Emission-line Metallicities from the Faint Infrared Grism Survey and VLT/MUSE

John Pharo[1] , Sangeeta Malhotra[1,2] , James Rhoads[1,2] , Lise Christensen[3] , Steven L. Finkelstein[4] , Norman Grogin[5] ,
Santosh Harish[1], Tianxing Jiang[1] , Keunho Kim[1] , Anton Koekemoer[5] , Norbert Pirzkal[5], Mark Smith[1], Huan Yang[1,6],
Andrea Cimatti[7,8] , Ignacio Ferreras[9], Nimish Hathi[5] , Pascale Hibon[10] , Gerhardt Meurer[11], Goeran Oestlin[12],
Anna Pasquali[13], Russell Ryan[5], Amber Straughn[2], and Rogier Windhorst[1]

[1] School of Earth & Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA
[2] NASAs Goddard Space Flight Center, Astrophysics Science Division, Code 660, Greenbelt, MD 20771, USA
[3] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark
[4] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA
[5] Space Telescope Science Institute, Baltimore, MD 21218, USA
[6] CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China,
People's Republic of China
[7] Department of Physics and Astronomy (DIFA), University of Bologna, Via Gobetti 93/2, I-40129, Bologna, Italy
[8] INAFOsservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125, Firenze, Italy
[9] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
[10] ESO, Alonso de Cordova 3107, Santiago, Chile
[11] International Centre for Radio Astronomy Research, The University of Western Australia, Crawley WA 6009, Australia
[12] Stockholm University, Stockholm, SE-10691, Sweden
[13] Astronomisches Rechen-Institut, Zentrum fuer Astronomie, Universitaet Heidelberg, Moenchhofstrasse 12-14, D-69120 Heidelberg, Germany

Abstract

We derive direct-measurement gas-phase metallicities of $7.4 < 12 + \log(O/H) < 8.4$ for 14 low-mass emission-line galaxies at $0.3 < z < 0.8$ identified in the Faint Infrared Grism Survey. We use deep slitless G102 grism spectroscopy of the Hubble Ultra Deep Field, dispersing light from all objects in the field at wavelengths between 0.85 and 1.15 μm. We run an automatic search routine on these spectra to robustly identify 71 emission-line sources, using archival data from Very Large Telescope (VLT)/Multi-Unit Spectroscopic Explorer (MUSE) to measure additional lines and confirm redshifts. We identify 14 objects with $0.3 < z < 0.8$ with measurable [O III] $\lambda4363$ Å emission lines in matching VLT/MUSE spectra. For these galaxies, we derive direct electron-temperature gas-phase metallicities with a range of $7.4 < 12 + \log(O/H) < 8.4$. With matching stellar masses in the range of $10^{7.9}\,M_\odot < M_\star < 10^{10.4}\,M_\odot$, we construct a mass–metallicity (MZ) relation and find that the relation is offset to lower metallicities compared to metallicities derived from alternative methods (e.g., R_{23}, O3N2, N2O2) and continuum selected samples. Using star formation rates derived from the Hα emission line, we calculate our galaxies' position on the Fundamental Metallicity Relation, where we also find an offset toward lower metallicities. This demonstrates that this emission-line-selected sample probes objects of low stellar masses but even lower metallicities than many comparable surveys. We detect a trend suggesting galaxies with higher Specific Star Formation (SSFR) are more likely to have lower metallicity. This could be due to cold accretion of metal-poor gas that drives star formation, or could be because outflows of metal-rich stellar winds and SNe ejecta are more common in galaxies with higher SSFR.

Key words: galaxies: abundances – galaxies: evolution

1. Introduction

The identification and study of nebular emission lines in galaxies can provide insight into star formation rates (SFRs), ionization parameters, and gas-phase metallicities, among other physical parameters. The gas-phase metallicity can be related to star formation and mass growth in galaxies via the mass–metallicity (MZ) relation, an observed correlation between a galaxy's stellar mass and its gas-phase metallicity, and by the Fundamental Metallicity Relation (FMR; Lara-López et al. 2010; Mannucci et al. 2010), an empirical plane relating the metallicity and the stellar mass to the SFR.

These relations have been well-established for local star-forming galaxies (Tremonti et al. 2004), which show an increase in gas-phase metallicity as stellar mass increases from $10^{8.5}\,M_\odot$ to $10^{10.5}\,M_\odot$, after which the metallicity flattens. Further surveys have pushed the study of the relation out to higher redshifts, typically finding lower levels of metallicity out to $z \sim 3$ (Lilly et al. 2003; Maier et al. 2005; Erb et al. 2006;

Mannucci et al. 2009). For these studies, the gas-phase metallicity is often measured through empirical and theoretical strong-line ratio calibrations, such as R_{23} (Kobulnicky & Kewley 2004), N2O2 (Kewley & Dopita 2002), and O3N2 (Pettini & Pagel 2004), using [O III], [O II], and Balmer-series hydrogen lines (see Table 1 for description of ratios), or via modeling UV indicators including [C III]$\lambda1907$ (Amorín et al. 2017). However, offsets between local and high-redshift galaxies on diagnostic plots such as the Baldwin–Phillips–Terlevich diagram (Baldwin et al. 1981; Steidel et al. 2014; Sanders et al. 2015), which compares the [O III]$\lambda5007$/Hβ line ratio to the [N II]$\lambda6568$/Hα line ratio, indicate that conditions in the interstellar medium may differ at different redshifts (Kewley et al. 2013). If so, there may be undetected biases in the line ratio calibrations. Some studies have also indicated, however, that the presence of very strong emission lines is itself an indicator of low gas-phase metallicity, regardless of the redshift (Finkelstein et al. 2011; Xia et al. 2012; Yang et al. 2017). Given these uncertainties and outliers, it is necessary to

Table 1
Common Strong Line Ratios

Name	Ratio
N2	$\log([\text{N II}]\lambda 6584/\text{H}\alpha)$
O2	$\log([\text{O II}]\lambda 3727 + 3729/\text{H}\beta)$
O3	$\log([\text{O III}]\lambda 4959 + 5007/\text{H}\beta)$
R_{23}	$\log(([\text{O III}]\lambda 4959 + 5007 + [\text{O II}]\lambda 3727 + 3729)/\text{H}\beta)$
N2O2	N2–O2
O3N2	O3–N2

seek out samples of emission-line galaxies (ELGs) for which we can precisely determine the metallicity and thus better understand its relationship to the other emission properties of galaxies.

A more direct method for measuring the gas-phase metallicity involves the ratio of the auroral [O III]λ4363 Å emission line to the [O III]$\lambda\lambda$4959,5007 Å lines, which is sensitive to the electron temperature of the ionized gas (Aller 1984; Dopita & Evans 1986; Kewley & Dopita 2002; Izotov et al. 2006). A direct measurement of the electron temperature allows for the derivation of abundances with a minimum of other assumptions compared to the more common strong-emission-line diagnostics described above. For example, the R_{23} relation is double-branched, with each R_{23} value corresponding to both a high-metallicity and a low-metallicity solution, requiring additional data or assumptions to break the degeneracy. Consequently, direct-temperature-derived metallicities are more reliable (Izotov et al. 2006). This method is not always practical, as the auroral line is typically quite weak (50–100 times weaker than typical strong lines, per Sanders et al. 2017) and may require stacking spectra to get a reliable signal (Andrews & Martini 2013), but it provides more accurate metallicity measurements.

In this paper, we describe our systematic search for ELGs in one-dimensional (1D) spectra from the Faint Infrared Grism Survey (FIGS). In Section 2, we describe the survey and procedures for data collection and reduction. In Section 3, we describe the line search method and the flux measurements for confirmed ELGs. In Section 4, we detail the measurement of the gas-phase metallicity, and in Section 5 we explore the mass–metallicity relation and other properties available from our line measurements. Finally, we summarize in Section 6. For this paper, we will use $H_0 = 67.3$ km s^{-1} Mpc^{-1} and $\Omega_M = 0.315$, $\Omega_\Lambda = 0.685$ (Planck Collaboration et al. 2014). All magnitudes are given in the AB system (Oke & Gunn 1983).

