

# THE MASS-METALLICITY RELATION AT $z \gtrsim 2$ <sup>1</sup>

DAWN K. ERB<sup>2</sup>

Astronomy Option, California Institute of Technology, MS 105-24, Pasadena, CA 91125

ALICE E. SHAPLEY<sup>3</sup>

Department of Astronomy, University of California, 601 Campbell Hall, Berkeley, CA 94720

MAX PETTINI

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

CHARLES C. STEIDEL AND NAVEEN A. REDDY

Astronomy Option, California Institute of Technology, MS 105-24, Pasadena, CA 91125

AND

KURT L. ADELBERGER

Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101

Received 2005 October 12; accepted 2006 February 17

## ABSTRACT

We use a sample of 87 rest-frame UV-selected star-forming galaxies with mean spectroscopic redshift  $\langle z \rangle = 2.26 \pm 0.17$  to study the correlation between metallicity and stellar mass at high redshift. Using stellar masses determined from SED fitting to observed  $0.3\text{--}8\ \mu\text{m}$  photometry, we divide the sample into six bins in stellar mass and construct six composite  $\text{H}\alpha + [\text{N II}]$  spectra from all of the objects in each bin. We estimate the mean oxygen abundance in each bin from the  $[\text{N II}]/\text{H}\alpha$  ratio and find a monotonic increase in metallicity with increasing stellar mass, from  $12 + \log(\text{O}/\text{H}) < 8.2$  for galaxies with  $\langle M_\star \rangle = 2.7 \times 10^9 M_\odot$  to  $12 + \log(\text{O}/\text{H}) = 8.6$  for galaxies with  $\langle M_\star \rangle = 1.0 \times 10^{11} M_\odot$ . We use the empirical relation between SFR density and gas density to estimate the gas fractions of the galaxies, finding an increase in gas fraction with decreasing stellar mass. These gas fractions, combined with the observed metallicities, allow the estimation of the effective yield  $y_{\text{eff}}$  as a function of stellar mass; in contrast to observations in the local universe, which show a decrease in  $y_{\text{eff}}$  with decreasing baryonic mass, we find a slight increase. Such a variation of metallicity with gas fraction is best fitted by a model with supersolar yield and an outflow rate  $\sim 4$  times higher than the SFR. We conclude that the mass-metallicity relation at high redshift is driven by the increase in metallicity as the gas fraction decreases through star formation and is likely modulated by metal loss from strong outflows in galaxies of all masses.

*Subject headings:* galaxies: abundances — galaxies: evolution — galaxies: high-redshift

*Online material:* color figures

## 1. INTRODUCTION

Correlations between mass and metallicity or luminosity and metallicity are well established in nearby galaxies, ranging over orders of magnitude in mass and luminosity and spanning  $\sim 2$  dex in chemical abundance. Lequeux et al. (1979) first observed a correlation between heavy-element abundance and the total mass of galaxies; since then, most investigations have focused on the metallicity-luminosity relationship (e.g., Skillman et al. 1989; Zaritsky et al. 1994; Garnett et al. 1997; Lamareille et al. 2004; Salzer et al. 2005), although others have studied correlations between metallicity and rotational velocity (Zaritsky et al. 1994; Garnett 2002). The relationship between metallicity and stellar mass has recently been quantified by Tremonti et al. (2004, hereafter T04), using a sample of  $\sim 53,000$  galaxies from the Sloan Digital Sky Survey (SDSS). Such correlations provide insight

into the process of galaxy evolution, as they allow the study of the history of star formation and gas enrichment or depletion through current, observable properties. Chemical enrichment is a record of star formation history, modulated by inflows and outflows of gas, while the stellar mass provides a more straightforward measure of the accumulated conversion of gas into stars and thus of the metals returned to the gas as the by-products of star formation.

It has long been recognized that a correlation between stellar mass and gas-phase metallicity is a natural consequence of the conversion of gas into stars in a closed system (van den Bergh 1962; Schmidt 1963; Searle & Sargent 1972). The principal ingredients of this “simple” or “closed-box” model are the metallicity, the yield from star formation (defined as the mass of metals produced and ejected by star formation, in units of the mass that remains locked in long-lived stars and remnants), and the gas fraction. If there are no inflows or outflows of gas, the metallicity is a simple function of the yield and the gas fraction and rises as the gas is converted into stars and enriched by star formation according to the yield. The model is subject to the further assumptions that the system is well mixed at all times, that it begins as pure gas with primordial abundances, that stellar evolution and nucleosynthesis take place instantaneously compared to the timescale of galactic evolution (the instantaneous

<sup>1</sup> Based on data obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA and was made possible by the generous financial support of the W. M. Keck Foundation.

<sup>2</sup> Current address: Harvard-Smithsonian Center for Astrophysics, MS 20, 60 Garden Street, Cambridge, MA 02138; derb@cfa.harvard.edu.

<sup>3</sup> Current address: Department of Astrophysical Sciences, Princeton University, Peyton Hall, Ivy Lane, Princeton, NJ 08544.

recycling approximation), and that the initial mass function (IMF) and the yield in primary elements of stars of a given mass are constant. This simple model, in combination with the observation that lower mass galaxies tend to have higher gas fractions (McGaugh & de Blok 1997; Bell & de Jong 2000), results in a correlation between metallicity and total mass.

One of the first, best known failures of the simple model is the so-called G dwarf problem: the closed-box model overpredicts the number of low-metallicity stars observed in the solar neighborhood. This is one aspect of the more fundamental problem that galaxies are clearly not closed boxes. On the one hand, infall and mergers are essential aspects of galaxy formation (e.g., Pagel & Patchett 1975; Naab & Ostriker 2006). On the other hand, galactic-scale winds driven by star formation (a process generally referred to as “feedback” from star formation activity) are a ubiquitous feature of starburst galaxies at low (e.g., Heckman et al. 1990; Lehnert & Heckman 1996; Martin 1999; Strickland et al. 2004) and high (Pettini et al. 2001; Shapley et al. 2003; Smail et al. 2003) redshifts. Metals are detected in the intergalactic medium (IGM; Ellison et al. 2000; Simcoe et al. 2004) at high redshifts, and, at  $z = 2-3$ , their locations are strongly correlated with the distribution of observed star-forming galaxies (Adelberger et al. 2003, 2005a). The potential of such supernova-powered winds to expel gas from galaxies and thus modify their chemical evolution has been known for some time; Larson (1974) showed how winds could account for the mass-metallicity relation in elliptical galaxies by preferentially ejecting metals from those with lower masses. This provides an alternative or additional explanation for the mass-metallicity correlation, and its appeal has increased in recent years as the physical evidence of feedback has multiplied and models of galaxy formation have recognized its importance (e.g., Hernquist & Springel 2003; Benson et al. 2003; Dekel & Woo 2003; Nagamine et al. 2004; Murray et al. 2005).

Motivated by these competing or complementary theories of the origin of the mass-metallicity relation, T04 use the statistical power of  $\sim 53,000$  star-forming SDSS galaxies to revisit the correlation, confirming its existence over  $\sim 3$  orders of magnitude in stellar mass and 1 dex in metallicity. Using estimates of the effective yield as a function of baryonic mass, they see evidence for metal depletion in low-mass galaxies, finding that a galaxy with baryonic mass  $M \sim 4 \times 10^9 M_\odot$  loses half its metals, while low-mass dwarf galaxies are 5 times more metal depleted than  $L_*$  galaxies. They also find that metal depletion occurs in galaxies with masses as high as  $10^{10} M_\odot$  and interpret their results as a signature of winds from an early starburst phase.

Given the strength of the mass-metallicity correlation and its plausible origin from either metal loss through winds or the change in gas fraction as gas is converted to stars (in combination with higher gas fractions in low-mass galaxies), it is not unreasonable to suppose that it may be present in galaxies at high redshift. This has been difficult to test, however. The strong optical emission lines usually used to determine metallicities shift into the infrared past  $z \sim 1$ , making the required large samples of spectra much more difficult to acquire. As luminosity can be determined with considerably greater ease than stellar mass, the first efforts focused on the luminosity-metallicity relation at high redshift (Kobulnicky & Koo 2000; Pettini et al. 2001; Shapley et al. 2004), finding that galaxies at  $z > 2$  are overluminous for their metallicities when compared to local galaxies. Shapley et al. (2004) also found that galaxies with  $M_* \sim 10^{11} M_\odot$  have approximately solar metallicity, while lower mass galaxies have  $Z \sim 0.5 Z_\odot$ .

In this paper we study the relationships among stellar mass, luminosity, and metallicity at  $z \geq 2$ , using a sample of 87 star-

forming galaxies with H $\alpha$  and [N II] spectra. We describe our sample selection, observations, and data reduction procedures in § 2 and discuss our methods of determining stellar mass and metallicity in § 3. In § 4 we give our results, and in § 5 we discuss their implications for the origin of the mass-metallicity relation. Our conclusions are presented in § 6. We use a cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$ , and  $\Omega_\Lambda = 0.7$ ; in such a cosmology, the universe at  $z = 2.26$  (the mean redshift of our sample) is 2.9 Gyr old, or 21% of its present age. For comparisons with solar metallicity, we use the most recent values of the solar oxygen abundance,  $12 + \log (\text{O}/\text{H})_\odot = 8.66$ , and the solar metal mass fraction  $Z_\odot = 0.0126$  (Asplund et al. 2004).

## 2. SAMPLE SELECTION, OBSERVATIONS, AND DATA REDUCTION

The galaxies discussed in this paper are drawn from a sample of 114 galaxies with H $\alpha$  spectra described in detail by Erb et al. (2006a, 2006b). The galaxies were selected by their rest-frame UV colors, and their redshifts were confirmed with rest-frame UV spectra from the LRIS-B spectrograph on the 10 m Keck I telescope; an overview of the  $z \sim 2$  sample is given by Steidel et al. (2004). We then obtained near-IR spectra for a subset of these galaxies. Galaxies were chosen for near-IR spectroscopy for a wide variety of reasons, the end result of which is a sample very similar to the full set of galaxies with spectroscopic redshifts but somewhat biased toward objects that are bright in  $K$  or red in  $\mathcal{R} - K$  or  $J - K$ . A detailed discussion of the sample selection and its relation to the larger sample of UV-selected galaxies is provided by Erb et al. (2006a). For measurements of metallicity in the present paper, we include all objects with H $\alpha$  spectra and  $K$ -band magnitudes (most have  $J$  magnitudes as well, and 32 have also been observed at 3.6, 4.5, 5.4, and  $8.0 \mu\text{m}$  with the IRAC on the *Spitzer Space Telescope*), except those with active galactic nucleus (AGN) signatures in either their rest-frame UV or optical spectra.

The H $\alpha$  spectra were obtained between 2002 May and 2004 September with the near-IR spectrograph NIRSPEC (McLean et al. 1998) on the Keck II telescope. For the redshifts of the galaxies presented here, H $\alpha$  falls in the  $K$  band; most observations were conducted with the N6 filter, which spans the wavelength range  $1.558-2.315 \mu\text{m}$ , and in low-dispersion mode, which provides a resolution of  $R \sim 1400$ . The data were reduced using standard procedures described by Erb et al. (2003, 2006a) and flux-calibrated with reference to near-IR standard stars.

The near-IR imaging was carried out with the Wide-Field IR Camera (WIRC; Wilson et al. 2003) on the Palomar 5 m Hale telescope. We obtained  $\sim 9' \times 9'$  images to  $K_s \sim 22.5$  and  $J \sim 24$  in four fields, with a typical integration time of  $\sim 11$  hr in each band per field. Data reduction and photometry were performed as described by Erb et al. (2006a) and Shapley et al. (2005b). For a description of the mid-IR IRAC data, reductions, and photometry, see Barmby et al. (2004), Shapley et al. (2005b), and Reddy et al. (2005).

## 3. MEASUREMENTS

### 3.1. Stellar Masses

Stellar masses are determined by fitting model spectral energy distributions (SEDs) to the  $U_nGRJK$  (and IRAC, when present) photometry, using the procedure described in detail by Shapley et al. (2005b) and Erb et al. (2006a), which uses the Bruzual & Charlot (2003) population synthesis models and the Calzetti et al. (2000) extinction law. We compare the model SEDs of galaxies with a variety of ages and amounts of extinction to

TABLE 1  
MEAN STELLAR MASSES AND STELLAR POPULATION PROPERTIES

Bin	$N^a$	$\langle z \rangle^b$	$\mathcal{R}^c$	$K_s^d$	$M_B^e$	Stellar Mass <sup>f</sup> ( $10^{10} M_\odot$ )	SFR <sup>g</sup> ( $M_\odot \text{ yr}^{-1}$ )	Age <sup>h</sup> (Myr)	$E(B - V)^i$
1.....	15	2.36	$24.20 \pm 0.72$	$21.51 \pm 0.69$	$-21.27 \pm 0.73$	$0.27 \pm 0.15$	$55 \pm 52$	$180 \pm 276$	$0.19 \pm 0.09$
2.....	15	2.27	$23.96 \pm 0.81$	$21.03 \pm 0.61$	$-21.66 \pm 0.59$	$0.71 \pm 0.17$	$24 \pm 10$	$448 \pm 374$	$0.14 \pm 0.07$
3.....	15	2.21	$24.01 \pm 0.72$	$20.86 \pm 0.47$	$-21.82 \pm 0.58$	$1.5 \pm 0.3$	$30 \pm 21$	$968 \pm 944$	$0.11 \pm 0.08$
4.....	14	2.28	$23.97 \pm 0.62$	$20.54 \pm 0.49$	$-22.15 \pm 0.51$	$2.6 \pm 0.4$	$28 \pm 17$	$1026 \pm 864$	$0.16 \pm 0.07$
5.....	14	2.21	$23.93 \pm 0.60$	$20.16 \pm 0.38$	$-22.33 \pm 0.51$	$4.1 \pm 0.6$	$47 \pm 30$	$1311 \pm 790$	$0.17 \pm 0.09$
6.....	14	2.26	$24.32 \pm 0.55$	$20.05 \pm 0.40$	$-22.30 \pm 0.44$	$10.5 \pm 5.4$	$47 \pm 24$	$2409 \pm 591$	$0.19 \pm 0.06$

<sup>a</sup> Number of galaxies in bin.

<sup>b</sup> Mean and standard deviation of  $z_{\text{H}\alpha}$ .

<sup>c</sup> Mean and standard deviation of  $\mathcal{R}$  magnitude.

