The Cosmic History of Star Formation

James S. Dunlop

Major advances in observational astronomy over the past 20 years have revolutionized our view of cosmic history, transforming our understanding of how the hot, smooth, early universe evolved into the complex and beautiful universe of stars and galaxies in which we now live. I describe how astronomers have used a range of complementary techniques to map out the rise and fall of star formation over 95% of cosmic time, back to the current observational frontier only ~500 million years after the Big Bang.

The cosmic history of star formation is our own history, or at least our prehistory. It is only through the lives and deaths of successive generations of stars that the atomic composition of the universe has been enriched (albeit only slightly) to contain atoms such as carbon, oxygen, and nitrogen, atoms that are essential for organic-based life. The past history of star-formation activity even affects today’s financial markets, with the seemingly ever-rising price of rare commodities such as gold being due, in large part, to the rarity and brevity of the violent supernova explosions in which all gold being due, in large part, to the rarity and brevity of the violent supernova explosions in which all
to suggest that the universe was once a much more violent place, with stars being formed at a much higher rate than is seen around us today. At the same time, we also know that, at very early times, in the so-called “dark ages,” there can have been no stars at all until the inferno following the Big Bang cooled to a temperature that allowed the first clouds of primordial gas to collapse. Charting the cosmic history of star-formation activity, from the first stars to the present day, has thus long been a fundamental goal of astronomy.

How to Measure Star-Formation History
Astronomers are fortunate in having four independent lines of evidence through which to map out the past history of star formation in galaxies.

We thus live in a fairly stable and peaceful corner of the universe, and indeed our entire Milky Way Galaxy of ~100 billion stars appears to be evolving gently and steadily, forming stars at the relatively modest rate of ~3 solar masses per year. However, there are many pieces of evidence to suggest that the universe was once a much more violent place, with stars being formed at a much higher rate than is seen around us today. At the same time, we also know that, at very early times, in the so-called “dark ages,” there can have been no stars at all until the inferno following the Big Bang cooled to a temperature that allowed the first clouds of primordial gas to collapse. Charting the cosmic history of star-formation activity, from the first stars to the present day, has thus long been a fundamental goal of astronomy.

How to Measure Star-Formation History
Astronomers are fortunate in having four independent lines of evidence through which to map out the past history of star formation in galaxies.

**References and Notes**

3. The parsec is a commonly used measure of distance in astronomy corresponding to the distance from the Sun to an object that has a parallax of 1 arc sec. This is equivalent to ~3.26 light years or 3.1 x 10^16 km.
4. The group of galaxies that includes the Milky Way and where all galaxies are bound together by gravity. It comprises around 50 small- to medium-sized dwarf galaxies, two large spiral galaxies (the Milky Way and M31), and two smaller spirals (the Large Magellanic Cloud and M33). It has a very irregular shape, but it is 2 x 10^17 parsecs at its broadest extent.

Acknowledgments: This was written while I was a visitor at l’Observatoire de la Cote d’Azur, and I am grateful for the support of the program host there. Supported by a VIGO grant from the Netherlands Organization for Scientific Research. I thank F. Fraternali, V. Hill, and E. Steffenburg for comments and discussion.

10.1126/science.1207392

**Fig. 1.** As shown by the Hubble Space Telescope image of the elliptical galaxy ESO 325-G004 (left), massive galaxies in the nearby universe are overwhelmingly dominated by old red/yellow stars formed more than 10 billion years ago and show little evidence of recent star-formation activity. By contrast, lower-mass galaxies generally continue to form stars in the present day, as shown by the young blue star clusters, red clumps of ionized hydrogen gas, and dark regions of cool dust and gas shown in the Hubble Space Telescope image of the spiral galaxy NGC 3982 (right).
First, like archaeologists, astronomers can examine the “fossil evidence” of past star-formation trends in the stellar populations of present-day galaxies. Detailed spectroscopic observations of the light from these relatively nearby galaxies are in principle straightforward, and the challenge is to dissect their integrated starlight into stellar populations of measurable mass, age, and chemical composition (so-called “metallicity”). This approach is especially powerful for probing recent star-formation activity but, like all archaeology, provides decreasing clarity with increasing look-back time. It also provides little direct information on where the various stellar populations found in present-day galaxies actually formed (i.e., in situ or in separate subunits that subsequently merged).