2. Survey Description and Data

2.1. FIGS Observations and Spectral Extraction

2.1.1. Survey Description

The FIGS (HST/Cycle 22, ID:13779, PI S. Malhotra) used the HST WFC3-G102 (see Figure 1) infrared grism to obtain deep slitless spectroscopy of ~6000 galaxies. FIGS achieved 40-orbit depth in four fields, designated GN1, GN2, GS1 (UDF), and GS2 (HDF-PAR2; see Table 2 for coordinates of each field). Objects in each field were observed in five different 8-orbit position angles (PAs) in order to mitigate contamination of the spectra by overlapping spectra from nearby objects. Each PA covers a $2'.05 \times 2'.27$ field of view. The area of coverage in
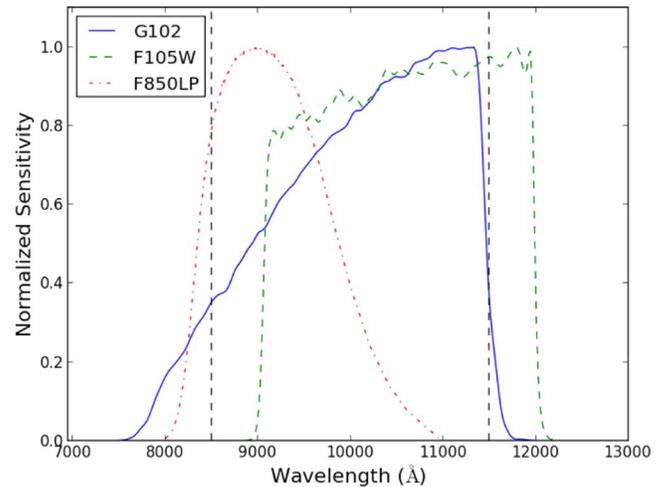
Figure 1. The sensitivity curves for the WFC3/G102 grism, as seen in Kuntschner et al. (2011) and Lee et al. (2014), and the WFC3-F105W and ACS-F850LP filters. The dashed vertical lines show the cutoffs for grism data used in the emission-line search. The curves have been normalized to their maximum sensitivity, so this plot gives the sensitivity at each wavelength in terms of its percentage of the peak sensitivity.

Table 2
A Description of the Four FIGS Fields

Field	R.A.	Decl.	Area[a]
GN1	12:36:41.467	+62:17:26.27	4.51
GN2	12:37:31.023	+62:18:26.91	5.06
GS1[b]	03:32:40.951	−27:46:47.92	4.09
GS2[c]	03:33:06.468	−27:51:21.56	4.02

Notes.
[a] Measured in arcmin^2.
[b] The HUDF.
[c] The HDF Parallel Field.

each field from which we derive the usable spectra is given in Table 2, for a total area of 17.7 square arcmin.

2.1.2. Spectral Extraction

In this paper, we used 1D spectra that were generated using the methods described in Pirzkal et al. (2017). Here we briefly summarize this process. We reduced FIGS data in a manner that loosely follows the method used for GRAPES and PEARS, previous HST grism surveys (Pirzkal et al. 2004; Xu et al. 2007; Rhoads et al. 2009; Straughn et al. 2009; Xia et al. 2012; Pirzkal et al. 2013). First, we generated a master catalog of sources from deep CANDELS survey mosaics in the F850LP filter in ACS and the F125W and F160W filters in WFC3 (approximately the z, J, and H bands; Grogin et al. 2011; Koekemoer et al. 2011). We astrometrically corrected the data to match the absolute astrometry of the GOODS catalogs. The background levels of the grism observations were estimated using a two-component model, which includes a constant Zodiacal light background as well as a varying He I light background. Individual spectra were generated using a Simulation Based Extraction approach that accounts for spectral contamination from overlapping spectra, as well as allow the use of an optimal extraction approach (Horne 1986) when generating 1D spectra from two-dimensional (2D)

spectra. We refer the reader to Pirzkal et al. (2017) for a complete description of these processes. When the extractions were complete, we had an average of ∼1700 spectra per field, with a typical 3σ detection limit of $m_{\rm F105W} = 26$ mag and an emission-line sensitivity of $10^{-17}\,{\rm erg\,cm^{-2}\,s^{-1}}$.

2.2. Optical Data

We supplemented our infrared FIGS spectra with archival high-resolution optical IFU spectra taken with the Multi-Unit Spectroscopic Explorer (MUSE) instrument (Bacon et al. 2010) from the Very Large Telescope (VLT). This expands the available spectroscopic wavelength coverage considerably, enabling confirmation of detected emission lines in FIGS via the identification of complementary emission lines at optical wavelengths. These lines also make possible the mass–metallicity results shown in Section 5. We used the publicly available IFU spectra from the MUSE Hubble Ultra Deep Survey (Bacon et al. 2017), a mosaic of nine 1×1 arcmin^2 MUSE fields in the Hubble Ultra Deep Field (HUDF). In order to extract spectra for emission-line objects in our sample, we applied the following procedure: Using the known sky coordinates for each object, 1D spectra were generated by summing up flux within a $2''$ aperture (centered on the object) at each wavelength slice, across the entire MUSE wavelength range. We extracted FIGS candidate spectra from the reduced MUSE datacube. The MUSE data wavelength coverage extends from 4752 Å to 9347 Å with a spectral resolution of 2.3 Å, though the sensitivity drops off precipitously at wavelengths lower than 5000 Å and higher than 9200 Å, so we restrict our usage to between these wavelengths. MUSE has a 3σ line sensitivity of $\sim3 \times 10^{-19}\,{\rm erg\,cm^{-2}\,s^{-1}}$, and thus is deep enough to detect the weak [O III]λ4363 line for FIGS-selected ELGs.

3. Line Identification and Flux Measurement Methods

3.1. Line Identification

Because we obtained our infrared spectra via slitless grism spectroscopy, there is no pre-selection of ELG candidates via the placement of slits or by broadband magnitude cutoffs. This has the advantage of enabling the detection of ELGs with potentially very low continuum levels, and so might allow for the study of smaller and/or fainter galaxies with nebular line emission. However, this does require an efficient method for selecting ELG candidates from the total sample of FIGS objects. In order to search the ∼6000 FIGS spectra for emission lines, we developed a code to automatically search for and identify peaks in a 1D spectrum.

First, the continuum flux needs to be estimated at each wavelength element. The G102 grism measures the spectrum every 24.5 Å, and we use the spectrum from 8500 to 11500 Å. The algorithm iterates over each wavelength element in a given spectrum, estimating the continuum flux at that wavelength and subtracting it. This estimation is accomplished via a median-flux filter, where, given a number of wavelength elements for the width of a prospective line, the algorithm measures the flux in a number of elements outside the guessed line width in both the blue and red wavelength directions. The median flux of all of these points is assumed to represent the continuum there and is subtracted from that point's flux. This serves to estimate the local value of the continuum while avoiding the influence of the line flux itself or of other features or changes in the
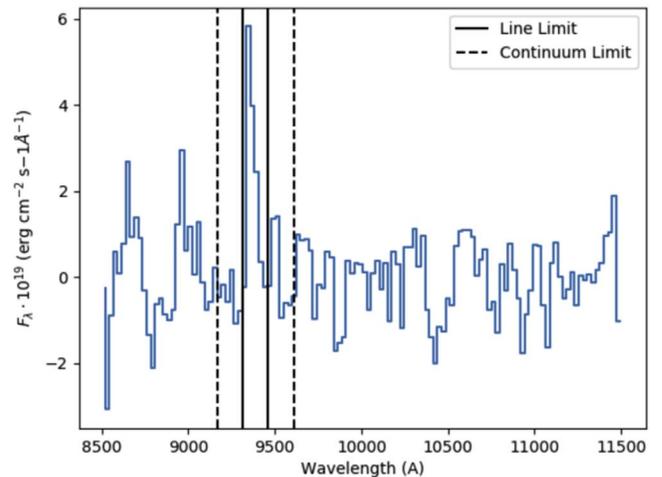
Figure 2. An example of the line-finding routine. This plot shows the continuum-subtracted flux for one PA of FIGS object GS1-2375. This shows an iteration of the line-finding routine when centered at 9388 Å. The routine sums the flux of the pixels within the solid black lines, which is considered to be the candidate line flux. Then, the continuum flux is estimated from the median flux of the pixels between the solid and dashed lines. This continuum is subtracted from the line flux, and the S/N is calculated.

spectrum. See Figure 2 for an example continuum-subtracted spectrum.