<sup>d</sup> Mean and standard deviation of  $K_s$  magnitude.

<sup>e</sup> Mean and standard deviation of absolute magnitude  $M_B$ , determined as described in the text.

<sup>f</sup> Mean and standard deviation of stellar mass from SED fitting; we use a Chabrier (2003) IMF.

<sup>g</sup> Mean and standard deviation of SFR from extinction-corrected  $\text{H}\alpha$  luminosity, including a factor of 2 aperture correction determined from narrowband imaging and comparison of the  $K$ -band continuum with broadband magnitudes. Because the dispersion is large, we also quote the semi-interquartile range for bins 1–6, respectively: 29, 7, 13, 12, 13, and  $15 M_\odot \text{ yr}^{-1}$ .

<sup>h</sup> Mean and standard deviation of best-fit age from SED fitting. Because the dispersion is large, we also quote the semi-interquartile range for bins 1–6, respectively: 123, 184, 677, 490, 541, and 375 Myr.

<sup>i</sup> Mean and standard deviation of best-fit  $E(B - V)$  from SED fitting.

our observed photometry and obtain the star formation rate (SFR) and stellar mass from the normalization of the best-fit model to the data. We try models with a constant star formation rate (CSFR) and models in which the SFR smoothly declines with time, parameterized by  $\text{SFR} \propto e^{-t/\tau}$ , with  $\tau = 10, 20, 100, 200, 500, 1000, 2000$ , and  $5000$  Myr. In practice, however, most star formation histories provide adequate fits to most objects, and we therefore use the CSFR models unless one of the  $\tau$  models provides a significantly better fit. The best-fit models for the galaxies discussed here are given by Erb et al. (2006a). We use a Chabrier (2003) IMF for the stellar masses and SFRs, which results in stellar masses and SFRs 1.8 times smaller than those computed using a Salpeter (1955) IMF. These stellar masses are the integral of the SFR over the lifetime of the galaxy and so represent the total mass of stars formed rather than the mass in living stars at the time of observation. For our adopted IMF, the current living stellar mass is  $\sim 10\%$ – $40\%$  lower depending on the age of the galaxy. We use the total rather than current stellar mass because the relevant quantity for the simple chemical evolution models we apply in § 5 is the fraction of the initial gas mass that has been turned into stars.

Uncertainties in the fitting are determined from a large number of Monte Carlo simulations in which the input photometry is varied according to the photometric errors. The simulations also take into account variations in the star formation history by treating  $\tau$  as a free parameter with the possible values given above. As has often been noted for SED modeling of this sort (e.g., Papovich et al. 2001; Shapley et al. 2001, 2005b), the stellar mass is the most securely determined parameter. For the current sample, the mean fractional uncertainty  $\langle \sigma_{M_*}/M_* \rangle = 0.4$ . The simulations and their results are described in more detail by Erb et al. (2006a).

As discussed by Papovich et al. (2001) and Shapley et al. (2005b), one limitation of such modeling is the insensitivity of the data to faint, old stellar populations, which could be obscured by current star formation and thus lead to an underestimate of the stellar mass. We have estimated the magnitude of this effect by fitting a variety of two-component models to the observed SEDs, in which the light from the observed-frame  $K$  band and redward is constrained to come from a maximally old burst, while a young population is fitted to the rest-frame UV residuals. Such

models make little difference to the stellar masses of already massive galaxies, since their stellar populations already approach the maximum age allowed by the age of the universe at their redshift. The masses of low-mass galaxies can be increased by an order of magnitude by such models, but the models are generally a poor fit to the SED and result in SFRs far higher than those determined by all other indicators, with an average SFR of  $\sim 900 M_\odot \text{ yr}^{-1}$ . While we cannot rule out such models in individual cases, they are very unlikely to be correct on average, and we therefore consider it unlikely that we have underestimated the stellar masses of the low-mass galaxies by such a large factor. More general two-component models, in which the relative contributions from a maximally old population and a young burst are allowed to vary, increase the stellar mass by a factor of a few at most. The two-component models are described in more detail by Erb et al. (2006a).

As described below, for the purpose of determining metallicities, we divide the sample into six bins by stellar mass, with 14 or 15 galaxies in each bin. The mean stellar mass in each bin ranges from  $2.7 \times 10^9$  to  $1.1 \times 10^{11} M_\odot$ . The means and standard deviations of the fitted parameters for each bin are given in Table 1, along with the SFRs determined from  $\text{H}\alpha$  luminosities. The  $\text{H}\alpha$  luminosities have been corrected for extinction using the Calzetti et al. (2000) extinction law and the best-fit value of  $E(B - V)$  from the SED modeling, and we have applied a factor of 2 aperture correction determined from the comparison of the NIRSPEC spectra and narrowband images (for details, and for a full discussion of the  $\text{H}\alpha$ -derived SFRs and their implications, see Erb et al. 2006b).

### 3.2. Metallicities

The most direct way to determine the abundances of metals from the observed emission-line fluxes in  $\text{H II}$  regions is through the measurement of the electron temperature  $T_e$ . As the metallicity of the gas increases, the cooling through metal emission lines also increases, resulting in a decrease in  $T_e$ . The ratio of the auroral (the transition from the second lowest to the lowest excited level) and nebular (the transition from the lowest excited level to the ground state) emission lines of the same ion is highly sensitive to the electron temperature; therefore, the measurement of such pairs of lines has been the preferred method of determining

abundances in H II regions. However, the auroral lines (in particular, the most widely used line, [O III]  $\lambda 4363$ ) become extremely weak at metallicities above  $\sim 0.5$  solar and undetectable at all metallicities in the low signal-to-noise ratio (S/N) spectra of distant galaxies. In most cases, therefore, we must use the empirical “strong-line” abundance indicators, which are based on the ratios of collisionally excited forbidden lines to hydrogen recombination lines.

These are calibrated with reference to the  $T_e$  method or, more commonly, with detailed photoionization models. The strong-line methods carry significant hazards, however. Substantial biases and offsets are observed between abundances determined with different methods and between different calibrations of the same method (e.g., Kennicutt et al. 2003; Kobulnicky & Kewley 2004), and many of the strong-line indicators are sensitive to the ionization parameter, as well as metallicity. The advent of large telescopes, sensitive detectors, and spectrographs with high throughput has enabled the measurement of abundances with the  $T_e$  method in an increasingly large sample of extragalactic H II regions. These new data suggest that the most widely used indicator,  $R_{23} \equiv ([\text{O II}] \lambda 3727 + [\text{O III}] \lambda \lambda 4959, 5007)/\text{H}\beta$ , may systematically overestimate metallicities in the high-abundance regime ( $Z \gtrsim Z_\odot$ ) by as much as 0.2–0.5 dex (Kennicutt et al. 2003; Garnett et al. 2004; Bresolin et al. 2005). The  $T_e$  method itself is not without difficulties, however, as temperature gradients or fluctuations in the H II regions may affect abundance determinations at high metallicities (Stasińska 2005). The result of all this is that absolute values of metal abundances are still quite uncertain, but, fortunately for our present purposes, relative abundances of similar objects, determined with the same method, are more reliable.

The available data limit the options for determining the chemical abundances of the  $z \sim 2$  galaxies. NIRSPEC can observe only one band in a single exposure, and at these redshifts H $\alpha$  and [N II]  $\lambda 6584$  fall in the  $K$  band, [O III]  $\lambda \lambda 5007, 4959$  and H $\beta$  in the  $H$  band, and [O II]  $\lambda 3727$  in  $J$ . We have focused our observations on H $\alpha$  in the  $K$  band, so obtaining the additional data needed to use the  $R_{23}$  indicator would triple the required observing time. Our only option to determine the metallicity of the vast majority of the galaxies in our sample is thus the ratio of [N II] to H $\alpha$ . In addition to minimizing the observing time required to obtain a large sample, this ratio has the advantages that it is insensitive to reddening and is not affected by the relative uncertainties in flux calibration of spectra taken in different bandpasses.

The use of  $N2 \equiv \log ([\text{N II}] \lambda 6584/\text{H}\alpha)$  as an abundance indicator was proposed by Storch-Bergmann et al. (1994) and has been further discussed and refined by Raimann et al. (2000), Denicoló et al. (2002), and Pettini & Pagel (2004). The [N II]/H $\alpha$  ratio is affected by metallicity in two ways. First, there is a tendency for the ionization parameter of an H II region to decrease with increasing metallicity (see, e.g., Fig. 3 of Dopita 2005), thereby increasing the ratio [N II]/[N III] and hence [N II]/H $\alpha$ . Second, nitrogen has both a primary component, whose abundance varies at the same rate as other primary elements, and a secondary component, which increases in abundance with increasing metallicity, further raising the [N II]/H $\alpha$  ratio. Using electron temperature measurements to determine ionic abundances in 20 extragalactic H II regions, Kennicutt et al. (2003) find that the nitrogen abundance is adequately described by a simple model with a primary component with constant  $\log (\text{N}/\text{O}) = -1.5$  and a secondary component for which  $\log (\text{N}/\text{O}) = \log (\text{O}/\text{H}) + 2.2$ . In other words, secondary nitrogen becomes important for  $12 + \log (\text{O}/\text{H}) \gtrsim 8.3$ ; unless we have significantly overestimated the metallicities of the galaxies in our sample, most of our objects are in this regime. As emphasized by Kewley

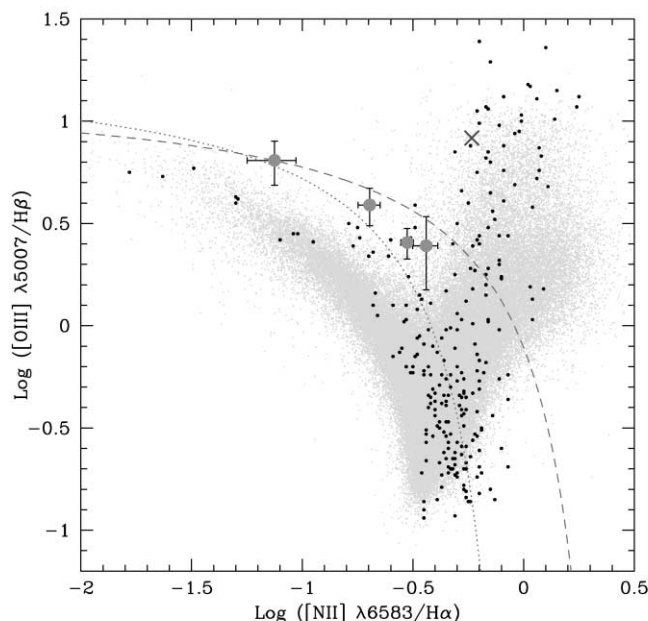


FIG. 1.—[O III]/H $\beta$  vs. [N II]/H $\alpha$  diagnostic diagram. The four galaxies in our sample for which we have measurements of all four lines are shown by the large gray filled circles. The  $z = 2.2$  galaxy discussed by van Dokkum et al. (2005), which shows evidence of an AGN or shock ionization by a wind, is shown by the cross. The small gray dots represent  $\sim 96,000$  objects from the SDSS, and the small black filled circles are the local starburst galaxies studied by Kewley et al. (2001b). The dashed line shows the maximum theoretical starburst line of Kewley et al. (2001a); for realistic combinations of metallicity and ionization parameter, star-forming galaxies fall below and to the left of this line. The dotted line is a similar, empirical determination by Kauffmann et al. (2003). [See the electronic edition of the Journal for a color version of this figure.]

& Dopita (2002), a drawback of the N2 indicator is its sensitivity to the ionization parameter, as well as to metallicity. In addition, the [N II]/H $\alpha$  ratio cannot be used to determine metallicities above approximately solar, at which the N2 index saturates as nitrogen becomes the dominant coolant.

Finally, the [N II]/H $\alpha$  ratio is sensitive to contamination from AGNs and shock excitation, and if possible these should be ruled out before using it as a metallicity indicator. This is generally done through a diagnostic line ratio diagram such as [O III]/H $\beta$  versus [N II]/H $\alpha$  (Baldwin et al. 1981; Veilleux & Osterbrock 1987); on such a diagram, normal galaxies and AGNs fall in generally well-defined regions. Unfortunately, we lack the data to place nearly all of our objects on such a diagram; because we have so far obtained very few [O III]/H $\beta$  spectra, we have measurements of all four lines for only four galaxies. The ratios [O III]/H $\beta$  versus [N II]/H $\alpha$  for these galaxies are shown in Figure 1, along with  $\sim 96,000$  objects from the SDSS (*small gray dots*), the local starbursts analyzed by Kewley et al. (2001b; *small black filled circles*), and the  $z = 2.2$  galaxy (*cross*) discussed by van Dokkum et al. (2005), which shows evidence of shock ionization or an AGN and falls in a clearly different region of the diagram than our galaxies. The dashed line shows the maximum theoretical starburst line determined from photoionization modeling by Kewley et al. (2001a); for realistic combinations of metallicity and ionization parameter, models of normal starbursts fall below and to the left of this line. The dotted line is a similar, empirically determined classification line derived for the SDSS objects by Kauffmann et al. (2003).

Our galaxies fall between the two lines, along a sequence with a higher [O III]/H $\beta$  ratio at a given [N II]/H $\alpha$  ratio (or a higher

[N II]/H $\alpha$  ratio for a given [O III]/H $\beta$  ratio) compared to the SDSS galaxies.<sup>4</sup> A similar offset is seen in star-forming galaxies at  $z \sim 1.4$  (Shapley et al. 2005a). The SDSS contains very few objects in this region, indicating that there are important physical differences between the local and high-redshift samples. One plausible way to produce such a shift in the [N II]/H $\alpha$ –[O III]/H $\beta$  diagram is through some combination of a harder ionizing spectrum and an increase in electron density. Although they are weak, the density-sensitive [S II] lines in the composite spectra indicate an average electron density of  $n_e \sim 500 \text{ cm}^{-3}$  (with no dependence on stellar mass), higher than that found in normal local galaxies but similar to densities seen in local starbursts (Kewley et al. 2001b). Furthermore, Kewley et al. (2001a) show that a relatively hard EUV radiation field is required to model the line ratios of local starbursts; note that a shift is also observed between the local starbursts and the SDSS sample. Detailed photoionization modeling is required to determine quantitatively the origin of this shift and its effect on the calibration of the N2 index, and for the moment the absolute calibration of the metallicity scale remains uncertain [although the relatively shallow slope of the relation between N2 and (O/H) means that offsets are unlikely to be large]. As diagnostic line ratios for more high-redshift galaxies are measured, it is hoped that full photoionization modeling, with improved spectra of the Wolf-Rayet stars that dominate the EUV radiation field, will clarify the physical reasons at the root of the shifts in the [N II]/H $\alpha$ –[O III]/H $\beta$  diagram.