Second, unlike any other area of science, astronomers have the luxury of being able to look directly back in time, exploiting the finite speed of light to study the star-formation activity in galaxies at ever earlier times simply by looking at galaxies at ever greater distances. The relevance of this approach relies on an extension of the Copernican principle that we live in an ordinary region of the universe and that, averaged over sufficiently large scales, the universe is essentially the same everywhere. Only then can the observed behavior of “galaxies back then over there” be meaningfully related to “galaxies back then over here,” allowing us to infer what our own region of the universe may have looked like at comparable times. Provided one accepts this (observationally supported) assumption, the challenges are primarily technical due to the extreme faintness of the most distant galaxies, the progressive redshifting of their light with increasing distance [due to the expansion of the universe, observed wavelength = (1 + z) × emitted wavelength, where z = redshift], the potentially confusing effects of interstellar dust on the visibility of young stars, and issues over the best observational tracers of star-formation activity.

Third, again using observations of high-redshift galaxies to look back to earlier epochs, astronomers can use infrared observations to measure the stellar masses of galaxies and hence chart the buildup of stellar mass with cosmic time. Because most stars (like our Sun) are relatively long-lived, the global stellar mass density at a particular time should reflect the time integral of all preceding star-formation activity. To some extent, this measurement can therefore be viewed as simply a consistency check on the second method described above. However, it is in fact of more value than this, because the stellar masses of galaxies are dominated by large numbers of relatively low-mass long-lived stars, whereas the direct measurement of star-formation activity is, in practice, confined to observations of the most massive, luminous short-lived stars. Thus, comparison of the results from this and the previous method has the potential to provide information on the ratio of the numbers of low-mass to high-mass stars formed in star-forming regions (the so-called “Initial Mass Function”) and whether this has changed substantially over cosmic time.

Finally, it is possible to set constraints on the history of cosmic star formation by attempting to measure how the average chemical composition of the universe has changed over cosmic time. All elements heavier than hydrogen have been produced by nuclear fusion, and the very early hot universe certainly provided the necessary high temperatures and densities for fusion to take place. However, the rapid expansion and consequent cooling of the universe in the immediate aftermath of the Big Bang meant that nuclear fusion could only have been sustained for ~15 min, with the result that only the first stage of fusion, from hydrogen to helium, was properly completed before the universe entered the “dark ages.” The first stars, referred to as Population III stars, must therefore have formed from material with the so-called “primordial composition” of 75% hydrogen and 25% helium (by mass; 92% and 8% by atomic number density), with only minute trace amounts of light metals such as lithium. These first Population III stars did not last long (none have ever been discovered surviving to the present day). However, through internal nuclear fusion followed by supernova explosions, they must have commenced the process of chemical enrichment that, through recycling in the interstellar medium, produces the successive generations of increasingly metal-rich stars (so called Population II, and then Population I, such as our own Sun). From sensitive spectroscopic observations of known emission and absorption lines from different elements in galaxies and more diffuse regions of the universe at various redshifts, it is possible to construct the cosmic history of chemical enrichment, which then sets constraints on the history of star formation.

The Fossil Evidence

Over the past decade there have been several major studies of the stellar populations in the present-day nearby galaxy population. A key advance has been the creation of very large spectroscopic surveys, in particular the Sloan Digital Sky Survey (SDSS) (1), which has now provided a public release of fully calibrated optical spectra for ~1 million galaxies in the local universe. The SDSS covers a large enough volume of the local universe for the results to be regarded as representative of the universe of galaxies in the present day. With such a large, complete, and representative galaxy sample, it is possible to explore how, for example, galaxy age depends on mass, size, or morphological type. In addition, because the consistent and well-calibrated SDSS data have now been made public, different groups of researchers have been able to undertake different independent analyses and openly explore the robustness and accuracy of the derived results (2–5).

Given such a vast and rich data set, the challenge is to maximize the information that can be reliably extracted from the spectral database. This is done by comparing the data with the predictions of computer-generated models of what the integrated spectra of stellar populations of different ages, initial mass functions, and metallicity should look like as a function of age. A lot of effort has been invested in the development of such models over the last ~30 years (6–8), and they rely on an accurate theoretical description of how a star of given mass and chemical composition will evolve over time (9) combined with an accurate prediction of precisely what spectrum of light will be produced by the stellar atmosphere of a given star at a given stage in its life. To tackle this latter issue, theoretical predictions can, at least for a subset of stars, be cross-checked with detailed spectroscopy of nearby stars (10).

The models are certainly not yet perfect, and even with clever and efficient data analysis techniques (4) there are limitations to how robustly the true star-formation history of a galaxy can be determined from the final integrated spectrum of all its constituent stellar populations. Nevertheless, some clear and unambiguous trends have been established by this work.