Next, we calculate the signal-to-noise ratio (S/N) at each wavelength with the residual flux and the flux error (determined by the standard deviation in the fluxes selected for measuring the continuum), once more iterating through the list of wavelength elements. The signal is determined by a sum of the fluxes across five points centered on the wavelength of the current iteration, and the noise is the same but added in quadrature. Then the algorithm identified the maximum S/N; if the ratio exceeds five, we fit a Gaussian at the central wavelength element, integrate it to get the flux, and subtract the Gaussian from the residual flux. Then we check the next-highest S/N, and if it still exceeds five, the routine repeats until the peak S/N is below the detection threshold.

We run this routine on the individual PA spectra in each field, and record all instances of S/N > 5. If the code finds a peak in at least two PAs with centroids at the same or adjacent wavelength elements (24.5 Å in either direction), it declares a detection. In this paper, we focus on only one of the fields, GS1/HUDF, and specifically on candidates with optical data available for line confirmation. A broader catalog of 1D-selected ELGs from FIGS will be explored in a future paper (J. Pharo et al. 2019, in preparation). A search for ELGs in the FIGS 2D spectra can be found in Pirzkal et al. (2018).

In the GS1/HUDF field, where we have matching optical MUSE spectra, this method produces 137 candidate emission-line objects. Of these, 131 had matches in the MUSE source catalog within 1 arcsec of separation. Using our FIGS redshift catalogs (Pharo et al. 2018), we matched the candidate list with their redshifts and sorted the candidates according to the likely spectral emission line at that redshift. We use the wavelength of the peak S/N pixel to get an approximate rest-frame line centroid. We also compared our candidates with emission lines identified in the GRAPES Survey with the *HST* ACS G800L grism (Xu et al. 2007).

When identifying the FIGS-spectra emission lines, we considered common strong lines: Lyα, Hα, Hβ, [Mg II]

Table 3
The GS1/HUDF Emission-line Candidates by Identification

ID	Initial	MUSE	GRAPES	Photo-z	Wrong	Total
[O III]	32	11	7	13	1[a]	31
Hα	22	15	3	5	0	23
[O II]	17	4	1	12	0	17

Note.
[a] Later confirmed to be Hα.

$\lambda 2798$ Å, [O III]$\lambda 5007$ Å, and [O II]$\lambda 3727$ Å, though fainter lines could often be seen in the higher-resolution MUSE spectra. We identified the FIGS lines by determining the ratio of observed wavelengths between two detected emission lines, a fixed property for a given pair of emission lines that is not dependent on the redshift. If no other emission line was detected, we identified the line with the object's photometric redshift. This produced 32 [O III]$\lambda 5007$ Å candidates ($z \simeq 0.82$–1.35), 22 Hα candidates ($z \simeq 0.30$–0.80), and 17 [O II] $\lambda 3727$ Å candidates ($z \simeq 1.30$–2.0). The remaining detections were ruled out as contamination or some other false detection (e.g., due to a sharp change in continuum slope) upon visual inspection, were ruled out by other line detections in MUSE, or had bad redshift calculations. These tended to be among the faintest objects, which are more susceptible to contamination from nearby objects. In order to cast a wide net for ELGs, we did not impose a continuum magnitude limit on the search, relying on visual inspection and other spectra to confirm our detections.

Of the 32 [O III] candidates from FIGS, 11 were confirmed by inspecting matching MUSE spectra, which means we either measured the same line in the region of overlapping wavelength coverage (8500–9300 Å), or measured a second line which produced a wavelength ratio consistent with an [O III]$\lambda\lambda 4959,5007 - $[O II]$\lambda\lambda 3727,3729$ pair. However, due to the presence of atmospheric emission lines in MUSE and the fact that we cannot know [O II] strength just from [O III] detection, lack of a clear [O II] detection does not rule out the line being [O III]. Matches with the GRAPES line list confirmed an additional 7 candidates, leaving 13 unconfirmed (though the line ID is still implied by the redshift) and 1 confirmed to be Hα.

We used a similar process for the FIGS Hα candidates, of which 15 were confirmed by MUSE, 3 by GRAPES, and 5 were unconfirmed except by photometric redshift (photo-z). For [O II], MUSE can only reliably provide confirmation if the [O II] line is in the overlap region, or if another feature (e.g., 4000 Å break), happens to be visible. Only 4 could be confirmed this way, and 1 more from GRAPES, leaving 12 candidates unconfirmed except by photo-z. See Table 3 for a summary of these results.

3.2. Flux Measurement

We calculated the emission-line fluxes for all of the emission-line candidates, regardless of their confirmation status. Beginning with the brightest FIGS line in the spectrum (Hα, [O III], or [O II], depending on the candidate line ID), we performed a Gaussian fit using the Kapteyn Package (Terlouw & Vogelaar 2015) at the wavelength of the peak in each PA where there was a 5σ detection, allowing the Gaussian amplitude and sigma to be free parameters with an initial guess based on the peak flux. The
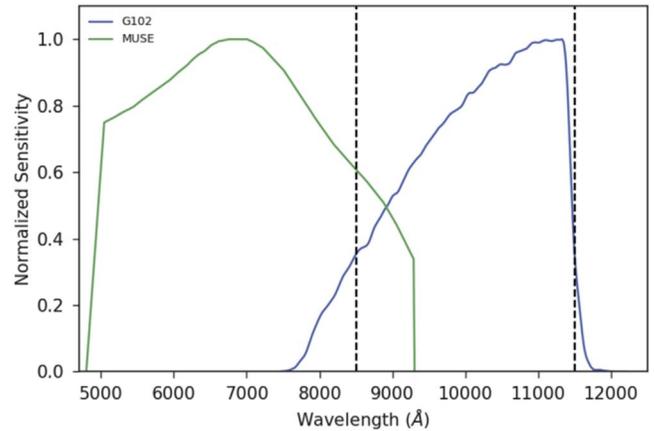
Figure 3. The sensitivity curve for the WFC3/G102 grism (Kuntschner et al. 2011; Lee et al. 2014), and for MUSE (Bacon et al. 2010). The dashed vertical lines show the cutoffs for grism data used in the emission-line search. The curves have been normalized to their maximum sensitivity.

centroid was allowed to vary between the adjacent pixels in order to determine the best-fit line center. We interpolated a Gaussian function from the fit, from which we derived the total line flux. Once all PAs for a single object and line had been fit, we averaged the individual fluxes and propagated the individual errors to get the final line measurement.