We can exclude an AGN contribution to our sample as a whole based on the X-ray and UV properties of the galaxies. Specifically, Reddy et al. (2005) find that only 3% of UV-selected galaxies at  $z \sim 2$  have X-ray detections in the ultra-deep 2 Ms Chandra Deep Field–North. These galaxies, almost all of which have  $K_s < 20$ , are not included in the composite spectra considered in the present analysis. Further information on AGN contamination comes from the rest-frame UV spectra of our galaxies, which allow the rejection of AGNs on an individual basis in fields without deep X-ray data. As described further in § 4, we have constructed six composite UV spectra, binned by stellar mass in the same way as the NIRSPEC spectra (see below), as well as larger composites of the two highest and two lowest mass bins. None of the composites show AGN features such as broad and/or high-ionization emission lines, placing further limits on low-level AGN activity. On the basis of all of these considerations, we conclude that a significant AGN component to the galaxies considered here is unlikely and proceed under the assumption that the emission lines we observe are produced in H II regions photoionized by hot stars.

This work is primarily concerned with relative, average abundances determined from composite spectra (see below), which can be determined with more accuracy than the metallicities of individual objects; this mitigates the uncertainties associated with the N2 method. We use the calibration of Pettini & Pagel (2004), which is based on H II regions whose ionic abundances have been determined from the electron temperature or from detailed photoionization modeling. From a sample of 137 such H II regions, 131 of which have abundances determined with the  $T_e$  method, Pettini & Pagel (2004) find

$$12 + \log (\text{O}/\text{H}) = 8.90 + 0.57 \times \text{N2}, \quad (1)$$

<sup>4</sup> We have not corrected the H $\beta$  fluxes for stellar absorption, but we do not expect this to be a significant effect. We expect the stellar H $\beta$  absorption line to have an equivalent width  $W_{\text{abs}} \gtrsim -5 \text{ \AA}$  (Kewley et al. 2001b; Kobulnicky et al. 1999), while, assuming a typical ratio of  $W_{\text{H}\alpha}/W_{\text{H}\beta} \sim 5$  (Leitherer et al. 1999), our galaxies have  $W_{\text{H}\beta} \sim 50 \text{ \AA}$ .

with a  $1 \sigma$  dispersion of 0.18 dex. They conclude that the N2 calibrator allows a determination of the oxygen abundance to within a factor of  $\sim 2.5$  with 95% confidence, an accuracy comparable to that of the  $R_{23}$  method.

A further difficulty of the N2 method is that the [N II] line is weak and is generally detected in the spectra of individual objects in our sample only when they approach solar metallicity (Shapley et al. 2004) or have especially strong line fluxes. In order to increase the S/N and improve the likelihood of detecting [N II] at lower metallicities, we have divided the sample into six bins by stellar mass, with 14 or 15 galaxies in each bin, and constructed a composite H $\alpha$  + [N II] spectrum of the objects in each bin. The use of composite spectra has the advantage that, in addition to increasing S/N, it minimizes the effects of the mass and metallicity uncertainties of individual objects because we are only concerned with the average properties of the galaxies in each bin. We first shift the flux-calibrated spectra into the rest frame and then average them, rejecting the minimum and maximum value at each dispersion point to suppress noise from large sky subtraction residuals in individual spectra. The six composite spectra, labeled with the mean stellar mass in each bin, are shown in Figure 2.

We measure the [N II]/H $\alpha$  ratio by first measuring the H $\alpha$  fluxes, central wavelengths, and widths and then constraining the [N II] line to have the same width and a central wavelength fixed by the position of H $\alpha$ . We use the rms of the spectrum between emission lines to determine the typical noise in each spectrum; because the galaxies are at different redshifts, the systematic effects of the night-sky lines are minimized in the composites, and the rms provides an adequate description of the noise. This procedure provides a good fit to the [N II] line for all of the composite spectra (except in the lowest mass bin, where we determine an upper limit on the [N II] flux). The measured H $\alpha$  and [N II] fluxes for each composite, as well as the inferred value of  $12 + \log (\text{O}/\text{H})$ , are given in Table 2. The listed uncertainties in  $12 + \log (\text{O}/\text{H})$  include the scatter in the N2 calibration (a  $1 \sigma$  uncertainty of 0.18 dex reduced by  $N^{1/2}$ , where  $N$  is the number of objects in the composite spectrum), as well as the uncertainties in the measurements of the H $\alpha$  and [N II] fluxes.

#### 4. THE MASS-METALLICITY RELATION

Figure 3 shows the mean metallicity of the galaxies in each mass bin plotted against their mean stellar mass (*large gray filled circles*); there is a monotonic increase in metallicity with stellar mass. The vertical error bars show the uncertainty in  $12 + \log (\text{O}/\text{H})$  from measurement uncertainties in the [N II]/H $\alpha$  ratio, while the additional vertical error bar in the lower right corner shows the uncertainty due to the scatter in the N2 calibration. The horizontal bars show the range of stellar masses in each bin. The most massive galaxies in the sample have close to solar metallicities, a result found previously by Shapley et al. (2004), who measured the [N II]/H $\alpha$  ratio from individual spectra of the brightest objects with  $K_s < 20$ . More typical galaxies with  $M_* \sim 10^{10} M_\odot$  have  $12 + \log (\text{O}/\text{H}) \sim 8.4$ , while for the lowest mass objects we can only place an upper limit  $12 + \log (\text{O}/\text{H}) < 8.2$ , or  $(\text{O}/\text{H}) < \frac{1}{3}(\text{O}/\text{H})_\odot$ .

We have not found any systematic effects that could spuriously produce the clear observed correlation between stellar mass and metallicity. Specifically, it is possible that the value of  $M_*$  in the lowest mass bins has been underestimated if an older stellar population is already in place in these galaxies (see the discussion in § 3.1). This would have the effect of steepening the observed correlation, as the lowest mass bins in Figure 3 would move to the right by 0.3–0.5 dex, while the high-mass bins

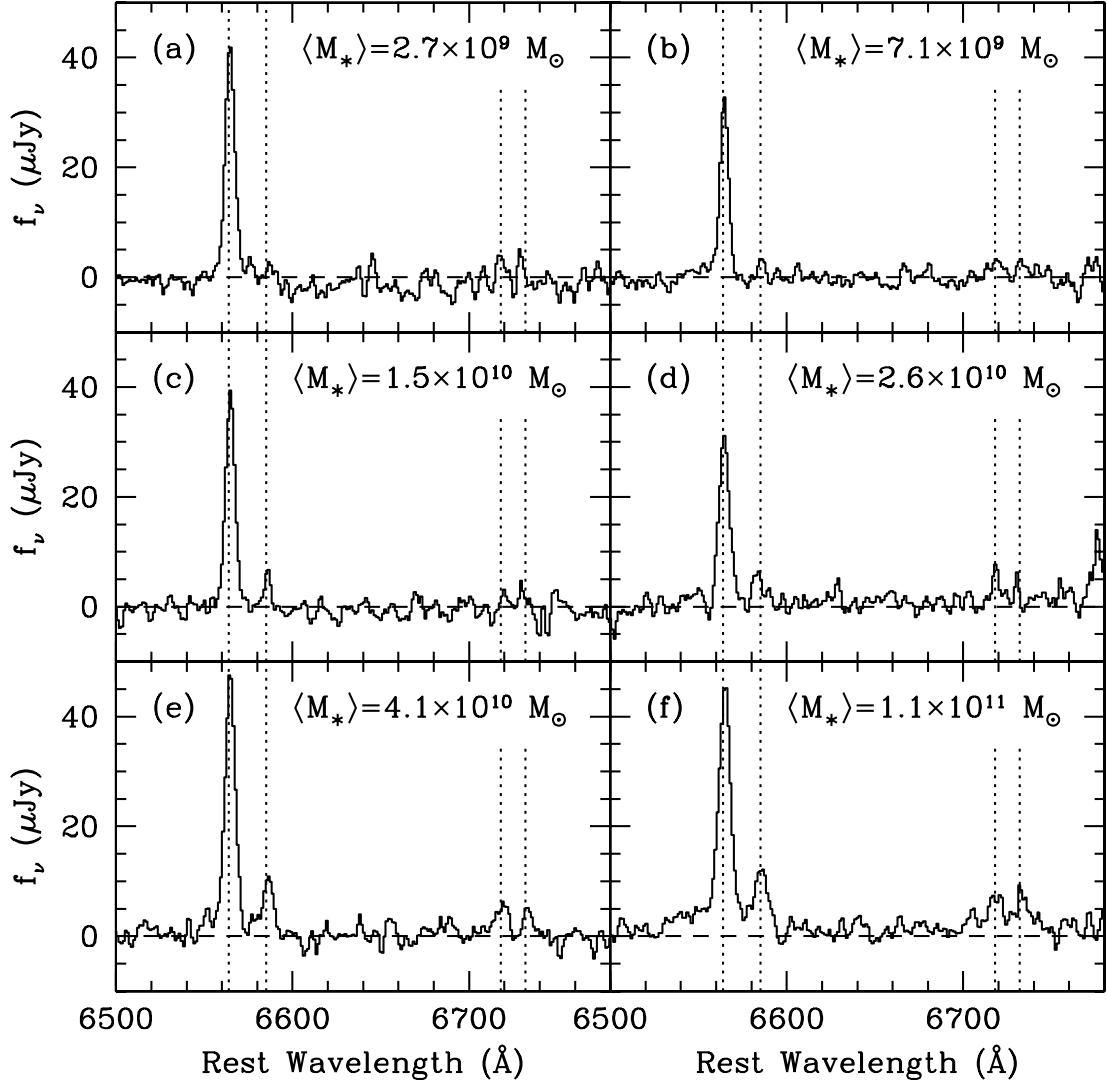


FIG. 2.—Composite NIRSPEC spectra of the 87 galaxies in our sample, divided into six bins of 14 or 15 objects each by increasing stellar mass. The spectra are labeled with the mean stellar mass in each bin, and the  $H\alpha$ ,  $[N\ II]$ , and  $[S\ II]$  lines are marked by dotted lines (*left to right, respectively*). The increase in the strength of  $[N\ II]$  with stellar mass can be seen clearly. The density-sensitive  $[S\ II]$  lines, while weak, indicate a typical electron density of  $n_e \sim 500\text{ cm}^{-3}$ , with no significant dependence on mass; this is a value comparable to that seen in local starburst galaxies (Kewley et al. 2001b).

TABLE 2  
OXYGEN ABUNDANCES AND GAS FRACTIONS

Bin	Stellar Mass <sup>a</sup> ( $10^{10} M_{\odot}$ )	$F_{H\alpha}$ <sup>b</sup> ( $10^{-17}\text{ ergs s}^{-1}\text{ cm}^{-2}$ )	$F_{[N\ II]}$ <sup>b</sup> ( $10^{-17}\text{ ergs s}^{-1}\text{ cm}^{-2}$ )	N2 <sup>c</sup>	$12 + \log(O/H)$ <sup>d</sup>	$M_{\text{bar}}$ <sup>e</sup> ( $10^{10} M_{\odot}$ )	$\mu_{\text{gas}}$ <sup>f</sup>	$y_{\text{eff}}$ <sup>g</sup>
1.....	$0.27 \pm 0.15$	$20.5 \pm 0.5$	$<1.2$	$<-1.22$	$<8.20$	$2.7 \pm 1.7$	$0.85 \pm 0.12$	$<0.027$
2.....	$0.71 \pm 0.17$	$13.9 \pm 0.3$	$1.4 \pm 0.2$	$-1.00^{+0.07}_{-0.09}$	$8.33^{+0.07}_{-0.07}$	$2.1 \pm 0.6$	$0.63 \pm 0.12$	$0.013 \pm 0.003$
3.....	$1.5 \pm 0.3$	$18.7 \pm 0.4$	$2.7 \pm 0.3$	$-0.85^{+0.05}_{-0.06}$	$8.42^{+0.06}_{-0.05}$	$3.2 \pm 1.1$	$0.48 \pm 0.19$	$0.010 \pm 0.002$
4.....	$2.6 \pm 0.4$	$15.9 \pm 0.4$	$2.6 \pm 0.3$	$-0.78^{+0.05}_{-0.05}$	$8.46^{+0.06}_{-0.05}$	$4.0 \pm 0.9$	$0.33 \pm 0.12$	$0.007 \pm 0.001$
5.....	$4.1 \pm 0.6$	$24.3 \pm 0.5$	$5.3 \pm 0.4$	$-0.66^{+0.03}_{-0.04}$	$8.52^{+0.06}_{-0.05}$	$6.6 \pm 1.1$	$0.36 \pm 0.10$	$0.009 \pm 0.002$
6.....	$10.5 \pm 5.4$	$27.0 \pm 0.4$	$7.4 \pm 0.3$	$-0.56^{+0.02}_{-0.02}$	$8.58^{+0.06}_{-0.04}$	$13.1 \pm 5.6$	$0.22 \pm 0.11$	$0.007 \pm 0.001$

<sup>a</sup> Mean and standard deviation of stellar mass from SED fitting; we use a Chabrier (2003) IMF.

<sup>b</sup> Fluxes of  $H\alpha$  and  $[N\ II]\ \lambda 6584$  from the composite spectra.

<sup>c</sup>  $N2 \equiv \log(F_{[N\ II]}/F_{H\alpha})$ .

<sup>d</sup> Oxygen abundance from N2, using the calibration of Pettini & Pagel (2004).

<sup>e</sup> Mean and standard deviation of the baryonic mass  $M_{\text{gas}} + M_*$ , with gas masses determined from the Schmidt law as described in the text.

<sup>f</sup> Mean and standard deviation of the gas fraction  $\mu = M_{\text{gas}}/(M_{\text{gas}} + M_*)$ .

<sup>g</sup> Effective yield  $y_{\text{eff}} = Z/\ln(1/\mu)$ .