First, it is clear that the most massive (generally elliptical) galaxies are the oldest, and the simplest, with generally very little recent star formation (4, 11) and spectra that can be described...
by between one and three distinct stellar populations (12). By contrast, galaxies with more typical masses have younger average stellar ages and more complicated star-formation histories, with as many as five different stellar populations being required to fit the SDSS spectra (Fig. 1). Second, integrating over the entire galaxy population, overall star-formation density (i.e., star formation per unit volume in the universe) is inferred to have declined monotonically since redshift \( z \approx 2 \), when the universe was -3 billion years old, at which point this approach to deducing cosmic star-formation history basically runs out of steam.

These are not all new results [for example, there has long been evidence that star formation has moved from the most massive to the least massive systems over cosmic time (13)], but this work has clarified and robustly quantified the key trends that must be explained by any successful model of galaxy formation and evolution. For example, the lack of substantial present-day star formation in massive galaxies has led theorists to invoke feedback from the central super-massive black hole to explain how residual gas in and around the most massive galaxies can be prevented from cooling and forming yet more stars (14).

**Looking Back in Time**

As already mentioned, astronomers are not just stuck with the fossil evidence, and the past 20 years have seen a veritable explosion in the detection and study of galaxies at ever greater distances, which are hence viewed at ever earlier times. This work has been technologically driven, both by the advent of giant (8- to 10-m-diameter) ground-based telescopes and by a series of spectacularly successful new space observatories (especially the Hubble Space Telescope, the Spitzer Space Telescope, and the Herschel Space Observatory) that, free from the effects of Earth’s atmosphere, have facilitated sensitive observations over the infrared region of the electromagnetic spectrum.

A multifrequency, multifacility approach is important in this work because there are a number of different probes of the level of star-formation activity in distant galaxies, and ideally one would like to be able to exploit these at all distances and epochs, despite the progressive redshifting of the light as one looks back to the earliest galaxies. It has also become clear that, because young stars are born in clouds of gas and dust, much of the blue or ultraviolet light produced by hot young stars is in fact often absorbed by interstellar dust and then re-emitted at much longer (infrared) wavelengths. Thus a pan-chromatic view is required to obtain a complete census of the “action.”

Observationally and theoretically, it is clear (and unsurprising) that stars form over a wide range of masses. However, although the precise shape of the initial mass function is a subject of continual study and debate (15-17), it is the relatively rare, most-massive stars that, in practice, provide the most useful tracers of star-formation activity, both because they shine only briefly (i.e., for a few million years) and because they are so bright and hot.

Young massive stars provide a number of useful indicators (at different wavelengths) of the level of star-formation activity (18, 19). These include bright optical-ultraviolet continuum light [from the stars themselves (20, 21)], bright hydrogen and oxygen emission lines [from the surrounding gas ionized by the hot stars (22-24)], enhanced emission at mid- and far-infrared wavelengths [from dust warmed by the ultraviolet light from the stars (25-27)], and radio emission [from relativistic electrons accelerated by the shock waves produced by the supernova explosions that mark the deaths of these short-lived massive objects (29)].

Various authors have discussed the pros and cons of each of these star-formation indicators and assessed the prospects for combining different measures to produce a complete, unbiased, and consistent measurement of star-formation activity over the widest possible range of redshifts (29-33). This is work in progress but, nevertheless, in recent years a consistent picture has emerged, at least out to redshift \( z \approx 2 \). As illustrated in Fig. 2, essentially all tracers of star-formation activity indicate that the star-formation rate (per unit volume) in the universe was an order of magnitude greater at \( z \approx 1 \) than in the present day (34, 20, 35) and that star-formation density stays at comparable or even higher levels out to at least redshift \( z \approx 2 \) (24, 36-38), with a plausible peak at \( z \approx 3 \) (39). Notably, this basic form of evolution over the past 12 billion years is in excellent agreement with that derived from the latest analyses of the fossil evidence.

At higher redshifts the picture becomes less clear, in part because the galaxies are more distant and hence fainter, but also because the most easily accessible star-formation indicators are ultraviolet continuum and emission lines (i.e., Lyman-alpha), which, although helpfully redshifted into the optical wavebands, are also the most susceptible to the highly uncertain effects of dust extinction (40, 41). Nevertheless, in the past ~10 years, enormous strides have been made in the discovery and study of large samples of galaxies at \( z \approx 3 \) to 6 (42-44), and the ultraviolet colors of the galaxies themselves can be used to make plausible corrections for dust extinction. At present, the available evidence suggests that beyond \( z \approx 3 \), as we look back into the first 2 billion years of cosmic history, star-formation density declines gradually but steadily out to at least \( z \approx 6 \), when the universe was ~1 billion years old (21, 45). At extreme redshifts, \( z > 7 \), even the ultraviolet light from young stars is redshifted out of the optical and into the infrared. As a result, the effective study of the first billion years of star-formation activity has had to await the advent of sensitive near-infrared imaging on the Hubble Space Telescope.