Once the primary line fit was completed, we recalculated the redshift based on the line center and used this new redshift measurement to predict the locations of other lines. The Gaussian function representing the previous fit was subtracted from the flux, and then we attempted to fit the next line. We repeated this process for any common, strong emission lines within the wavelength coverage. In the FIGS spectra, the only non-primary line detected with any significance was Hβ. If an object had matching MUSE spectra, we applied the same process there as well. Wes estimated the total flux errors based on the propagation of errors in the Gaussian fit parameters, which the Kapteyn fitting package determined in part based on flux errors in the constituent pixels.

For the 18 objects where the Hα emission line was detected in FIGS and for which a matching optical spectrum was available, we measured the extinction via the Balmer decrement. All the galaxies for which we later derive T_e-based metallicities are included in this set. In order to correct for stellar absorption of the Balmer lines, we follow the procedure used in Ly et al. (2014), which covers objects a similar redshift. For objects without measurements in both Hα and Hβ, we applied an extinction correction using the Calzetti et al. (2000) reddening model, following the procedure used in a study of ELGs with R_{23} at comparable redshift (Xia et al. 2011). The full catalog of flux measurements is listed in Table 6.

3.3. Line Comparisons in FIGS and MUSE

In addition to using emission lines in optical spectra to confirm line detections in FIGS, the measurement of additional line fluxes for an ELG makes it possible to measure gas-phase metallicities, but it is first necessary to check the consistency of the flux measurements between the two sources of spectra. We were able to check this by looking at emission lines that appeared in both the G102 and MUSE spectra. For emission lines observed between 8500 and 9200 Å (See Figure 3), where the throughput of both instruments is good, we were usually

Table 4
FIGS-MUSE Objects

FIGS ID	R.A.	Decl.	λ (Å)	z	FIGS Flux[a]	MUSE Flux[a]
1467	53.151047	−27.777309	8735	0.736	347.7 ± 27.1	365.3 ± 36.2
1689	53.162483	−27.780346	8615	0.719	890.6 ± 39.0	860.9 ± 56.7
1851	53.152782	−27.782698	8855	0.766	967.2 ± 58.4	1007.1 ± 97.3
2560	53.184158	−27.792637	8687	0.738	1821.9 ± 208.1	2041.8 ± 121.9
2654	53.182205	−27.793993	8687	0.735	315.0[b] ± 45.5	352.1 ± 50.6
8178	53.187664	−27.783779	8663	0.734	176.6 ± 28.7	94.3 ± 34.3

Notes.
[a] Measured in 10^{-19} erg cm^{-2} s^{-1}.
[b] Flux calculated after removing one PA for uncorrected contamination, which significantly altered the flux average.

able to measure the line flux in both FIGS and MUSE. This provided the opportunity to compare line measurements between the space-based *HST*/WFC3 and the ground-based VLT/MUSE instruments. In Table 4, we show the flux measurements of the six matching objects, where the [O III] $\lambda\lambda4959,5007$ were measured. The matching fluxes are within the measured 1σ flux errors for four of the seven objects, including the two which are part of the later analysis in this paper. Object 2654 was primarily detected by the Hα line, and one PA of the FIGS spectra contains non-removed contamination at the predicted location of the [O III] line, which skewed its average flux measurement high. Removing this one PA from the flux measurement brings the FIGS spectra flux into agreement with what we measure in MUSE, bringing the number of well-matched spectra to five of seven objects.

Because so few objects have an emission line appear in both spectra, it is difficult to judge whether any systematic offset is present from the few where the flux differs. Examining these cases does, however, suggest some possible causes for difference in FIGS and MUSE flux due to contamination or other artifacts, which we checked for visually in our further results. We examined the individual PAs for the 14 objects used in the mass–metallicity analysis to search for any unnoticed contamination that could affect the FIGS lines as with Object 2654, or for any other issues. We discovered no such contamination in any lines required for the metallicity measurement. Object 1299 possibly suffers from oversubtraction of the Hα line. However, this object is detected in five PAs, so the effect is small.

We also compared the redshifts derived individually from the FIGS line detection and the MUSE line. We calculated the redshift of each object in Table 4 based on the best-fit central wavelength of the line fit for each spectrum. The differences are plotted as a function of MUSE redshift in Figure 4. We find a root-mean-square (rms) redshift difference ($\Delta z/(1 + z)$) between the two sources of spectra of $\sigma_z = 0.002$. Xia et al. (2011) found an rms of $\sigma_z = 0.006$ when comparing redshifts derived from the *HST* ACS PEARS grism survey and from *Magellan*'s LDSS-3 spectrograph ($R = 100$, as compared to $R = 210$ for G102). This indicates FIGS has a wavelength calibration of at least comparable accuracy to PEARS, given the spectral resolution in each.

4. Metallicity Measurements

Strong nebular emission lines provide the means to measure the gas-phase oxygen abundance in a galaxy, which serves as a proxy for the metallicity. For objects with a [O III]$\lambda4363$ Å auroral line detected at S/N ⩾ 3, we used the direct metallicity measurement described in Ly et al. (2014), based on the
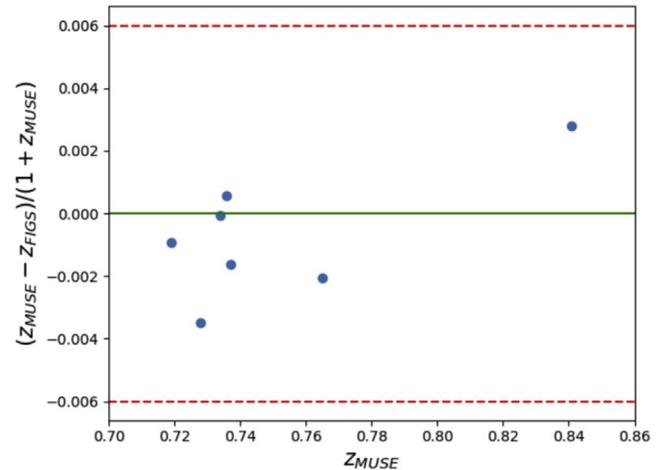
Figure 4. The differences in redshift calculated from the FIGS emission lines and the MUSE emission lines are shown here in blue, with the green line showing a difference of 0. The dashed red lines give the bounds of the rms wavelength error from (Xia et al. 2011) between the ACS grism and the LDSS3 spectrograph. We measure and rms redshift difference of $\sigma_z = 0.002$.

empirical relations in Izotov et al. (2006). This method first estimates the [O III]$\lambda4363$ electron temperature (T_e) based on the nebular-to-auroral flux ratio:

$$\log\left(\frac{[\text{O III}]\lambda\lambda4959, 5007}{[\text{O III}]\lambda4363}\right) = \frac{1.432}{t_3} + \log C_T \quad (1)$$

where $t_3 = T_e([\text{O III}])/10^4$ K, and

$$C_T = (8.44 - 1.09t_3 + 0.5t_3^2 - 0.08t_3^3) \cdot \frac{1 + 0.0004x}{1 + 0.044x} \quad (2)$$

where $x = 10^{-4}n_e t_3^{-0.5}$ and n_e is the electron density (cm^{-3}). Since we are unable to resolve the [S II]$\lambda\lambda6717,6732$ doublet in FIGS, and it is too red to appear in MUSE spectra, we cannot directly measure n_e, but C_T is only strongly dependent on n_e in the high-density regime ($n_e > 10^4$ cm^{-3}), where n_e is large enough for the x term to be important. We tested the temperature calculation with $n_e = 10, 100, 1000$ cm^{-3} using a range of measured line ratios from Ly et al. (2014), and the resulting temperatures were virtually identical for the different density measurements. Thus, we can safely adopt the assumption of Ly et al. (2014) that $n_e \approx 100$ cm^{-3} for our calculations.