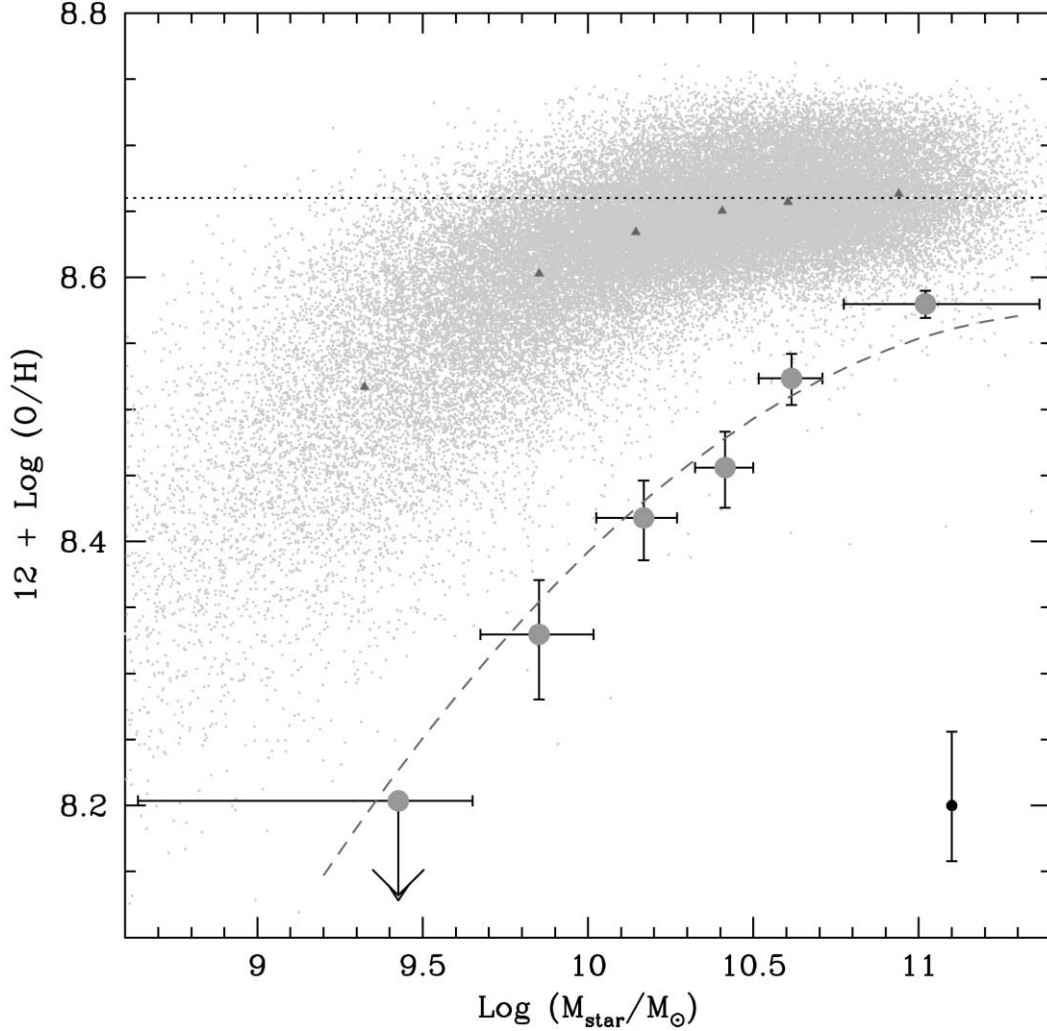


FIG. 3.—Observed relation between stellar mass and oxygen abundance at  $z \sim 2$ , shown by the large gray filled circles. Each point represents the average value of 14 or 15 galaxies, with the metallicity estimated from the  $[\text{N II}]/\text{H}\alpha$  ratio of their composite spectrum. Horizontal bars indicate the range of stellar masses in each bin, while the vertical error bars show the uncertainty in the  $[\text{N II}]/\text{H}\alpha$  ratio. The additional error bar in the lower right corner shows the additional uncertainty in the N2 calibration itself. The dashed line is the best-fit mass-metallicity relation of T04, shifted downward by 0.56 dex. The metallicities of different samples are best compared using the same calibration; we therefore show, with small gray dots, the metallicities of the  $\sim 53,000$  SDSS galaxies of T04 determined with the N2 index. Note that the  $[\text{N II}]/\text{H}\alpha$  ratio saturates near solar metallicity (*horizontal dotted line*). The filled triangles indicate the mean metallicity of the SDSS galaxies in the same mass bins we use for our sample; using the more reliable, low-metallicity bins, our galaxies are  $\sim 0.3$  dex lower in metallicity at a given mass. [See the electronic edition of the *Journal* for a color version of this figure.]

would remain unaffected. The use of the integrated SFR as the stellar mass rather than the mass in currently existing stars (see § 3.1) would also steepen and shift the correlation somewhat, with the lower mass points moving  $\sim 0.1$  dex to the left and the upper mass points shifting left by  $\sim 0.2$  dex. AGN contamination is highly unlikely to produce the correlation, given the low fraction of AGNs in the sample. Variations in the ionization parameter are also unlikely to be correlated with the assembled stellar mass. The ionization parameter of an  $\text{H II}$  region depends on the age of the ionizing cluster, which is very much less than the age of the galaxy (see, e.g., Dopita 2005); since each galaxy in our sample presumably contains many  $\text{H II}$  regions of different ages, the overall variation in ionization parameter from galaxy to galaxy should be small. There is also no dependence of electron density on galaxy mass; the density-sensitive  $[\text{S II}]$  lines in the composite spectra indicate an average density  $n_e \sim 500 \text{ cm}^{-3}$ , with no trend with stellar mass. We therefore have no reason to

expect the ionization parameter to depend on the total stellar mass (other than the dependence on metallicity, which is included in the N2 calibration).

The dashed line in Figure 3 shows the mass-metallicity relation determined for  $\sim 53,000$  star-forming SDSS galaxies by T04, after an arbitrary downward shift of 0.56 dex. With this shift the SDSS relation matches the  $z \sim 2$  galaxies remarkably well, although it is slightly shallower. The empirical shift of 0.56 dex includes an offset due to the different abundance diagnostics used in the two studies: while the SDSS metallicity determinations take into consideration all of the strong nebular lines, ours are based on the N2 index alone, for the reasons explained above. For a more consistent comparison, we use the N2 calibration to calculate the metallicities of the same 53,000 SDSS galaxies, shown by small gray dots in Figure 3; the mean N2 metallicity, in bins spanning the same range of stellar masses used for our sample, is shown by the small filled triangles. The saturation of

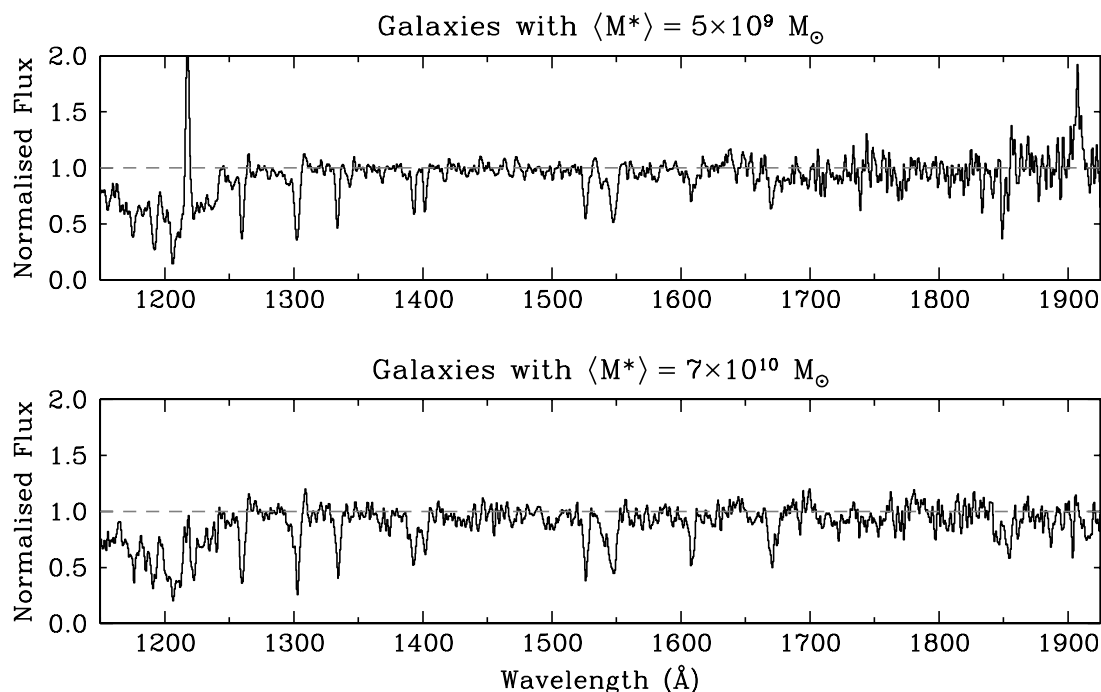


FIG. 4.—Composite rest-frame UV spectra of galaxies in the two lower (*top*) and two higher (*bottom*) mass bins in our sample. The upper spectrum is the average of 30 LRIS-B spectra of galaxies with mean stellar mass  $\langle M_* \rangle = 5 \times 10^9 M_\odot$ , while 28 spectra contribute to the lower composite for which  $\langle M_* \rangle = 7 \times 10^{10} M_\odot$ . The spectra have been divided by the underlying stellar continuum estimated according to the prescription by Rix et al. (2004). Differences in the stellar, interstellar, and nebular lines between the two composites are discussed in the text. [See the electronic edition of the *Journal* for a color version of this figure.]

the  $[\text{N II}]/\text{H}\alpha$  ratio at metallicities approaching solar (*horizontal dotted line*) is clearly apparent in the SDSS sample and makes the determination of the true offset between the two samples difficult. For galaxies with  $M_* \sim 2 \times 10^9 M_\odot$ , the oxygen abundance of the  $z \sim 2$  sample is lower by  $>0.3$  dex; at higher stellar masses the offset is smaller, due at least in part to the saturation of the  $[\text{N II}]/\text{H}\alpha$  ratio. Taking the two lowest mass bins as the most reliable indicators and assuming that the shape of the relation remains unchanged, we find that galaxies at  $z \sim 2$  are  $\sim 0.3$  dex lower in metallicity than galaxies of the same stellar mass today. We discuss the likely reasons for this offset and its uncertainties in § 5.1.

#### 4.1. Composite Ultraviolet Spectra

Given the uncertainties in the absolute metallicity scale associated with the N2 index, it would be highly desirable to obtain independent abundance measures for the galaxies considered here. As discussed by Rix et al. (2004), the UV spectrum of star-forming galaxies is rich in stellar spectral features that provide abundance diagnostics for the young stellar populations. The difficulty is that these are low-contrast features, usually requiring data of higher quality than can be obtained with current instrumentation. Nevertheless, it is worthwhile to examine whether the rest-frame UV spectra of the galaxies under study are consistent with the abundance trend revealed by Figure 3.

To this end, we have constructed two composite spectra, each consisting of approximately 30 galaxies, by averaging the LRIS-B spectra of the galaxies in, respectively, the two lower and the two higher mass bins in Figure 3. The corresponding mean stellar masses are  $\langle M_* \rangle = 5 \times 10^9$  and  $7 \times 10^{10} M_\odot$ . The coarser mass binning was required to improve the S/N of the rest-UV composites to the level where the stellar absorption features that are sensitive to metallicity could be clearly discerned. The two composite spectra are shown in Figure 4, after normalization to the

underlying stellar continua following the prescription by Rix et al. (2004).

Within the spectral region of 1150–1925 Å covered by the composites, the interval near 1400 Å is particularly suitable for an abundance analysis; we show this portion on an expanded scale in Figure 5. The region includes two blends of stellar photospheric lines, centered at 1370 and 1425 Å, whose strengths were shown by Leitherer et al. (2001) to be mostly sensitive to metallicity. The “1370” feature is a blend of  $\text{O V } \lambda 1371$  and  $\text{Fe V } \lambda \lambda 1360\text{--}1380$ , while  $\text{Si III } \lambda 1417$ ,  $\text{C III } \lambda 1427$ , and  $\text{Fe V } \lambda 1430$  make up the “1425” feature. This portion of the spectra also encompasses the  $\text{Si IV } \lambda \lambda 1393, 1402$  doublet, which has a broad stellar P Cygni component formed in the expanding atmospheres of late O and early B supergiants; superposed on this broad component are narrower interstellar absorption lines due to the ambient interstellar medium that lies in front of the stars.

In Figure 5 we have also reproduced synthetic spectra generated with the Starburst99 code (Leitherer et al. 1999, 2001) for the standard case of continuous star formation and Salpeter IMF, using in turn the two empirical stellar libraries available in the code, built from spectra of OB stars in, respectively, the Milky Way (the solar metallicity, or  $Z = Z_\odot$ , library) and the Magellanic Clouds (the MC metallicity, or  $Z = \frac{1}{3} Z_\odot$ , library). Figure 5 shows that the former is a plausible match to the composite spectrum of the higher stellar mass galaxies, while the latter is consistent with the composite UV spectrum of the galaxies in the lower stellar mass bins. We now discuss this comparison in more detail.