**The Buildup of Mass and Metals**

Sustained star formation over the majority of cosmic time must inevitably produce a gradual
buildups in the total mass of stars in the universe. In recent years, improvements in near- and mid-infrared observational facilities have made it possible to estimate the stellar masses of galaxies out to the highest redshifts, hence enabling a direct check on whether the whole picture hangs together (46, 47). The answer seems to be that there is excellent agreement out to \(z \approx 1\), but at higher redshifts some researchers have argued that there do not seem to be as many stars in place as would be expected, given the integrated total of all the observed preexisting star-formation activity (after accounting for mass loss caused by stellar evaporation processes such as supernovae and stellar winds (49)). It would be a mistake to exaggerate the severity of this discrepancy [the numbers are only in disagreement by a factor \( \sim 2 \) to 3, or arguably less (49)], but the direction of disagreement suggests that either the level of star-formation activity in the young universe has been overestimated (e.g., by excessive upward corrections for dust extinction) or the integrated stellar masses of all galaxies at each epoch have been systematically underestimated. Both of these options are still possible (49), as is the more speculative possibility of time evolution of the stellar initial mass function (50, 51). However, the direction of this tension certainly makes it hard to argue that the star-formation density at early times has been seriously underestimated, thus reinforcing the argument for a gradual decline in universal star-formation activity beyond \(z \sim 5\) (Fig. 3). Broadly speaking, the data all suggest that about half the stars in the present-day universe were in place by \(z \sim 2\), when the universe was\( \sim 3\) billion years old.

A second long-lived legacy of preceding star-formation activity is the abundance of the heavier chemical elements. Because they are only formed in massive stars that live for less than 10 million years, the growth in the cosmic abundance of elements such as carbon or iron should provide a fairly prompt (in cosmological terms) indicator of recent star-formation activity, delayed only by the time scale for expulsion of the newly formed atomic nuclei into the wider interstellar medium by supernovae explosions. However, conducting a full observational census of the cosmic budget of chemical elements at all epochs is extremely challenging because the “metals” (i.e., any element heavier than helium) can hide in different places at different redshifts (52). Specifically, in the present day most of the heavier elements are now either locked up in stars and planets (like our Earth) or are found in the hot-gas phase between galaxies. By contrast, when the universe was young, those heavy elements that had already been produced almost certainly resided primarily in cool gas and dust grains within galaxies (53). In addition, heavy element abundance is also a strong function of environment; it appears that massive galaxies are, and always have been, more metal rich than their lower-mass counterparts (54). Still, the search has now been conducted with sufficient thoroughness that astronomers are now reassured that the measured abundances of the elements at \(z \sim 2\) to 3 are in agreement with that expected from the integrated star-formation activity at earlier epochs (55). Interestingly, there now appears to be evidence of a relatively rapid downturn in carbon abundance as we look back toward \(z \sim 6\), suggestive of an extremely rapid buildup of stars in the immediately preceding \(\sim 500,000\) million years (55).

Epilogue: Searching for the First Galaxies and Stars

The recent refurbishment of the Hubble Space Telescope with the sensitive near-infrared Wide Field Camera 3 (WFC3) has enabled the first detections of star-forming galaxies at redshifts \(z \gtrsim 7\) (Fig. 3 (56–59)), providing a first glimpse into the “epoch of reionization,” the \(\sim 600\)-million-year era in which a range of circumstantial evidence indicates the first stars and galaxies switched on and reionized the previously cold, neutral, and dark universe (60).

We now know that stars and galaxies existed at \(z \sim 8.5\), extending our study of cosmic star-formation history back to within 500 million years of the Big Bang (56, 61, 62), and a lot of work is currently being expended in trying to determine whether there are enough young galaxies at these early times to explain cosmic reionization (63). The very first galaxies are of course, by definition, expected to contain very young stellar populations of very low metallicity, and claims that some of the galaxies at \(z \gtrsim 7\) discovered with WFC3 are exceptionally blue have sparked an ongoing and lively debate as to whether such primitive stellar populations have indeed now been observed (64–66). It remains to be seen whether the next generation of more powerful and sensitive new facilities, in particular the James Webb Space Telescope, can discover the elusive first generation of Population III stars, thus completing our journey through the cosmic history of star formation.

References and Notes