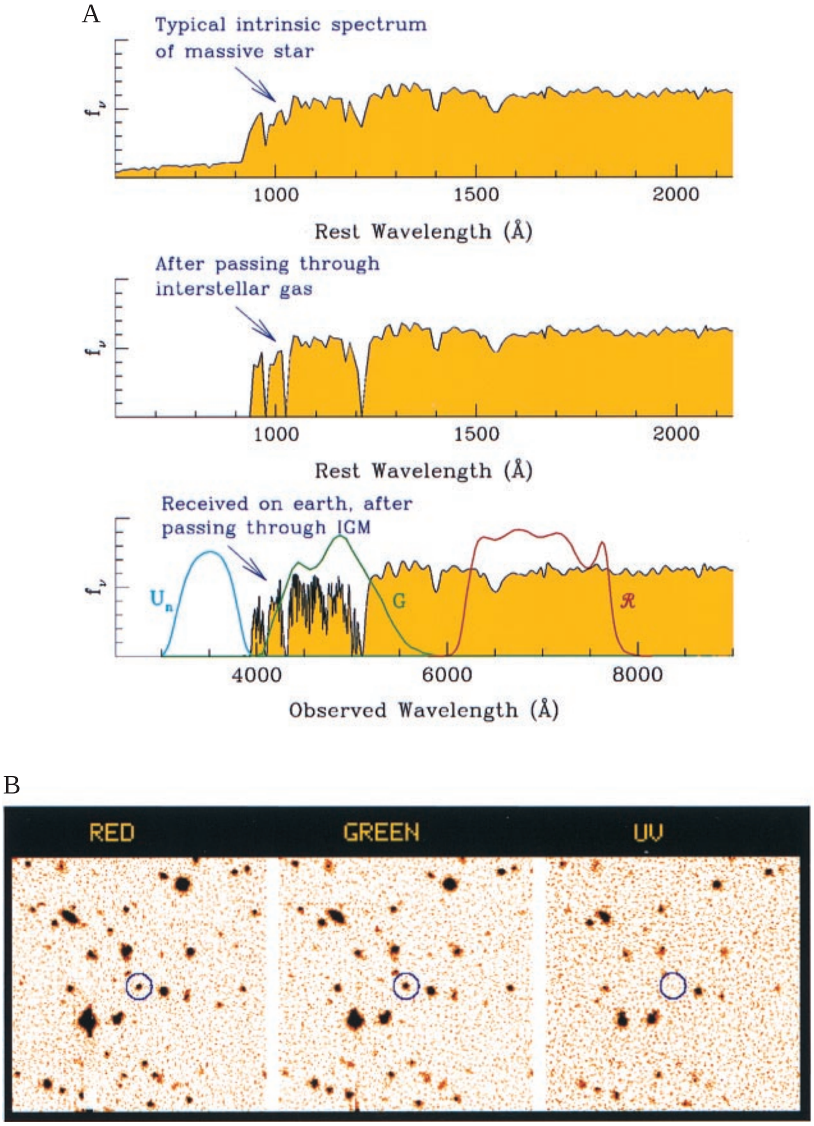

FIG. 2. An illustration of the so called "Lyman-break" or "UV-dropout" method for isolating galaxies at very high redshifts. This technique takes advantage of the sharp break expected at wavelengths <912 Å in the spectrum of a galaxy dominated (in the UV spectrum) by massive, young stars (*A Top*). The break is accentuated by the photoelectric absorption both in the galaxy hosting the stars (*A Middle*) and in the intervening intergalactic medium (*A Bottom*). The observed wavelength of the Lyman-break feature is found at $912(1 + z)$ Å, where z is the redshift. In the case illustrated here, the galaxy redshift is 3.15, bringing the feature into the optical window observable from the ground. Shown in *A Bottom* are broad-filter passbands that can be used effectively to isolate Lyman-break galaxies in the vicinity of redshift $z = 3$. (*B*) An example of small regions of charge-coupled device images taken through these filters. Note that the circled galaxy is seen clearly through the red and green filters, but it disappears completely through the UV filter. Only a few percent of all comparably faint galaxies will behave in this way, and such abrupt changes in spectral energy are not mimicked by anything except very distant galaxies. This procedure allows the efficient identification of large numbers galaxies in a prescribed range of redshifts at early epochs. The two highest redshift points in Fig. 1 were measured with this technique (17).

those selected in the UV; they are not necessarily mutually exclusive. These are very early but very exciting days for submillimeter-wave extragalactic astronomy (a completely new field emerging in only the last year), and the implications of the results are very much in a state of flux.

Large-Scale Distribution of Distant Galaxies

In addition to allowing a census of the total (relatively unobscured) star formation at early epochs, the new photometric techniques for isolating very high-redshift galaxies, combined with the ability to obtain spectra of extremely faint objects, also allow for the first systematic studies of large samples of very distant galaxies, and a first assessment of the clustering properties of early galaxies. Clustering is, in principle, a very powerful test of the modern theoretical paradigm for galaxy formation, as the models make specific predictions about what the larger-scale spatial distribution of galaxies of a given mass should be. A general prediction of hierarchical models is that the most massive, and presumably the most luminous, galaxies at early epochs should be very strongly clustered in space, as they are expected to trace quite unusual locations, many of which have become large clusters of galaxies by the present epoch (20). By extension, less massive objects should be less strongly clustered, in a predictable manner. The first observational results (21) strongly support both the strong clustering of the most (UV) luminous high-redshift galaxies and the dependence of clustering properties on UV luminosity (which at present seems to be a reasonable proxy for galactic mass). In addition, the structures that are in place at high redshift closely resemble incipient galaxy clusters as expected in most hierarchical-structure-formation models (22). Thus, the relationship of luminous galaxies to the expected dark-matter distribution seems to be relatively simple at these redshifts; numerical simulations that include only gravity can be made to resemble the observed universe by placing a single luminous galaxy within each gravitationally collapsed object and making the most massive objects host the brightest galaxies. Although the apparent agreement of simple models and the observations is not a proof of the general theoretical paradigm, it represents a significant success and an indication that our collective ideas about galaxy formation may well be on the right track. In general, the relative simplicity of the distant universe and the fact that different theories diverge most in their predictions at early epochs make it a very attractive laboratory for further testing of ideas about how galaxies actually form.