Focusing on the composite spectrum of the most massive galaxies (Fig. 5, *bottom panel*), we find that the overall equivalent widths of the 1370 and 1425 photospheric features are in agreement between empirical and synthetic spectra. The comparison of the  $\text{Si IV } \lambda \lambda 1393, 1402$  P Cygni component is complicated by the blending with the interstellar lines, which, as is



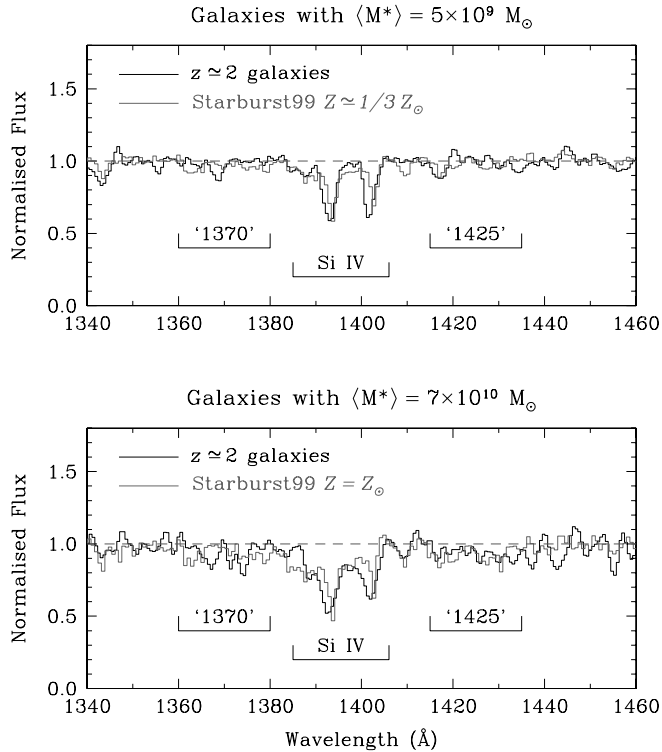


FIG. 5.—Close-up of the 1400 Å region in the two composite spectra shown in Fig. 4 (black histogram). This region contains two blends of stellar photospheric lines, labeled “1370” and “1425,” whose strength is thought to depend primarily on metallicity, and the Si IV  $\lambda\lambda 1393, 1402$  doublet, which consists of a broad P Cygni stellar absorption on which narrower interstellar lines are superposed. The gray histogram shows the Starburst99 spectrum for the standard case of continuous star formation with a Salpeter IMF and Magellanic Cloud (top) or solar (bottom) metallicities. [See the electronic edition of the *Journal* for a color version of this figure.]

usually the case, are stronger (and blueshifted) in starburst galaxies than in the individual stars that make up the stellar libraries. Nevertheless, the broad component of the Si IV feature has comparable optical depth to the solar metallicity Starburst99 model. Turning to the lower mass galaxies (Fig. 5, top panel), we see that all three spectral features are undetected in the lower stellar mass UV composite. The Magellanic Cloud metallicity model spectrum also shows these lines to be much reduced in strength.

The S/N of the data and the subtlety of the spectral features in question limit the above comparison to qualitative statements, and we do not consider a more quantitative approach (such as a  $\chi^2$  analysis, for example) to be warranted in the present circumstances. Nevertheless, with only the UV spectra at our disposal, we would have concluded that the galaxies with a mean stellar mass  $\langle M_* \rangle = 7 \times 10^{10} M_\odot$  have metallicity  $Z \sim Z_\odot$  [or  $12 + \log(\text{O}/\text{H}) \sim 8.6$ ] and that those with  $\langle M_* \rangle = 5 \times 10^9 M_\odot$  have  $Z \lesssim \frac{1}{3} Z_\odot$  [or  $12 + \log(\text{O}/\text{H}) \lesssim 8.1$ ]. Given the uncertainties, these conclusions are broadly consistent with those deduced from our analysis of the [N II]/H $\alpha$  ratios.

There are other differences between the two composite UV spectra reproduced in Figure 4; for example, the interstellar absorption lines are significantly stronger in the galaxies with higher stellar mass. The strongest UV interstellar lines, Si II  $\lambda 1260$ , O I  $\lambda 1302$ +Si II  $\lambda 1304$ , C II  $\lambda 1334$ , Si II  $\lambda 1526$ , Fe II  $\lambda 1608$ , and Al II  $\lambda 1670$ , have rest-frame equivalent widths  $W_0 = 2\text{--}3$  Å in the  $\langle M_* \rangle = 7 \times 10^{10} M_\odot$  composite and  $W_0 = 1.5\text{--}2$  Å in the  $\langle M_* \rangle = 5 \times 10^9 M_\odot$  composite. Since these lines are all strongly saturated (e.g., Pettini et al. 2002), the higher equivalent

widths are much more likely to be due to a larger velocity dispersion of the absorbing gas than to an increase in the column densities of the metals. This may be related to differences in the star formation histories of the two samples of galaxies; on average, star formation has been in progress for longer in galaxies with higher values of  $\langle M_* \rangle$  and presumably more kinetic energy has been deposited in their interstellar media, stirring the gas to higher velocity dispersions.

Finally, the most obvious difference between the two composite UV spectra is in the strengths of the two nebular emission lines that fall within our wavelength range, Ly $\alpha$  and C III]  $\lambda 1908$ , which are much stronger in the lower stellar mass galaxies. Qualitatively, this difference is in agreement with the metallicity and kinematics differences discussed above. The C III]  $\lambda\lambda 1907, 1909$  doublet is similar to [O III]  $\lambda\lambda 5007, 4959$  in showing a steep *inverse* dependence on metallicity in the high-metallicity regime: as the metallicity increases, the temperature of the H II regions decreases and so do the relative populations of the collisionally excited levels from which these emission lines originate (e.g., Pettini & Pagel 2004). Thus, it is not surprising to find that C III]  $\lambda 1908$  is strong in the UV composite with metallicity  $Z \lesssim \frac{1}{3} Z_\odot$  and below our detection limit in the galaxies with metallicities close to solar. The dominant factor that determines the escape of resonantly scattered Ly $\alpha$  photons from a star-forming galaxy is the velocity dispersion of the interstellar gas through which the photons propagate: the larger the velocity dispersion, the smaller (and more blueshifted) is the fraction of Ly $\alpha$  photons that escape before being absorbed by dust and turned into infrared photons (e.g., Mas-Hesse et al. 2003). Thus, the stronger Ly $\alpha$  flux of the galaxies making up the lower stellar mass composite in the top panel of Figure 4 goes hand in hand with the smaller velocity dispersion of their interstellar media, as discussed above.

#### 4.2. The Luminosity-Metallicity Relation

Because the luminosity of a galaxy is far more easily determined than its mass, there are a great many more luminosity-metallicity ( $L$ - $Z$ ) relations than mass-metallicity relations in the literature (e.g., Skillman et al. 1989; Zaritsky et al. 1994; Garnett et al. 1997; Lamareille et al. 2004; Salzer et al. 2005). These correlations span 11 orders of magnitude in luminosity and 2 dex in metallicity and are seen in galaxies of all types. Both the slope and the zero point of the relation shift depending on the bandpass in which the correlation is determined (Salzer et al. 2005); it is traditional to use the absolute  $B$  magnitude, but both the slope and the dispersion of the relation decrease as wavelength increases to the IR, probably due to decreased extinction and the closer correspondence of the infrared luminosity to stellar mass. The metallicity-luminosity relation has been observed in galaxies at redshifts up to  $z \sim 1$  (Kobulnicky et al. 2003; Lilly et al. 2003; Kobulnicky & Kewley 2004; Maier et al. 2004; Liang et al. 2004), at which point the shifting of most of the strong nebular lines into the IR makes spectroscopy much more difficult. Most of these studies show that the zero point of the  $L$ - $Z$  relation evolves with redshift, so that galaxies of a given luminosity have decreasing metallicity with increasing redshift. The small number so far of metallicity-luminosity comparisons at  $z > 2$  confirms this trend, as high-redshift galaxies are 2–4 mag brighter than local galaxies of comparable metallicity (Kobulnicky & Koo 2000; Pettini et al. 2001; Shapley et al. 2004).

We construct a luminosity-metallicity relation analogous to our mass-metallicity relation by dividing our sample of 87 galaxies into six bins by rest-frame absolute  $B$  magnitude  $M_B$ , which we determine by multiplying the best-fit SED of each object by the redshifted  $B$ -band transmission curve. We construct a

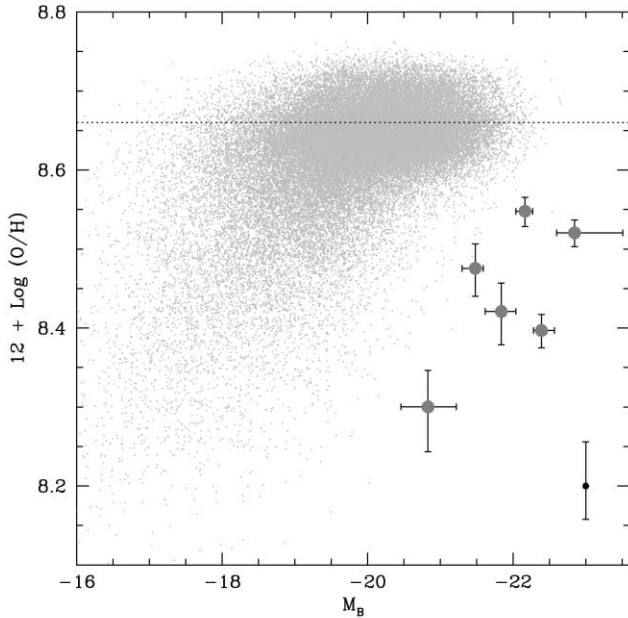


FIG. 6.—Luminosity-metallicity relation at  $z \approx 2$ . We have divided the sample into six bins by rest-frame absolute  $B$  magnitude and estimated the metallicity in each bin. The symbols are the same as in Fig. 3. The points are not significantly correlated, although the faintest galaxies do have the lowest metallicities. The lack of correlation can be understood through the large variation in the optical mass-to-light ratio at high redshift. The  $z \sim 2$  galaxies are  $\sim 3$  mag brighter than the SDSS galaxies at a given metallicity, as estimated by the  $[\text{N II}]/\text{H}\alpha$  ratio. [See the electronic edition of the Journal for a color version of this figure.]

composite spectrum of the galaxies in each bin, and we measure the  $[\text{N II}]/\text{H}\alpha$  ratio and determine the oxygen abundance in the manner described in § 3.2. The results are shown in Figure 6, again including SDSS metallicities determined from the N2 indicator. It is immediately apparent that the correlation between luminosity and metallicity is weaker than that between mass and metallicity; the trend with luminosity is not monotonic and is not statistically significant. A Spearman correlation test finds a 33% probability that the points are uncorrelated, giving a significance of  $1\sigma$ . The faintest galaxies do have the lowest metallicities, however; any appearance of correlation is driven by this bin.

The comparison with the SDSS sample is again made difficult by the saturation of the  $[\text{N II}]/\text{H}\alpha$  ratio around solar metallicity, but it is clear that the high-redshift galaxies have both lower metallicities (subject to the caveats discussed in § 4) and higher luminosities than most of the local sample. Considering the offset of the more reliable lower metallicity bins, we see from Figure 6 that the  $z \sim 2$  galaxies are approximately 3 mag brighter than local galaxies with the same oxygen abundance. It is difficult to invert the comparison to determine the difference in metallicity between galaxies of a given luminosity in the two samples, however, since virtually all of the local galaxies as bright as the  $z \sim 2$  sample have solar or greater abundances that cannot be accurately determined by the N2 method. In order for the best-fit  $L-Z$  relation determined by T04 to pass through the average luminosity and metallicity of our sample, it must be shifted downward by  $\sim 0.9$  dex. After allowing for the  $\sim 0.25$  dex systematic difference between the metallicity diagnostics used by T04 and here (as discussed in § 4), we are still left with a shift of  $0.6\text{--}0.7$  dex between the local and high-redshift samples. Our mean values are offset from other local  $L-Z$  relations by amounts ranging from  $\sim 0.4$  (Skillman et al. 1989) to  $\sim 1.0$  dex (Lamareille et al. 2004).

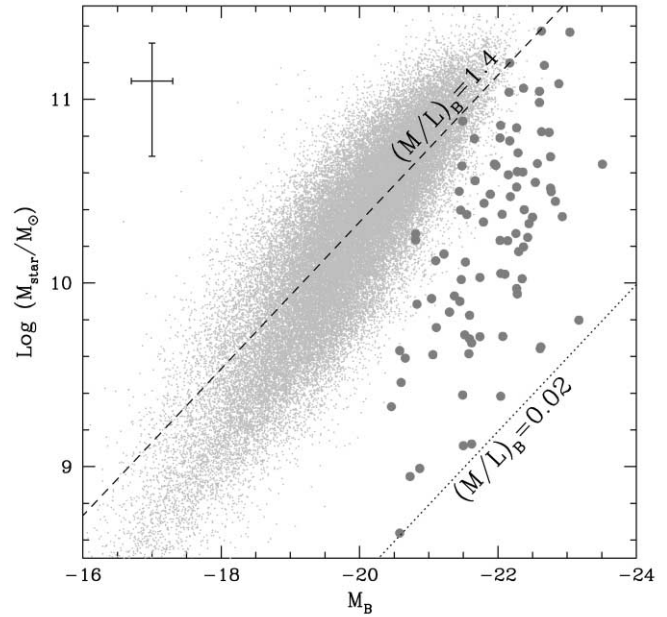


FIG. 7.—Absolute rest-frame  $B$  magnitude vs. stellar mass, for the individual galaxies in the  $z \sim 2$  sample (large gray filled circles) and the SDSS (small gray dots). The dashed and dotted lines show constant mass-to-light ratios  $M/L$ ; in the  $z \sim 2$  sample, the total range in luminosity is narrow and  $M/L$  varies by a factor of  $\sim 70$  at most luminosities, while for most SDSS galaxies,  $M/L$  varies by a factor of  $2\text{--}5$  at a given luminosity. [See the electronic edition of the Journal for a color version of this figure.]

The large offsets between the various local relations are due to differences in calibration methods and sample selection; the relations are not directly comparable, and it is not yet clear which of them, if any, provides the most appropriate comparison for our sample. While it is probably a robust conclusion that the  $z \sim 2$  galaxies are more metal-poor than their luminous local counterparts, a better understanding of systematic effects is required to reliably quantify the differences.

Our results are consistent with previous  $L-Z$  determinations at  $z > 2$ , and, moreover, they are not surprising given our knowledge of the high-redshift sample. Star-forming galaxies at  $z \sim 2$  have lower average mass-to-light ratios  $M/L$  than local galaxies, and the variation in  $M/L$  at a given rest-frame optical luminosity can be as much as a factor of  $\sim 70$  (Shapley et al. 2005b). This large variation in  $M/L$  explains the lack of correlation in the  $z \sim 2$   $L-Z$  relation compared to the local relation. Figure 7 shows absolute magnitude  $M_B$  plotted against stellar mass for the individual galaxies in the  $z \sim 2$  and SDSS samples. The  $z \sim 2$  sample is shown by large filled circles; at a given luminosity,  $M/L$  varies by up to a factor of 70, and the range in luminosity is small, so that galaxies with a wide range of stellar masses fall in each of the six bins in luminosity. In contrast,  $M_*$  and  $M_B$  are much more tightly correlated in the SDSS sample (small gray dots); at a given luminosity, most of the points are within a factor of a few in  $M/L$ . If the fundamental correlation is between metallicity and mass, these differences in  $M/L$  would shift the relation to higher luminosities and dramatically increase the scatter, as we have observed. A mass-metallicity correlation is clearly more physically meaningful than a luminosity-metallicity correlation at high redshifts; a corollary is that the local metallicity-luminosity relation is simply a result of the strong correlation between mass and luminosity at low redshift. While it is likely that an  $L-Z$  relation determined in the rest-frame IR would show

a stronger correlation, investigation of this possibility must wait until a larger sample of high- $z$  galaxies with both mid-IR photometry and nebular line spectra has been assembled.