Table 5
FIGS T_e Metallicities

FIGS ID	12 + Log(O/H)	Log($M_\star$)	Log(T_e/K)	SFR($M_\odot$ yr^{-1})	z
950	7.81 ± 0.15	8.92	4.46 ± 0.09	0.71 ± 0.18	0.678
1016	8.25 ± 0.23	9.12	4.01 ± 0.10	0.17 ± 0.15	0.622
1103	8.17 ± 0.23	9.80	4.03 ± 0.06	0.59 ± 0.23	0.334
1171	7.52 ± 0.07	8.58	4.35 ± 0.09	0.76 ± 0.13	0.606
1295	8.19 ± 0.05	9.74	4.00 ± 0.08	1.41 ± 0.17	0.420
1299	8.31 ± 0.68	9.87	4.02 ± 0.13	0.59 ± 0.76	0.622
1392	8.19 ± 0.09	9.87	4.01 ± 0.08	3.18 ± 1.27	0.668
1689	7.41 ± 0.12	8.20	4.34 ± 0.08	0.65 ± 0.18	0.719
2168	7.53 ± 0.09	7.99	4.38 ± 0.05	0.12 ± 0.02	0.468
2378	8.09 ± 0.24	10.06	4.08 ± 0.05	2.92 ± 1.20	0.436
2517	8.13 ± 0.17	9.64	4.09 ± 0.07	5.99 ± 2.13	0.459
2560[a]	8.24 ± 0.11	9.88	4.51 ± 0.07	5.61 ± 1.01	0.738
2783	8.31 ± 0.79	7.88	4.03 ± 0.13	0.06 ± 0.13	0.532
4198	8.36 ± 0.06	10.44	4.00 ± 0.14	3.10 ± 0.39	0.669

Note. For object coordinates and line fluxes, see Table 6. Stellar mass error is ⩽0.1 dex. See Section 5.1 for discussion.
[a] Object 2560 has possible indicators of being a Seyfert galaxy. See Section 4 for details.

With the temperature estimated, the ionic abundances of oxygen can be determined from the line ratios [O II]$\lambda\lambda$ 3726,3729/Hβ and [O III]$\lambda\lambda$4959,5007/Hβ:

$$12 + \log\left(\frac{O^+}{H^+}\right) = \log\left(\frac{[O\ II]}{H\beta}\right) + 5.961 + \frac{1.676}{t_2} \quad (3)$$

$$12 + \log\left(\frac{O^{++}}{H^+}\right) = \log\left(\frac{[O\ III]}{H\beta}\right) + 6.200 + \frac{1.251}{t_3} \quad (4)$$

where t_2 is the [O II] electron temperature, assuming a two-temperature model $t_2 = T_e$ ([O II])/10^4 K $= -0.577 + t_3(2.065 - 0.498t_3)$ from Izotov et al. (2006). In nebular regions, oxygen ions in ionization states other than O$^+$ and O^{++} make up a negligible fraction of the population, so the total oxygen abundance can be determined from

$$\frac{O}{H} = \left(\frac{O^+}{H^+}\right) + \left(\frac{O^{++}}{H^+}\right). \quad (5)$$

Fourteen objects have sufficient [O III]4363 signal (S/N ⩾ 3) in the MUSE optical spectra, as well as the other requisite [O III], [O II], and Hβ lines from FIGS and MUSE, to perform this direct metallicity measurement. We refer the reader to Tables 5 and 6 for the measured line fluxes and spectroscopic sources for individual ELGs. We summarize metallicity and electron-temperature measurements for these objects in Table 5. For one of these objects, FIGS ID 2560, we observed a strong peak at the location of the [O I]λ6300 emission line in one PA, a possible indicator of Seyfert or LINER properties. By consulting line ratio diagnostics in Kewley et al. (2006), our measurements match the characteristics of a Seyfert galaxy, which could explain the very high temperature measurement, and could skew the metallicity calculation if the [O I] line is real.

For ELGs without a significant [O III]4363 detection, we computed metallicities iteratively using the R_{23} diagnostic (Pagel et al. 1979), given by the ratio R_{23} = ([O II] + [O III])/Hβ. We tested the effectiveness of this method compared to the direct measurement by calculating metallicities using both methods for the 14 objects where this was possible. We found some significant disagreement in the results between the two, particularly at low

metallicity, as can be seen in Figure 5. This is not unusual: Kewley & Ellison (2008) show that different metallicity diagnostics can produce different measurements of oxygen abundance with a scatter of up to 0.7 dex. However, Christensen et al. (2012) demonstrated that using an R_{23} calibration with a correction for the ionization parameter based on the [O III]/[O II] ratio (Pilyugin & Thuan 2005) agreed well with direct metallicities of $z \sim 2$ galaxies. To address this, we applied a new empirical R_{23} calibration with an [O III]/[O II] ratio correction, based on a sample of 800 "green pea" galaxies at $0.011 < z < 0.411$ with reliable direct metallicity measurements (Jiang et al. 2019). This new calibration reduced the scatter between T_e-derived metallicities and R_{23}-derived metallicities, as can be seen in Figure 5, demonstrating that we could obtain reliable metallicity measurements using R_{23}. Thus, we were able to add eight additional objects to our metallicity sample via the calibrated R_{23} method.

Error measurements for the metallicities are obtained via the propagation of the initial flux errors through the electron-temperature calculation combined with error introduced by the extinction correction, and are reported at 1σ. Electron-temperature errors are determined by the errors of the line fluxes going into the [O III] line ratio: [O III]$\lambda\lambda$5007,4959 and [O III]λ4363. These are summarized in Table 5.

5. Results and Discussion

5.1. Mass–Metallicity Relation

We obtained stellar masses from the catalogs in Santini et al. (2015), hereafter S15. S15 presented a series of mass catalogs derived from CANDELS photometry (UV to through mid-IR in GOODS-S) and redshifts. The catalogs were computed using a variety of stellar mass codes and a range of preferred modeling parameters. We considered only the mass catalogs whose fits included contributions from nebular emission, which restricted our choice to four of the mass catalogs presented by S15. We use the mass values from one of these, their $6a_\tau NEB$ method, which is fit to BC03 templates (Bruzual & Charlot 2003) using a Chabrier Initial Mass Function (IMF) and includes the widest range of considered metallicities out of the four methods that consider nebular emission in their SED fits. S15 do not provide