The Next Step

Knowledge of the rough dependence of the total star formation as a function of cosmic epoch and the ability to simulate the evolution of the structure of the dark matter under various flavors of hierarchical models are quite far from an understanding of the process of galaxy formation. It is likely that the focus of galaxy formation/evolution studies in the next decade will need to turn to the detailed astrophysics of individual objects, in a move away from population studies and global statistics (23). For example, a key pieces of missing information, required to tie together the observable luminous matter and the simulated dark matter, are measurements of the masses of distant galaxies. Such measurements, although extremely difficult to make, are essential for ultimately trying to understand how star formation works on galactic scales at early epochs and for being able to follow the evolution over time of a particular type of galaxy. In general, the most challenging aspect of deducing the history of galaxies is establishing evolutionary links between samples observed at different cosmic epochs; subtle details of galaxy morphologies from high-spatial-resolution studies, spectroscopic analyses of the chemical content of the stars and gas, and spectrophotometric observations over a very wide-wavelength baseline will all play important roles in this regard.

The health of general theoretical ideas about galaxy formation is quite robust at present, although the details are still very immature. For example, star formation is treated in a heuristic manner in most models (e.g., refs. 20 and 24), which use relatively simple scaling relations in lieu of a full treatment of the hydrodynamics of stars and gas. It is likely that huge improvements in the modeling can be expected as computing capabilities improve, but substantial progress also may require delving into a very complicated parameter space where the small-scale physical processes are not yet well understood. The largest gains are expected to be observational for the time being, and by 2005, there will be nearly 20 state-of-the-art large (8-m-class aperture) ground-based optical/IR telescopes in operation, and there are planned facilities that will exploit the new far-IR and submillimeter windows on galaxy formation (19). The uncovering of the universe of galaxies beyond the epoch that has been explored already and more detailed views of the currently accessible epochs will be made possible by the Next Generation Space Telescope, a large, IR-optimized telescope planned for 2007 or 2008. The next decade promises to bring a literal avalanche of data relevant to the question of galaxy formation. The challenge will be to make sense of it all.

1. Eggen, O. J., Lynden-Bell, D. & Sandage, A. R. (1962) *Astrophys. J.* **136,** 748–758.
2. Tinsley, B. M. & Larson, R. B. (1977) *Evolution of Galaxies and Stellar Populations* (Yale Univ. Observatory, New Haven, CT).
3. White, S. D. M. & Rees, M. J. (1978) *Mon. Notices R. Astron. Soc.* **183,** 341–358.
4. Evrard, A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 4228–4231.
5. Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. (1984) *Nature (London)* **311,** 517–525.
6. Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M. (1985) *Astrophys. J.* **292,** 371–394.
7. Bahcall, N. A. & Cen, R. (1992) *Astrophys. J.* **398,** L81–L84.
8. Lilly, S. J., Tresse, L., Hammer, F., Crampton, D. & Le Fèvre, O. (1995) *Astrophys. J.* **455,** 108–112.
9. Cowie, L. L., Songaila, A., Hu, E. M. & Cohen, J. G. (1996) *Astron. J.* **112,** 839–848.
10. Ellis, R. S. (1997) *Annu. Rev. Astron. Astrophys.* **35,** 389–443.
11. Kirshner, R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 4224–4227.
12. Madau, P., Ferguson, H. C., Dickinson, M., Giavalisco, M., Steidel, C. C. & Fruchter, A. (1996) *Mon. Notices R. Astron. Soc.* **283,** 1388–1404.
13. Gallego, J., Zamorano, J., Aragon-Salamanca, A. & Rego, M. (1995) *Astrophys. J.* **455,** L1–L4.
14. Tresse, L. & Maddox, S. J. (1998) *Astrophys. J.* **495,** 691–698.
15. Lilly, S. J., Le Fèvre, O., Hammer, F. & Crampton, D. (1996) *Astrophys. J.* **460,** L1–L4.
16. Connolly, A. J., Szalay, A. S., Dickinson, M., Subbarao, M. U. & Brunner, R. J. (1997) *Astrophys. J.* **486,** L11–L14.
17. Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M. & Pettini, M. (1999) *Astroph. J.*, in press.
18. Blain, A. W., Smail, I., Ivison, R. J. & Kneib, J.-P. (1999) *Mon. Notices R. Astron. Soc.* **302,** 632–642.
19. Blain, A. W. (1999) in *COSMOS-98: Particle Physics and the Early Universe* (American Institute of Physics, Woodbury, NY), in press.
20. Baugh, C. M., Cole, S. M., Frenk, C. S. & Lacey, C. G. (1998) *Astrophys. J.* **498,** 504–517.
21. Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., Pettini, M. & Kellogg, M. (1999) *Philos. Trans. R. Soc. London Ser. A* **357,** 153–166.
22. Governato, F., Baugh, C. M., Frenk, C. S., Cole, S., Lacey, C. G., Quinn, T. & Stadel, J. (1998) *Nature (London)* **392,** 359–361.
23. Ellis, R. S. (1998) *Nature (London)* **395**, Suppl., A3–A8.
24. Kauffmann, G., White, S. D. M. & Guiderdoni, B. (1993) *Mon. Notices R. Astron. Soc.* **264,** 201–210.