### 5. THE ORIGIN OF THE MASS-METALLICITY RELATION

A correlation between gas-phase metallicity and stellar mass can plausibly be explained either by the tendency of lower mass galaxies to have larger gas fractions (McGaugh & de Blok 1997; Bell & de Jong 2000) and thus be less enriched, or by the preferential loss of metals from galaxies with shallow potential wells by galactic-scale winds. With the relevant information on the star, gas, and metal content of the galaxies, the two effects can be differentiated. In the simple, closed-box model of chemical evolution with no inflows or outflows, the mass fraction of metals  $Z$  is a simple function of the gas fraction  $\mu \equiv M_{\text{gas}}/(M_{\text{gas}} + M_*)$  and the true yield  $y$ , which represents the ratio of the mass of metals produced and ejected by star formation to the mass locked in long-lived stars and remnants. The true yield is a function of stellar nucleosynthesis, and as in previous, similar studies (T04; Garnett 2002), we assume that it is constant (for a discussion see Garnett 2002). The metallicity is then given by

$$Z = y \ln(1/\mu). \quad (2)$$

This equation can be inverted to determine the effective yield  $y_{\text{eff}}$  from the observed metallicity and gas fraction,  $y_{\text{eff}} = Z/\ln(1/\mu)$ . If the simple model applies,  $y_{\text{eff}}$  will be constant for all masses and equal to the true yield  $y$ , while a decrease in  $y_{\text{eff}}$  (either with respect to the expected true yield or, more commonly, as a function of mass) is a signature of outflows or of dilution by the infall of metal-poor gas (Edmunds 1990).

T04 determined the effective yields of the SDSS galaxies by using the empirical Schmidt law (Kennicutt 1998), which relates the SFR per unit area to the gas surface density, to estimate gas masses. They found lower effective yields in galaxies with lower baryonic masses (the baryonic mass is expected to correlate with the dark matter content and thus indicate the depth of the potential well; McGaugh et al. 2000) and interpreted this result as evidence for the preferential loss of metals in low-mass galaxies. We carry out a similar analysis on our sample of  $z \sim 2$  galaxies. As described by Erb et al. (2006a), we use a galaxy's  $H\alpha$  luminosity and the spatial extent of its  $H\alpha$  emission  $r_{H\alpha}$  (after deconvolution with the seeing point-spread function) to calculate the  $H\alpha$  luminosity per unit area  $\Sigma_{H\alpha}$ . We then estimate each galaxy's gas surface density  $\Sigma_{\text{gas}}$  by combining the Kennicutt (1998) relation between SFR per unit area and gas density with the conversion from  $H\alpha$  luminosity to SFR from the same paper as follows:

$$\Sigma_{\text{gas}} = 1.6 \times 10^{-27} \left( \frac{\Sigma_{H\alpha}}{\text{ergs s}^{-1} \text{ kpc}^{-2}} \right)^{0.71} M_{\odot} \text{ pc}^{-2}. \quad (3)$$

The estimated gas mass is then  $M_{\text{gas}} \sim \Sigma_{\text{gas}} r_{H\alpha}^2$ . We next combine  $M_{\text{gas}}$  and  $M_*$  to obtain an estimate of the gas fraction  $\mu$  and compute the mean value of  $\mu$  in each mass bin. Although the considerable (factor of  $\sim 2$ ) uncertainties in both the corrected  $H\alpha$  flux and the galaxy's size translate into a significant uncertainty in  $\mu$  for individual objects, the number of galaxies in each bin allows us to determine the mean value of  $\mu$  to within  $\sim 10\%$ .

Galaxies at low and intermediate redshifts show significant correlations between metallicity and extinction (e.g., Cortese

et al. 2006). Galaxies at  $z \sim 2$  also show a strong correlation between dust obscuration and stellar mass (and hence metallicity), as shown by Reddy et al. (2006), who use the  $24 \mu\text{m}$  luminosity as observed by the *Spitzer Space Telescope* to infer the infrared luminosity  $L_{\text{IR}}$  and the extinction as parameterized by  $L_{\text{IR}}/L_{\text{UV}}$ . One implication of the work of Reddy et al. (2006) is that  $E(B - V)$  as determined from the UV slope may overestimate the extinction correction for galaxies with ages less than 100 Myr. This effect can be seen in the high mean value of  $E(B - V)$  for the lowest mass bin in Table 1; for galaxies with ages greater than 100 Myr in the current sample,  $E(B - V)$  and stellar mass are correlated with  $4 \sigma$  significance. This overestimation of the extinction in young galaxies has a negligible effect on our current analysis; we estimate that it will lead to an overestimation of the SFRs by a typical factor of  $\sim 1.2$ , considerably less than other uncertainties. Only the lowest mass bin is affected, since it is the only one that contains a significant number of young objects; because we derive only an upper limit on the mean metallicity of the galaxies in this bin, it does not affect our conclusions.

As in the local universe (McGaugh & de Blok 1997; Bell & de Jong 2000), there is a strong trend of decreasing gas fraction with increasing stellar mass. The lowest mass bin in the sample has a mean gas fraction  $\langle \mu \rangle = 0.85$ , while the highest mass bin has  $\langle \mu \rangle = 0.22$ . The median value of  $\mu$  in our sample is  $\sim 0.5$ , significantly higher than the corresponding median value of  $\sim 0.2$  for the SDSS galaxies considered by T04. The low metallicities, low stellar masses, and high gas fractions of the objects in the lowest mass bin suggest that they are young objects just beginning to form stars; this is also indicated by the ages from the SED modeling (see Table 1). The correlations between age and other properties are discussed in detail by Erb et al. (2006a) in the context of the comparison of stellar and dynamical masses.

A consequence of the large gas fractions of the low stellar mass bins is that the baryonic mass  $M_{\text{gas}} + M_*$  of the galaxies in the sample spans a much smaller range than the stellar mass. The difference in mean stellar mass between the highest and lowest mass bins is a factor of 39, while the difference in mean baryonic mass between the same bins is only a factor of 6. This relatively small range in baryonic mass also limits our ability to detect differential loss of metals as a function of mass. We show the variations of  $M_*$ ,  $M_{\text{gas}}$ , and  $M_{\text{bar}}$  with metallicity in Figure 8. Notably, the increase in baryonic mass along our sequence of six bins is driven almost entirely by the increase in stellar mass, while the gas mass remains relatively constant. This is because the average SFR and  $H\alpha$  luminosity vary much less than the average stellar mass across the bins; in particular, galaxies in the lowest mass bin have higher than average SFRs. This may seem contrary to the trend of increasing SFR at brighter  $K$  magnitudes found by Reddy et al. (2005), but we account for the difference both because we are less likely to detect  $H\alpha$  emission in  $K$ -faint galaxies unless it is especially strong (as discussed by Erb et al. 2006a, 2006b) and because galaxies are more likely to be detected in the  $K$  band if they have high SFRs. Reddy et al. (2006) also show that the correlation of stellar mass and bolometric luminosity is relatively weak and that low stellar mass galaxies span a wide range in  $L_{\text{bol}}$ . For present purposes, there is a lower limit on the gas masses we are able to detect corresponding to our  $H\alpha$  flux limit, and the absence of galaxies with both low stellar masses and low gas masses is probably a selection effect. We also note that the dynamical masses derived from the  $H\alpha$  line widths are a better match to the baryonic masses than the stellar masses, with  $M_{\text{dyn}} \gg M_*$  for the objects in the lowest mass bin. Erb et al. (2006a) discuss this comparison in detail.

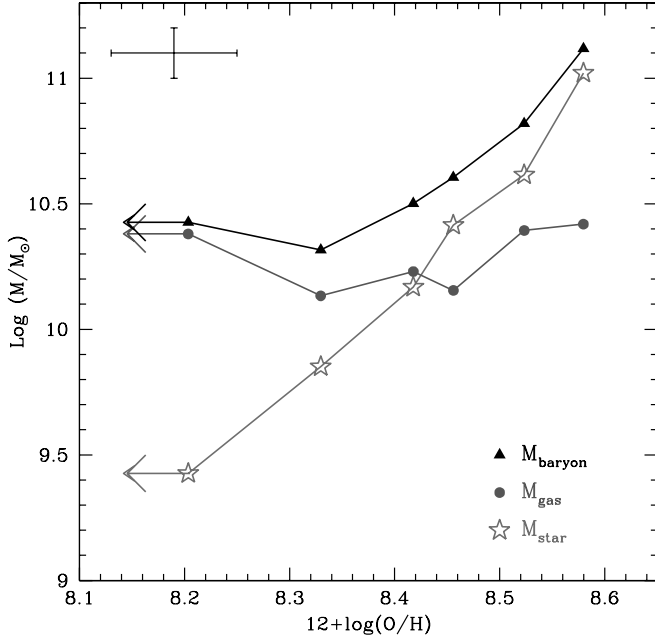


FIG. 8.—Variation of stellar, gas, and baryonic mass in each of the six bins with metallicity. Across the observed range in oxygen abundance, stellar mass increases strongly, baryonic mass increases weakly, and gas mass remains approximately constant. We thus see an increase in metallicity with decreasing gas fraction. The strong correlation between stellar mass and age means that metallicity also increases with the age of the stellar population. The error bars at upper left show typical uncertainties in gas and baryonic masses and metallicities; uncertainties in stellar masses are smaller. [See the electronic edition of the *Journal* for a color version of this figure.]

The mean gas fractions and effective yields for each mass bin are given in Table 2. In contrast to the SDSS sample, there is no decline in the effective yield with decreasing baryonic mass; in fact, we see the opposite, as  $y_{\text{eff}}$  increases slightly with decreasing mass. The filled circles in each panel of Figure 9 show the metallicity  $Z$  in each mass bin plotted against the mean gas fraction  $\mu$ , with decreasing gas fraction from left to right to show the increase in  $Z$  as  $\mu$  declines. We first discuss the data in the context of the closed-box model with no inflows or outflows. For such a model, lines of constant yield are curves given by equation (2). The uppermost (black) line on each of the three plots

shows the variation of  $Z$  with gas fraction in this model for three different values of the true yield  $y$ ; from left to right we show  $y$  equal to our observed effective yield  $y_{\text{eff}} = 0.008 - 0.6 Z_{\odot}$ , solar yield  $y = 0.0126$ , and a supersolar yield  $y = 1.5 Z_{\odot}$ . The points are formally consistent with the  $y \sim 0.6 Z_{\odot}$  model in the left panel, as the measured value of  $y_{\text{eff}}$  in each bin is within  $1 \sigma$  of the weighted mean  $y_{\text{eff}} = 0.008$  in four of the bins and within  $2 \sigma$  of this value in the fifth (we do not consider the lowest mass bin, for which we find only an upper limit on  $Z$  and  $y_{\text{eff}}$ ); however, the systematic tendency of the lower mass points to fall above the black line indicates that the model is not a good fit. A  $\chi^2$  test confirms this, giving a value of 9.

Given the ubiquitous signature of galactic-scale outflows in the kinematics of the  $z \sim 2$  galaxies, we now consider the effects of such outflows on the metallicities of the galaxies. To address this question, we modify the simple model to include gas outflow at rate  $\dot{M} M_{\odot} \text{ yr}^{-1}$ , which is a fraction  $f$  of the SFR; note that in this model the outflow rate does not directly depend on the mass of the galaxy. We assume that the metallicity of the outflowing gas is the same as the metallicity of the gas that remains in the H II regions in the galaxy. It can then be shown that the metallicity is given by

$$Z = y(1+f)^{-1} \ln[1 + (1+f)(\mu^{-1} - 1)]. \quad (4)$$

Observations at low and high redshift suggest that  $\dot{M}$  is comparable to the SFR and may be higher (e.g., Martin 1999, 2003; Pettini et al. 2000). The lines in each panel of Figure 9 show the evolution of  $Z$  with gas fraction for, from top to bottom,  $f = 0$  (black; the closed-box model), 0.5, 1, 2, and 4. A decrease in  $y_{\text{eff}}$  with decreasing gas fraction such as we have inferred is a general feature of models with constant  $\dot{M}/\text{SFR}$ ; more specifically, from comparison with the points in each panel, it is apparent that the data are best matched by a model with supersolar yield  $y \sim 1.5 Z_{\odot}$  and a high outflow rate  $\dot{M} \sim 4 \times \text{SFR}$  (the bottom line in the right panel). We caution that the uncertainties in both metallicity and gas fraction are too large to distinguish between the above models with confidence; nevertheless, the best-fitting model is not implausible, as we discuss below.

An outflow rate of several times the SFR has already been suggested by observations of the lensed Lyman break galaxy (LBG)

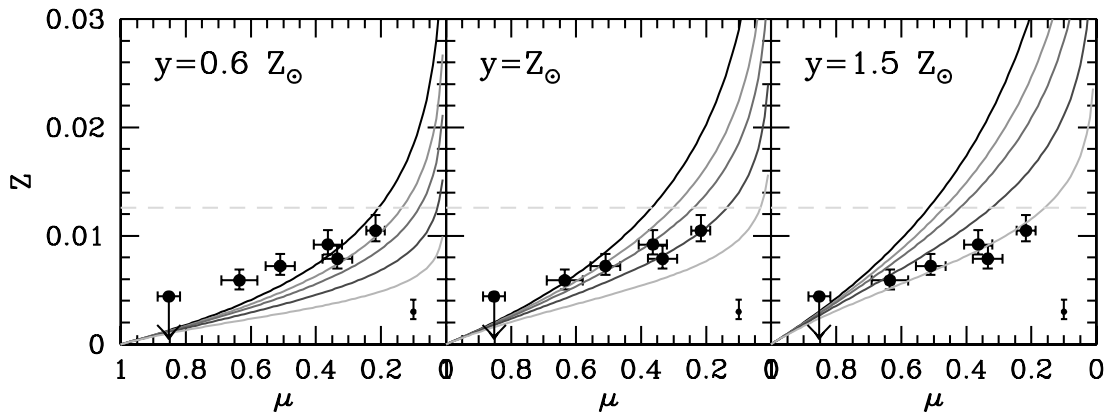


FIG. 9.—Filled circles: Mean metallicity  $Z$  in each mass bin, plotted against the mean gas fraction  $\mu$ . Gas fraction decreases from left to right, to show the increase in metallicity with decreasing gas fraction. The lines illustrate the variation of  $Z$  with  $\mu$  for three different values of the true yield: from left to right,  $y = 0.6$  (our observed effective yield), 1, and  $1.5 Z_{\odot}$ . Within each panel, the lines show varying values of the mass outflow rate  $\dot{M}$ , parameterized as a fraction  $f$  of the SFR. From top to bottom in each panel, the lines correspond to  $f = 0$  (black; the closed-box model), 0.5, 1, 2, and 4. The data are best matched by a model with supersolar yield and an outflow rate  $\dot{M} \sim 4 \times \text{SFR}$ . The horizontal dashed gray line corresponds to solar metallicity, and the small error bar in the lower right corner of each panel shows the systematic uncertainty in the metallicity calibration. [See the electronic edition of the *Journal* for a color version of this figure.]