Table 6
All Observed Emission-line Fluxes for 71 Galaxies in GS1/HUDF

FIGS ID	R.A.	Decl.	F105W	Hα	[O III]	Hβ	[O III]4363	[O II]3727	z[a]
724	53.17226	−27.76062	22.56	⋯	⋯	⋯	⋯	532.6 ± 53.3	1.550
950	53.16150	−27.76762	23.43	788.8 ± 78.9	254.4 ± 31.3[b]	101.3 ± 14.4[b]	17.23 ± 8.76[b]	259.6 ± 29.8[b]	0.678
970	53.16018	−27.76931	24.06	⋯	621.2 ± 36.0	138.8 ± 16.9	⋯	216.4 ± 51.9[b]	1.037
1013	53.16993	−27.77103	19.97	772.1 ± 77.2	368.0 ± 42.2[b]	674.3 ± 95.8[b]	⋯	1362 ± 109[b]	0.622
1016	53.17210	−27.77038	23.62	215.3 ± 21.5	49.7 ± 12.0[b]	33.6 ± 23.4[b]	17.1 ± 9.2[b]	88.5 ± 12.5[b]	0.622
1056	53.16245	−27.77091	24.30	⋯	260.7 ± 13.4	33.2 ± 3.3	⋯	43.1 ± 10.3[b]	1.038
1103	53.17400	−27.77206	20.72	2019 ± 285[c]	80.0 ± 37.2[b]	178.2 ± 20.3[b]	70.64 ± 17.81[b]	302.6 ± 30.1[b]	0.334
1132	53.18448	−27.77225	24.58	⋯	197.5 ± 42.5[c]	59.8 ± 36.4[c]	⋯	47.8 ± 24.4[b]	0.840
1171	53.15122	−27.77284	23.79	866.7 ± 86.7	756.8 ± 43.5[b]	158.4 ± 22.0[b]	21.1 ± 11.3[b]	331.8 ± 30.0[b]	0.606
1239	53.19146	−27.77389	23.48	666.9 ± 66.7	⋯	⋯	⋯	⋯	0.420
1295	53.16236	−27.77506	20.58	1869 ± 522[c]	198.4 ± 21.4[b]	476.2 ± 49.1[b]	108.9 ± 46.0[b]	655.0 ± 54.8[b]	0.420
1296	53.15936	−27.77503	23.09	⋯	354.3 ± 31.9	131.6 ± 13.2	⋯	⋯	1.219
1299	53.16080	−27.77537	21.24	657.2 ± 65.7	126.0 ± 76.3[b]	⋯	18.1 ± 8.9[b]	671.0 ± 48.4[b]	0.622
1316	53.16531	−27.77486	27.01	⋯	91.6 ± 8.0	77.2 ± 7.7	⋯	79.4 ± 31.6[b]	1.253
1359	53.18591	−27.77561	22.90	⋯	⋯	⋯	⋯	240.3 ± 76.7[c]	1.414
1392	53.18105	−27.77618	22.08	1335 ± 71.2	150.6 ± 52.4[b]	72.0 ± 9.2[b]	31.4 ± 10.3[b]	211.0 ± 29.1[b]	0.668
1467	53.15105	−27.77731	24.14	355.3 ± 21.1	336.6 ± 50.6[c]	69.0 ± 38.6[b]	⋯	194.1 ± 25.9[b]	0.736
1476	53.14744	−27.77760	23.63	⋯	⋯	⋯	⋯	189.9 ± 11.3	1.859
1477	53.15829	−27.77745	24.49	⋯	⋯	⋯	⋯	310.6 ± 19.6	1.556
1481	53.14661	−27.77749	25.03	⋯	281.3 ± 21.9	56.7 ± 6.0	⋯	106.7 ± 27.5[b]	1.088
1500	53.15234	−27.77795	24.32	⋯	⋯	⋯	⋯	214.0 ± 74.1[c]	1.413
1552	53.15720	−27.77852	23.83	⋯	⋯	78.0 ± 12.7	⋯	299.0 ± 44.3[c]	1.307
1689	53.16248	−27.78035	25.11	559.4 ± 48.9	707.6 ± 83.1[c]	103.0 ± 32.4[b]	31.23 ± 14.41[b]	39.7 ± 10.5[b]	0.719
1711	53.19700	−27.78060	23.86	⋯	575.4 ± 45.5	⋯	⋯	⋯	0.733
1728	53.17633	−27.78086	24.99	684.8 ± 44.5	660.2 ± 43.9[c]	89.6 ± 10.9[b]	⋯	109.1 ± 15.7[b]	0.535
1803	53.17007	−27.78207	26.68	⋯	⋯	56.2 ± 7.5	⋯	134.7 ± 27.2[c]	1.351
1829	53.15076	−27.78256	24.22	⋯	⋯	90.4 ± 9.0	⋯	136.6 ± 38.1[c]	1.352
1851	53.15278	−27.78270	24.44	⋯	926.3 ± 100.5[c]	124.7 ± 20.9[b]	⋯	260.4 ± 23.5[b]	0.764
1900	53.18457	−27.78332	24.42	⋯	979.0 ± 69.8	100.5 ± 10.1	⋯	195.6 ± 27.7[b]	1.136
1901	53.18433	−27.78337	25.78	⋯	254.9 ± 18.6	13.8 ± 1.4	⋯	⋯	1.257
1946	53.19259	−27.78379	24.71	⋯	429.1 ± 32.5	42.1 ± 7.6	⋯	⋯	0.869
2023	53.15186	−27.78475	25.58	⋯	285.4 ± 25.9	36.2 ± 3.6	⋯	36.2 ± 18.2[b]	1.219
2039	53.16657	−27.78486	27.62	⋯	⋯	29.6 ± 5.2	⋯	102.7 ± 27.0[c]	1.320
2049	53.16935	−27.78499	27.13	⋯	⋯	27.3 ± 2.7	⋯	100.3 ± 26.5[c]	1.344
2138	53.16048	−27.78630	24.31	⋯	356.4 ± 28.9	66.5 ± 6.5	⋯	⋯	0.984
2168	53.16347	−27.78664	25.14	284.9 ± 19.6	227.7 ± 18.9[b]	41.5 ± 3.9[b]	13.73 ± 4.08[b]	64.2 ± 5.7[b]	0.468
2187	53.17775	−27.78697	24.37	⋯	336.3 ± 21.7	99.7 ± 11.0	⋯	143.2 ± 22.8[b]	0.955
2221	53.16410	−27.78730	23.78	⋯	2168 ± 156	318.2 ± 31.8	⋯	284.1 ± 50.1[b]	1.097
2291	53.14930	−27.78853	23.05	⋯	⋯	⋯	⋯	453.0 ± 45.3	1.917
2338	53.15736	−27.78922	25.11	⋯	346.1 ± 11.6	3.6 ± 0.4	⋯	18.7 ± 17.8[b]	1.015
2363	53.16802	−27.78967	22.80	717.5 ± 61.3	373.2 ± 30.6[b]	30.9 ± 20.9[b]	⋯	626.8 ± 48.2[b]	0.619
2378	53.18795	−27.79000	20.25	8916 ± 568	198.1 ± 36.0[b]	489.6 ± 44.1[b]	81.29 ± 17.89[b]	640.2 ± 56.5[b]	0.436
2385	53.18481	−27.78993	23.13	⋯	392.3 ± 31.2	160.0 ± 25.4	⋯	431.7 ± 62.6[b]	0.954
2417	53.16042	−27.79037	23.48	⋯	⋯	⋯	⋯	354.6 ± 35.5	1.614
2495	53.18413	−27.79153	23.07	⋯	251.2 ± 14.5	50.6 ± 5.1	⋯	⋯	1.224
2517	53.16161	−27.79230	20.62	16439 ± 1374	3238 ± 198[b]	1127 ± 88[b]	45.2 ± 17.46[b]	3279.0 ± 202.2[b]	0.459
2560	53.18416	−27.79264	21.41	5114 ± 225	1810 ± 226[c]	716 ± 54[b]	146.72 ± 54.11[b]	1495 ± 101[b]	0.738
2570	53.16412	−27.79265	27.02	⋯	⋯	30.9 ± 3.1	⋯	89.6 ± 25.7[c]	1.311
2654	53.18221	−27.79399	24.81	278.9 ± 27.9	996.6 ± 53.5[c]	33.3 ± 28.8[b]	⋯	80.6 ± 16.4[b]	0.734
2669	53.15663	−27.79430	24.33	⋯	329.3 ± 32.3	118.8 ± 11.9	⋯	⋯	1.094
2696	53.15586	−27.79490	22.94	⋯	785.2 ± 42.9	215.5 ± 17.6	⋯	⋯	1.104
2720	53.15675	−27.79558	21.80	⋯	731.9 ± 35.8	330.3 ± 33.0	⋯	398.5 ± 76.1[b]	1.099
2732	53.16133	−27.79580	23.59	⋯	⋯	⋯	⋯	485.7 ± 134.9[c]	1.498
2783	53.18808	−27.79574	24.03	125.6 ± 12.6	43.3 ± 16.6[b]	⋯	9.3 ± 6.8[b]	15.4 ± 10.0[b]	0.532
2872	53.16687	−27.79771	23.68	⋯	615.7 ± 42.9	66.3 ± 6.6	⋯	409.5 ± 28.0[b]	0.984
2942	53.16112	−27.79880	25.53	⋯	387.9 ± 34.1	19.5 ± 1.9	⋯	231.4 ± 90.6[b]	1.238
4198	53.17838	−27.76824	20.22	1964.6 ± 115.3	317.8 ± 276.7[b]	474.1 ± 184.8[b]	110.9 ± 16.5[b]	1073.9 ± 97.2[b]	0.669
4258	53.15229	−27.77009	23.66	⋯	⋯	⋯	⋯	420.6 ± 32.9	1.859
4284	53.18454	−27.76822	25.22	⋯	⋯	⋯	⋯	74.4 ± 7.4	1.839
6865	53.19033	−27.77430	26.84	⋯	195.7 ± 12.1	119.7 ± 12.0	⋯	⋯	0.902
8178	53.18766	−27.78378	26.95	84.3 ± 8.4	126.8 ± 37.2[c]	91.9 ± 44.5[b]	⋯	16.8 ± 15.3[b]	0.739

Notes. Fluxes are given in units of 10^{-19} erg s^{-1} cm^{-2}.
[a] Derived from central wavelength of the most significantly detected line, averaged from PAs and MUSE.
[b] Line measured only in MUSE.
[c] Line measured in FIGS and MUSE.

7
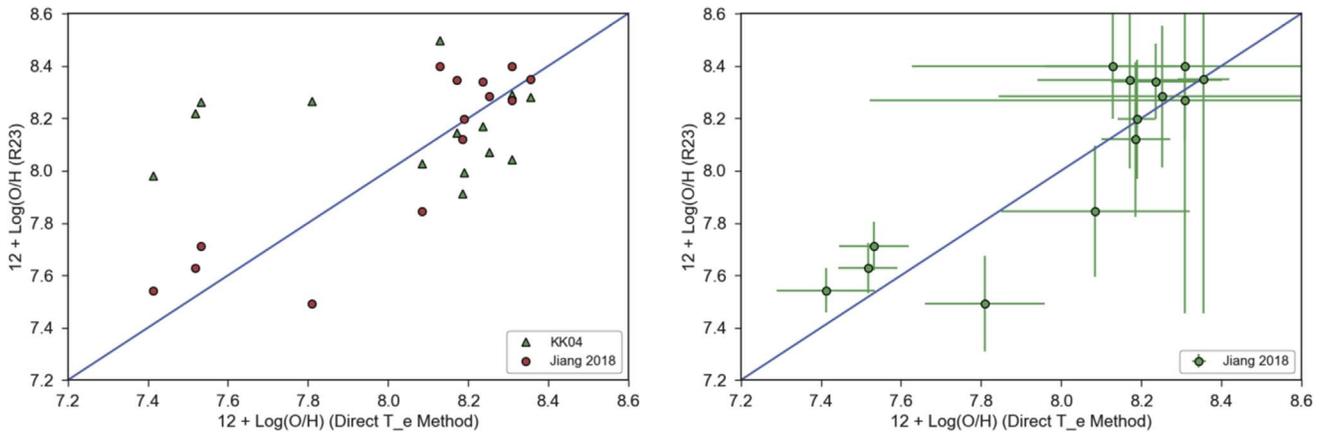
Figure 5. Left: R_{23} metallicities compared with T_e-derived metallicities, using both the KK04 (Kobulnicky & Kewley 2004) parameterization (green triangles) and new calibration based on Green Peas (Jiang et al. 2019) (red circles). Both R_{23} methods place all of the FIGS objects on the lower branch, but the KK04 parameterization tends to overestimate the lowest-metallicity objects, while the Green Pea calibration reduces the scatter considerably. Right: the comparison of the calibrated R_{23} metallicities with T_e metallicities including errors.
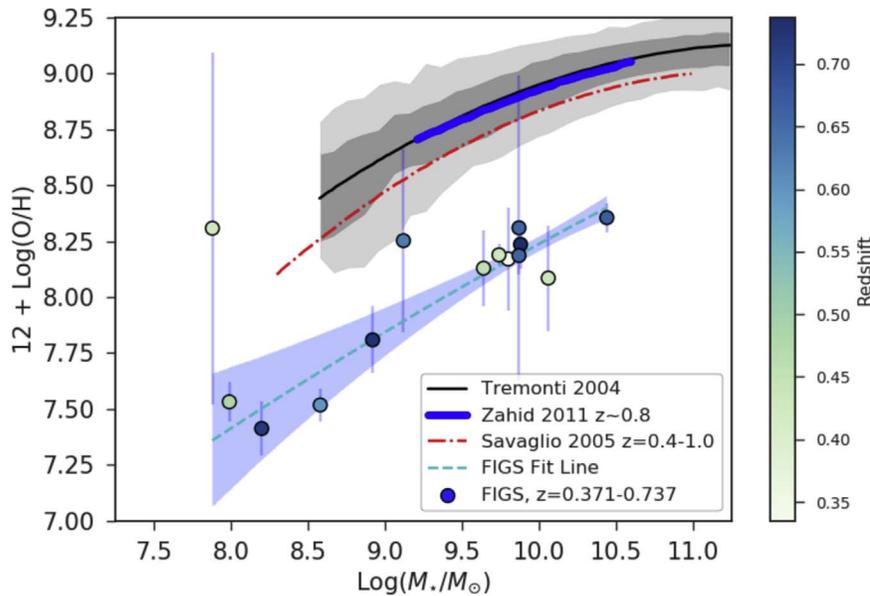
Figure 6. The MZ relation between the stellar masses as measured by Santini et al. (2015) and the gas-phase oxygen abundances for FIGS objects as measured by [O III]λ 4363 in matching MUSE spectra. The FIGS objects are given by circles with errorbars in metallicity, and are shaded by redshift, with a median of $z = 0.614$. The black line and contours represent local SDSS galaxies as measured in Tremonti et al. (2004). The thick blue, solid line represents the $z \sim 0.8$ upper-branch R_{23} metallicities from Zahid et al. (2011), and the red, dotted–dashed line represents the $0.4 < z < 1.0$ upper-branch R_{23} metallicities from Savaglio et al. (2005). The blue dashed line is the nonlinear least-squares fit to the FIGS objects, using SciPy's curve-fit function (see Equation (7) for parameters; Jones et al. 2001). The blue shaded region is the 1σ uncertainty of the fit as measured from a Monte Carlo of the 1σ uncertainties in the fit parameters.

individual estimates of the mass error, but did investigate the distributions of mass estimates as compared to the median masses from the list of mass catalogs. They quantified the typical deviation from the median mass with the distribution's semi interquartile range, which they found to be usually less than 0.1 dex, giving a reasonable upper bound on the mass uncertainty.

We matched the S15 catalog with our 22 objects with T_e or R_{23} metallicity measurements within an angular separation of 1 arcsec and confirmed that the CANDELS redshifts provided by the S15 catalog matched the line-derived redshifts for the objects. Then we produced a relation between the stellar mass and the gas-phase oxygen abundance for the 14 objects with T_e-derived metallicities, as can be seen in Figure 6. This subsample has a median redshift of $z = 0.614$. We parameterize the FIGS mass–metallicity relation

with a quadratic function of the form

$$12 + \log\left(\frac{O}{H}\right) = A + Bx + Cx^2 \qquad (6)$$

where $x = \log(M_\star/M_\odot) - 10$. We use a Python function, *curve-fit* from the SciPy package (Jones et al. 2001), to perform a least-squares fit of the FIGS data to this parameterization. The MZ relation is best fitted by