MS 1512–cB58; Pettini et al. (2002) found a value of  $\dot{M}$  higher than the SFR using conservative assumptions for the size of the galaxy and outflow velocity. This scenario is also in good agreement with observations and predictions of the enrichment of the intracluster medium (ICM). The ICM contains several times more mass in baryons than is found in the cluster galaxies themselves; thus, it contains most of the metals in clusters as well (e.g., Mushotzky & Loewenstein 1997). Recent models for the enrichment of the ICM (De Lucia et al. 2004; Nagashima et al. 2005) use feedback from supernovae to transport metals out of galaxies, finding that enrichment of the ICM occurs at high redshift, with of order half of the metals produced by  $z \sim 2$ , and that half or more of the metals in the ICM are produced in massive galaxies ( $L \gtrsim L_*$ , Nagashima et al. 2005; or baryonic mass  $\gtrsim 10^{10} h^{-1} M_\odot$ , De Lucia et al. 2004). These models also require either a top-heavy IMF in starbursts or a significantly enhanced yield in order to reproduce the observed metal abundances in clusters; some such variant IMF is a feature of many models of star formation and chemical enrichment in clusters (e.g., Zepf & Silk 1996; Finoguenov et al. 2003; Tornatore et al. 2004). The supersolar yield suggested by the best-fitting model shown in Figure 9 may support such a scenario, although we emphasize that our data do not require an IMF with more high-mass stars than the Chabrier (2003) IMF we use here. Such an IMF, when combined with estimates of yields from stellar nucleosynthesis (e.g., Meynet & Maeder 2002), results in a metal yield that is solar or higher.

### 5.1. Redshift Evolution of the Mass-Metallicity Relation

With estimates of the gas fractions in hand, we now return to the  $\sim 0.3$  dex offset between the local and  $z \sim 2$  mass-metallicity relations. There are several possible reasons for this factor of  $\sim 2$  difference. The magnitude of the offset itself is not well determined because of the uncertainty in our metallicity scale indicated by the offsets between the  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  ratios relative to the SDSS galaxies seen in Figure 1. The effects of these offsets cannot be quantified reliably without detailed modeling of a larger sample of galaxies with all four emission lines, but meanwhile we can use several techniques to estimate the possible uncertainty in metallicity that may result. We have another method of determining the metallicity of the four galaxies with all four emission lines in the O3N2 index of Pettini & Pagel (2004):  $\text{O3N2} \equiv \log [([\text{O III}] \lambda 5007/\text{H}\beta)/([\text{N II}] \lambda 6583/\text{H}\alpha)]$ . Pettini & Pagel (2004) find that O3N2 varies with oxygen abundance as

$$12 + \log (\text{O}/\text{H}) = 8.73 - 0.32 \times \text{O3N2}. \quad (5)$$

Using this calibration to determine the metallicities of the four galaxies with all four lines, we find oxygen abundances that are 0.17 dex lower on average than those determined using  $[\text{N II}]/\text{H}\alpha$ . This is one plausible estimate of the metallicity uncertainty due to the differing physical conditions in the  $\text{H II}$  regions of the  $z \sim 2$  galaxies. We can obtain another by assuming that the observed shift in the  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  diagram is entirely in the  $[\text{N II}]/\text{H}\alpha$  direction; then the relatively shallow slope of the relationship between  $[\text{N II}]/\text{H}\alpha$  and  $(\text{O}/\text{H})$  leads to a maximum offset in metallicity of  $\sim 0.3$  dex. The effect of an offset in the  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  diagram on metallicity determinations is considered in detail by Shapley et al. (2005a), and we refer the reader to § 5.3 of that paper for a discussion. These authors consider the effect of a harder ionizing spectrum in high-redshift galaxies

and conclude that this could plausibly lead to a factor of  $\sim 1.5$ –2 uncertainty in metallicity.

Another possible issue is that the SDSS metallicities are biased toward the innermost regions of the galaxies observed, where the spectrograph fibers were positioned. According to a recent reanalysis of this effect by Ellison & Kewley (2005), this fact alone can account for a  $\sim 0.15$  dex offset between nuclear and global (i.e., integrated over the whole galaxy) metallicities. However, this effect may be mitigated by the fact that our global spectra are also biased toward the galaxies' central regions, which have the highest surface brightness. We conclude that the uncertainty in the metallicity offset between the  $z \sim 2$  and local galaxies is approximately a factor of 2, about the same as the offset itself.

There are reasons to believe that there may be a real evolutionary offset in the mass-metallicity relation, and we proceed with the discussion under this assumption. As described above, we find that the galaxies in our sample have an average gas fraction more than 2 times higher than the galaxies studied by T04, placing them at an earlier stage in the process of converting their gas to stars. Given this less evolved state, it is not surprising that they should have lower gas-phase metallicities. Redshift evolution of the mass-metallicity relation has also recently been investigated by Savaglio et al. (2005) with a sample of galaxies at  $0.4 < z < 1$ , finding that a galaxy of a given stellar mass tends to have lower metallicity at  $z \sim 0.7$  than at  $z \sim 0$ . Using the SDSS relation of T04, their sample at  $z \sim 0.7$ , and the  $z \sim 2$  galaxies presented by Shapley et al. (2004), they develop an empirical model for the redshift evolution of the relation. This model predicts metallicities an average of  $\sim 0.2$  dex higher than we observe, although at least some of this offset is likely to be due to the different metallicity indicators used (it is not possible to place our sample and that of Savaglio et al. [2005] on the same metallicity scale, as an accurate conversion between the two indicators used has not yet been established for the samples in question). Such a shift to lower metallicities (or higher masses) with increasing redshift can qualitatively be explained via a scenario in which star formation takes place over a more protracted period in lower mass galaxies. This scenario is consistent with our results; we find that the more massive galaxies in our sample have smaller gas fractions and are thus likely to exhaust their gas supply before the less massive galaxies.

The Savaglio et al. (2005) model also predicts a steeper relation at  $z \sim 2$  than we have observed; this is a consequence of their assumption that the shape of the T04 relation remains unchanged but shifts to higher masses, whereas we find that our data are well approximated by the T04 relation shifted to lower metallicities at the same stellar mass. In physical terms, the slope of the stellar mass-metallicity relation depends on both the yield and the presence or absence of outflows (and inflows). As shown in Figure 9, a closed-box model such as that assumed by Savaglio et al. (2005) results in a steep relation between metallicity and gas fraction (and stellar mass), while the  $z \sim 2$  data are better described by the shallower slope of a model with significant outflows.

While comparisons between the metallicities of star-forming galaxies at high and low redshift are of obvious interest, it is very likely that our sample and that of T04 do not form an evolutionary sequence. The likely descendants of the  $z \sim 2$  sample can be identified by comparing their clustering properties, evolved to  $z \sim 0$ , with those of objects in the local universe; such a comparison shows that early-type galaxies in the SDSS match the clustering properties of the  $z \sim 2$  sample, while the later type

star-forming objects studied by T04 are too weakly clustered to be the descendants of the  $z \sim 2$  population (Adelberger et al. 2005b). In this sense it is more relevant to compare the metallicities of the  $z \sim 2$  galaxies with local early-type objects than with star-forming galaxies at  $z \sim 0$ , although such a comparison would undoubtedly be complicated by systematic offsets resulting from the very different methods used to determine metal abundances in elliptical galaxies (usually absorption features in the integrated spectra of old stellar populations). Broadly speaking, however, most early-type galaxies in the SDSS with  $M_* > 10^{10} M_\odot$  have metallicities ranging from  $Z \sim 0.6$  to  $\sim 1.5 Z_\odot$  (Gallazzi et al. 2005); these results are consistent with the probable final stellar masses and metallicities of the  $z \sim 2$  galaxies.

Substantial uncertainties in the mass-metallicity relation at  $z \sim 2$  remain, of course. It is important to confirm the trend in metallicity with additional measurements and with abundance indicators that use a broader set of emission lines. The absolute values of the abundances in the sample are uncertain, and the improved understanding of the galaxies' physical conditions that will result from a larger sample of lines will be essential for determining this absolute scale. Furthermore, our derivation of the gas masses and gas fractions is indirect and assumes that the Schmidt law takes the same form at  $z \sim 2$  as in the local universe. This has not yet been tested, although the one similar galaxy with a direct measurement of the gas mass, the lensed  $z = 2.7$  LBG MS 1512–cB58, appears to be consistent with the local Schmidt law (Baker et al. 2004). As long as some form of the Schmidt law applies, our results will be qualitatively similar. A related question concerns the appropriateness of the gas masses derived from the Schmidt law. These represent only the gas associated with current star formation and are therefore almost certainly an underestimate of the total gas masses. In a typical disk galaxy today,  $\sim 40\%$  of the gas mass is not included by the Schmidt law (Martin & Kennicutt 2001); this fraction could plausibly be higher in the young starbursts in our sample. It is not clear how or if this gas affects the metal enrichment and star formation. Given our lack of information on this gas, we do not consider it. T04 discussed these questions of gas masses derived from the Schmidt law in somewhat more detail but arrived at similar conclusions.

Our simple model for the effect of winds on gas-phase abundances assumes that the metallicity of the outflowing gas is the same as the observed abundances in the galaxy; this may not be correct. Metal-enhanced hot winds have been observed in X-rays in local starbursts; in dwarf galaxies, such winds may carry away nearly all of the oxygen produced by the burst (Martin et al. 2002). Such a metal-enhanced outflow would increase metal losses for a given outflow rate; thus, it may be possible to produce the relative shallow relationship between  $Z$  and  $\mu$  that we have observed with a mass outflow rate smaller than the value  $\dot{M} \sim 4 \times \text{SFR}$  we find above.

We also consider the possibility, discussed in §§ 3.1 and 4, that we have underestimated the stellar masses of the galaxies with the smallest  $M_*$  by a factor of a few. This would primarily affect the objects in the lowest mass bins and would result in a decrease in the gas fraction and a decrease in the effective yield. A factor of 3 increase in the stellar mass of the galaxies in the two lowest mass bins would result in a decrease in  $y_{\text{eff}}$  of approximately a factor of 2 for those bins; this is enough to remove the observed trend of decreasing  $y_{\text{eff}}$  with increasing stellar mass and results instead in an approximately constant effective yield, roughly consistent with the closed-box model. Given the significant evidence for strong outflows in these galaxies, however, we regard the best-fit model obtained above as more plausible.

Finally, the sample of  $z \sim 2$  galaxies considered here is by no means complete. The UV selection technique is most likely to miss galaxies that are very dusty or have little current star formation; Reddy et al. (2005) show that there is relatively little overlap ( $\sim 12\%$ ) between the UV-selected sample and  $z \sim 2$  galaxies selected by their  $J - K_s > 2.3$  colors (distant red galaxies [DRGs]; Franx et al. 2003). This technique selects galaxies with a strong optical break, due to either an evolved stellar population or strong reddening (e.g., Papovich et al. 2006). Little is known about the metallicities of the DRGs. A few rest-frame optical spectra of brighter DRGs with  $K < 20$  have been presented by van Dokkum et al. (2004, 2005); these suggest approximately solar metallicities, although they frequently show evidence for AGNs or shock ionization. It will be extremely difficult to measure metallicities for passively evolving galaxies at high redshift, since they lack strong emission lines. Massive galaxies that have consumed or expelled most of their gas might be expected to be among the most metal-rich objects at high redshift. Dusty, rapidly star-forming galaxies are likely to have large gas masses, but their metallicities will depend on how many generations of stars have enriched the gas. There is probably a range of metallicities within this population. Some confirmation of this can be found in Figure 15 of Reddy et al. (2006), who plot inferred gas fraction against stellar mass for near-IR–selected galaxies, as well as for UV-selected galaxies similar to those considered here. All samples considered follow the same trend of decreasing gas fraction with increasing stellar mass, with the DRGs at the high-mass and low gas fraction end of the spectrum. This suggests that galaxies selected by near-IR techniques may also follow a similar mass-metallicity relation to that discussed here, with a possible extension to higher metallicities for the oldest and most massive objects. It will be interesting to see if this proves to be true when larger samples of metallicity measurements become available.

In summary, the  $z \sim 2$  galaxies show a strong trend in oxygen abundance over a range of only a factor of  $\sim 6$  in baryonic mass. The effective yield increases slightly with decreasing mass, rather than decreasing as would be expected if low-mass galaxies lost a larger fraction of their metals to outflows. We conclude that the mass-metallicity relation at high redshift is primarily a product of varying stages of galaxy evolution, caused by the increase in metallicity as gas is converted into stars and metals are returned to the remaining gas. It may be modulated by metal loss from strong outflows in galaxies of all masses, which results in a shallower increase of metallicity with decreasing gas fraction than predicted by the closed-box model. There is no evidence for preferential loss of metals from lower mass galaxies, as has been inferred in the local universe, although the small mass range spanned by our sample limits our sensitivity to such an effect.

## 6. SUMMARY AND CONCLUSIONS

We have used composite  $\text{H}\alpha + [\text{N II}]$  spectra of 87 star-forming galaxies at  $z \gtrsim 2$ , divided into six bins by stellar mass, to study the correlation between stellar mass and metallicity at high redshift. Our conclusions are summarized as follows:

1. There is a strong correlation between stellar mass and metallicity at  $z \gtrsim 2$ , as the quantity  $12 + \log(\text{O}/\text{H})$  increases monotonically from  $< 8.2$  for galaxies with  $\langle M_* \rangle = 2.7 \times 10^9 M_\odot$  to  $8.6$  for galaxies with  $\langle M_* \rangle = 1.0 \times 10^{11} M_\odot$ . The relation is offset by  $\sim 0.3$  dex from the local mass-metallicity relation, in the sense that galaxies of a given stellar mass have lower metallicities at high redshift. The absolute values of the oxygen abundances are

uncertain, but the trend is unlikely to be due to systematic effects such as AGN contamination, hidden stellar mass, or variations in the ionization parameter of the H II regions.

2. Rest-frame  $B$ -band luminosity and metallicity are not significantly correlated. The  $z \sim 2$  galaxies are systematically brighter than local star-forming galaxies spanning the same range in stellar mass, showing that they have smaller mass-to-light ratios  $M/L$ . At a given metallicity, local galaxies are  $\sim 3$  mag fainter than the  $z \sim 2$  sample. The known large scatter in the rest-frame optical  $M/L$  at  $z \sim 2$  accounts for the lack of correlation between luminosity and metallicity and indicates that the correlation with stellar mass is (as expected) more fundamental.

3. We use the Schmidt law to estimate the gas masses and gas fractions of the  $z \sim 2$  galaxies, finding that the gas fraction increases substantially with decreasing stellar mass. The lowest mass bin in the sample has a mean gas fraction of 85%, while the highest stellar mass bin has a fraction of 20%. Our median gas fraction is  $\sim 50\%$ , as compared to  $\sim 20\%$  in local star-forming galaxies. Galaxies with low stellar masses, low metallicities, and high gas fractions also have young ages. A consequence of the trend in gas fraction with stellar mass is a much smaller range in baryonic mass (a factor of 6) than stellar mass (a factor of 39) across the sample.

4. The observed metallicities and gas fractions allow an estimate of the effective yield in each mass bin. In contrast to the results of a similar study in the local universe, we find no decrease in  $y_{\text{eff}}$  at low masses; instead,  $y_{\text{eff}}$  increases slightly with decreasing mass. Qualitatively, such an increase is a feature of models in which galaxies of all masses lose metals from outflows. More quantitatively, comparison with simple models shows that the observed variation of metallicity with gas fraction is best described by a model with supersolar yield and an outflow rate  $\sim 4$  times higher than the SFR. We conclude that the mass-metallicity relation at high redshift is caused by the increase in metallicity as gas is converted to stars and may be modulated by strong outflows in galaxies of all masses. Our ability to detect

differential metal loss as a function of mass is limited by the small range of baryonic masses spanned by the galaxies in the sample, but there is no evidence for preferential loss of metals from low-mass galaxies as is suggested locally.

Much remains to be done in order to improve on the substantial uncertainties inherent in the present work. Independent measurements of the gas masses of galaxies at high redshift, although very difficult, are essential to determine whether our derived gas fractions and effective yields are valid. Additional metallicity measurements, based on other indicators that use a wider set of emission lines, are needed to confirm the trend with stellar mass revealed by the N2 index, to provide a better understanding of the physical conditions in the H II regions, and to establish a secure absolute calibration of the metallicity scale. Further investigation of the question of differential metal loss from outflows at high redshift requires observations of fainter galaxies with smaller potential wells, to expand the dynamic range in baryonic mass. We anticipate that all of these measurements will greatly increase our understanding of the interplay between stars and gas, within and outside of galaxies, at high redshift.

We thank the anonymous referee for useful comments and a careful reading of the paper. We also thank Mike Dopita, Lisa Kewley, and Evan Skillman for illuminating discussions and the staffs of the Keck and Palomar observatories for their assistance with the observations. C. C. S., D. K. E., and N. A. R. have been supported by grant AST 03-07263 from the US National Science Foundation and by the David and Lucile Packard Foundation. A. E. S. acknowledges support from the Miller Institute for Basic Research in Science, and K. L. A. from the Carnegie Institution of Washington. Finally, we wish to extend special thanks to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests. Without their generous hospitality, most of the observations presented herein would not have been possible.

#### REFERENCES

- Adelberger, K. L., Shapley, A. E., Steidel, C. C., Pettini, M., Erb, D. K., & Reddy, N. A. 2005a, *ApJ*, 629, 636
- Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N. A., & Erb, D. K. 2005b, *ApJ*, 619, 697
- Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, *ApJ*, 584, 45
- Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, *A&A*, 417, 751
- Baker, A. J., Tacconi, L. J., Genzel, R., Lehnert, M. D., & Lutz, D. 2004, *ApJ*, 604, 125
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Barnby, P., et al. 2004, *ApJS*, 154, 97
- Bell, E. F., & de Jong, R. S. 2000, *MNRAS*, 312, 497
- Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2003, *ApJ*, 599, 38
- Bresolin, F., Schaerer, D., González Delgado, R. M., & Stasińska, G. 2005, *A&A*, 441, 981
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, *ApJ*, 533, 682
- Chabrier, G. 2003, *PASP*, 115, 763
- Cortese, L., et al. 2006, *ApJ*, 637, 242
- Dekel, A., & Woo, J. 2003, *MNRAS*, 344, 1131
- De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, *MNRAS*, 349, 1101
- Denicoló, G., Terlevich, R., & Terlevich, E. 2002, *MNRAS*, 330, 69
- Dopita, M. A. 2005, in *AIP Conf. Proc.* 761, *The Spectral Energy Distributions of Gas-rich Galaxies: Confronting Models with Data*, ed. C. C. Popescu & R. J. Tuffs (New York: AIP), 203
- Edmunds, M. G. 1990, *MNRAS*, 246, 678
- Ellison, S. L., & Kewley, L. J. 2005, preprint (astro-ph/0508627)
- Ellison, S. L., Songaila, A., Schaye, J., & Pettini, M. 2000, *AJ*, 120, 1175
- Erb, D. K., Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., Hunt, M. P., Moorwood, A. F. M., & Cuby, J. 2003, *ApJ*, 591, 101
- Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., & Adelberger, K. L. 2006a, *ApJ*, in press (astro-ph/0604041)
- . 2006b, *ApJ*, in press (astro-ph/0604388)
- Finoguenov, A., Burkert, A., & Böhringer, H. 2003, *ApJ*, 594, 136
- Franx, M., et al. 2003, *ApJ*, 587, L79
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, 362, 41
- Garnett, D. R. 2002, *ApJ*, 581, 1019
- Garnett, D. R., Kennicutt, R. C., Jr., & Bresolin, F. 2004, *ApJ*, 607, L21
- Garnett, D. R., Shields, G. A., Skillman, E. D., Sagan, S. P., & Dufour, R. J. 1997, *ApJ*, 489, 63
- Heckman, T. M., Armus, L., & Miley, G. K. 1990, *ApJS*, 74, 833
- Hernquist, L., & Springel, V. 2003, *MNRAS*, 341, 1253
- Kauffmann, G., et al. 2003, *MNRAS*, 346, 1055
- Kennicutt, R. C., Jr. 1998, *ApJ*, 498, 541
- Kennicutt, R. C., Jr., Bresolin, F., & Garnett, D. R. 2003, *ApJ*, 591, 801
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, 142, 35
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001a, *ApJ*, 556, 121
- Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001b, *ApJS*, 132, 37
- Kobulnicky, H. A., Kennicutt, R. C., Jr., & Pizagno, J. L. 1999, *ApJ*, 514, 544
- Kobulnicky, H. A., & Kewley, L. J. 2004, *ApJ*, 617, 240
- Kobulnicky, H. A., & Koo, D. C. 2000, *ApJ*, 545, 712
- Kobulnicky, H. A., et al. 2003, *ApJ*, 599, 1006
- Lamareille, F., Mouhcine, M., Contini, T., Lewis, I., & Maddox, S. 2004, *MNRAS*, 350, 396
- Larson, R. B. 1974, *MNRAS*, 169, 229
- Lehnert, M. D., & Heckman, T. M. 1996, *ApJ*, 472, 546

- Leitherer, C., Leão, J. R. S., Heckman, T. M., Lennon, D. J., Pettini, M., & Robert, C. 2001, *ApJ*, 550, 724
- Leitherer, C., et al. 1999, *ApJS*, 123, 3
- Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, *A&A*, 80, 155
- Liang, Y. C., Hammer, F., Flores, H., Elbaz, D., Marcillac, D., & Cesarsky, C. J. 2004, *A&A*, 423, 867
- Lilly, S. J., Carollo, C. M., & Stockton, A. N. 2003, *ApJ*, 597, 730
- Maier, C., Meisenheimer, K., & Hippelein, H. 2004, *A&A*, 418, 475
- Martin, C. L. 1999, *ApJ*, 513, 156
- . 2003, in *Rev. Mex. AA Conf. Ser.*, 17, 56
- Martin, C. L., & Kennicutt, R. C., Jr. 2001, *ApJ*, 555, 301
- Martin, C. L., Kobulnicky, H. A., & Heckman, T. M. 2002, *ApJ*, 574, 663
- Mas-Hesse, J. M., Kunth, D., Tenorio-Tagle, G., Leitherer, C., Terlevich, R. J., & Terlevich, E. 2003, *ApJ*, 598, 858
- McGaugh, S. S., & de Blok, W. J. G. 1997, *ApJ*, 481, 689
- McGaugh, S. S., Schombert, J. M., Bothun, G. D., & de Blok, W. J. G. 2000, *ApJ*, 533, L99
- McLean, I. S., et al. 1998, *Proc. SPIE*, 3354, 566
- Meynet, G., & Maeder, A. 2002, *A&A*, 390, 561
- Murray, N., Quataert, E., & Thompson, T. A. 2005, *ApJ*, 618, 569
- Mushotzky, R. F., & Loewenstein, M. 1997, *ApJ*, 481, L63
- Naab, T., & Ostriker, J. P. 2006, *MNRAS*, 366, 899
- Nagamine, K., Springel, V., Hernquist, L., & Machacek, M. 2004, *MNRAS*, 350, 385
- Nagashima, M., Lacey, C. G., Baugh, C. M., Frenk, C. S., & Cole, S. 2005, *MNRAS*, 358, 1247
- Pagel, B. E. J., & Patchett, B. E. 1975, *MNRAS*, 172, 13
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, 559, 620
- Papovich, C., et al. 2006, *ApJ*, 640, 92
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, 348, L59
- Pettini, M., Rix, S. A., Steidel, C. C., Adelberger, K. L., Hunt, M. P., & Shapley, A. E. 2002, *ApJ*, 569, 742
- Pettini, M., Shapley, A. E., Steidel, C. C., Cuby, J., Dickinson, M., Moorwood, A. F. M., Adelberger, K. L., & Giavalisco, M. 2001, *ApJ*, 554, 981
- Pettini, M., Steidel, C. C., Adelberger, K. L., Dickinson, M., & Giavalisco, M. 2000, *ApJ*, 528, 96
- Raimann, D., Storch-Bergmann, T., Bica, E., Melnick, J., & Schmitt, H. 2000, *MNRAS*, 316, 559
- Reddy, N. A., Erb, D. K., Steidel, C. C., Shapley, A. E., Adelberger, K. L., & Pettini, M. 2005, *ApJ*, 633, 748
- Reddy, N. A., Steidel, C. C., Fadda, D., Yan, L., Pettini, M., Shapley, A. E., Erb, D. K., & Adelberger, K. L. 2006, *ApJ*, 644, 792
- Rix, S. A., Pettini, M., Leitherer, C., Bresolin, F., Kudritzki, R., & Steidel, C. C. 2004, *ApJ*, 615, 98
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Salzer, J. J., Lee, J. C., Melbourne, J., Hinz, J. L., Alonso-Herrero, A., & Jangren, A. 2005, *ApJ*, 624, 661
- Savaglio, S., et al. 2005, *ApJ*, 635, 260
- Schmidt, M. 1963, *ApJ*, 137, 758
- Searle, L., & Sargent, W. L. W. 1972, *ApJ*, 173, 25
- Shapley, A. E., Coil, A. L., Ma, C.-P., & Bundy, K. 2005a, *ApJ*, 635, 1006
- Shapley, A. E., Erb, D. K., Pettini, M., Steidel, C. C., & Adelberger, K. L. 2004, *ApJ*, 612, 108
- Shapley, A. E., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, *ApJ*, 562, 95
- Shapley, A. E., Steidel, C. C., Erb, D. K., Reddy, N. A., Adelberger, K. L., Pettini, M., Barmby, P., & Huang, J. 2005b, *ApJ*, 626, 698
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *ApJ*, 588, 65
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, *ApJ*, 606, 92
- Skillman, E. D., Kennicutt, R. C., Jr., & Hodge, P. W. 1989, *ApJ*, 347, 875
- Smail, I., Chapman, S. C., Ivison, R. J., Blain, A. W., Takata, T., Heckman, T. M., Dunlop, J. S., & Sekiguchi, K. 2003, *MNRAS*, 342, 1185
- Stasińska, G. 2005, *A&A*, 434, 507
- Steidel, C. C., Shapley, A. E., Pettini, M., Adelberger, K. L., Erb, D. K., Reddy, N. A., & Hunt, M. P. 2004, *ApJ*, 604, 534
- Storch-Bergmann, T., Calzetti, D., & Kinney, A. L. 1994, *ApJ*, 429, 572
- Strickland, D. K., Heckman, T. M., Colbert, E. J. M., Hoopes, C. G., & Weaver, K. A. 2004, *ApJ*, 606, 829
- Tornatore, L., Borgani, S., Matteucci, F., Recchi, S., & Tozzi, P. 2004, *MNRAS*, 349, L19
- Tremonti, C. A., et al. 2004, *ApJ*, 613, 898 (T04)
- van den Bergh, S. 1962, *AJ*, 67, 486
- van Dokkum, P. G., Kriek, M., Rodgers, B., Franx, M., & Puxley, P. 2005, *ApJ*, 622, L13
- van Dokkum, P. G., et al. 2004, *ApJ*, 611, 703
- Veilleux, S., & Osterbrock, D. E. 1987, *ApJS*, 63, 295
- Wilson, J. C., et al. 2003, *Proc. SPIE*, 4841, 451
- Zaritsky, D., Kennicutt, R. C., Jr., & Huchra, J. P. 1994, *ApJ*, 420, 87
- Zepf, S. E., & Silk, J. 1996, *ApJ*, 466, 114