$$12 + \log\left(\frac{O}{H}\right) = 8.240 + 0.367x - 0.018x^2. \qquad (7)$$

The 1σ errors in the parameters are determined from the diagonal of the covariance matrix, which gives $\sigma_A = 0.033$,
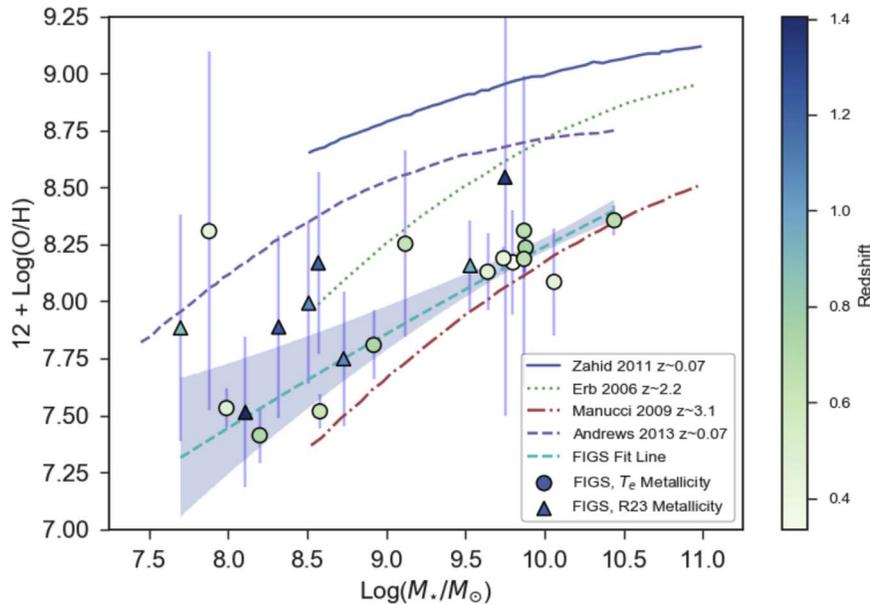
Figure 7. The relation between the stellar masses as measured by Santini et al. (2015) and the gas-phase oxygen abundances as measured by [O III]λ4363. The FIGS objects are given by circles (direct metallicity measurement) and triangles (calibrated R_{23}) with blue errorbars in metallicity. The FIGS objects are shaded by redshift according to the colorbar to the right. The blue solid line represents the $z \sim 0.07$ relation for local galaxies described in Zahid et al. (2011); the purple dashed line represents the $z \sim 0.07$ direct-metallicity relation from Andrews & Martini (2013); the green dotted line Erb et al. (2006) and the red dashed–dotted line Mannucci et al. (2009) represent R_{23}-derived relations at $z = 2.2$ and $z = 3.1$. The blue dashed line is the nonlinear least-squares fit to the FIGS objects, using SciPy's curve-fit function (see Equation (7) for parameters; Jones et al. 2001). The blue shaded region is the 1σ uncertainty of the fit as measured from a Monte Carlo of the 1σ uncertainties in the fit parameters.

$\sigma_B = 0.089$, and $\sigma_C = 0.051$. We estimate the uncertainty in the fit by performing a Monte Carlo simulation at each stellar mass in the range of the fit (1000 points between $\log(M_\star/M_\odot) = 7.70$ and 10.44), assuming a Gaussian distribution around these errors. For each mass point, the fit parameters are sampled 10,000 times, and the standard deviation of the result is used to estimate the 1σ uncertainty in the fit. This is represented by the shaded region in Figure 6.

5.2. Comparison with Other MZ Relations

Figure 6 shows the FIGS-MUSE mass–metallicity relation for the T_e-measured objects plotted alongside mass–metallicity relations from other surveys at similar redshift. Our measurements are offset to lower metallicity by ~ 0.6–0.7 dex compared to these surveys. The curve from Zahid et al. (2011) is fit from stacks of DEEP2 objects at $z \sim 0.8$, for which metallicities were derived using the R_{23} method. Zahid et al. (2011) notes that since [N II]/Hα measurements were not available, they were unable to break the R_{23} degeneracy and instead assume the metallicities lie on the R_{23} upper branch, though they observe that this assumption breaks down at $M_\star < 10^9 M_\odot$. Savaglio et al. (2005) derived an MZ relation for 56 $0.4 < z < 1.0$ galaxies from the Gemini Deep Deep Survey and the Canada–France Redshift Survey, also using the R_{23} upper branch for metallicity. As described in Section 4, when we apply the R_{23} calculation to the 14 FIGS objects, both methods place all of them on the low-Z branch, which is itself enough to alter the metallicity measurement by up to ~ 1 dex, enough to explain the offset in metallicity between the two surveys.

A more "direct" MZ comparison can be made from the Andrews & Martini (2013, hereafter AM13) result, shown in Figure 7 as a purple dashed curve. The AM13 MZ relation is

derived from stacks of direct-method metallicity calculations of local SDSS galaxies at $z = 0.07$. The direct-metallicity FIGS measurements are denoted by circles and the calibrated-R_{23} measurements denoted by triangles. Despite also using the [O III]4363 T_e method, AM13 find higher metallicities than we find, with a median metallicity offset of $+0.65$ dex.

5.3. Discussion of the Offset

In this section, we examine possible causes for the low-metallicity offset of our sample.

5.3.1. Redshift

In Figures 6 and 7, the FIGS points are colored according to the line-centroid-derived redshift, but no significant trend in redshift emerges from among these 14 ELGs. This agrees with the results of Savaglio et al. (2005), who also found no significant redshift evolution in metallicity in their sample at a similar redshift range.

The median redshift in the AM13 sample is $z = 0.07$, with a maximum of $z = 0.25$. This is lower than the minimum redshift in the FIGS sample ($z = 0.371$), and the median redshift in the FIGS T_e-derived sample is $z = 0.614$. Previous surveys (Maiolino et al. 2008; Zahid et al. 2013) suggest the metallicity evolution from $z = 0$ to $z \sim 0.8$ is roughly 0.1–0.2 dex at a given stellar mass. This is not large enough to account for the offset between FIGS and AM13, though possibly the R_{23} measurements used by the previous surveys underestimate this evolution. This offset does allow for the FIGS objects to fall within the scatter of the metal-poor galaxies in the AM13 sample.

Jones et al. (2015) selected a sample of 32 DEEP2 galaxies with [O III]4363 emission at $z \sim 0.8$ from which they calculated gas-phase metallicities in the range $7.8 < 12 + \log(O/H) < 8.4$.

They do not include a mass–metallicity relation, but most of the FIGS objects have metallicities that are consistent with this metallicity range to within the 1σ error. Of the three objects with significantly lower measured metallicity, only one is at a redshift at the higher end of the sample redshift ($z = 0.719$), and thus at a comparable redshift to the Jones sample, and all three are at relatively low mass.

5.3.2. IMF

Inconsistency in the IMF used to derive stellar masses for different studies can produce offsets in stellar mass, which in turn would affect the MZ Relation. Masses for the Zahid et al. (2011) relation were also calculated using a Chabrier IMF, while the Savaglio et al. (2005) relation uses masses with an IMF derived by Baldry & Glazebrook (2003) that produces masses 1.2 times smaller than Kroupa. A calculation of IMF offsets (Zahid et al. 2012) suggests an offset of $+0.03$ dex between Kroupa-derived and Chabrier-derived stellar masses, and an offset of -0.07 dex between Chabrier masses and those used in Savaglio et al. (2005). The Andrews & Martini (2013) relation used masses derived from a Kroupa IMF, which should result in a $+0.03$ mass offset compared to the Chabrier masses used in the FIGS relation. These offsets are all comparable to the <0.1 dex scatter in stellar masses in the Sanders et al. (2015) catalogs and are much too small to explain the metallicity offset.

5.3.3. Contributions of Multiple H II Regions

Another explanation for the low metallicities we measure is the possibility that the lines we detect are dominated by emission from particularly extreme regions within the galaxy. An H II region with an especially low metallicity and large electron temperature could produce stronger [O III]λ4363 emission for that region. In a small galaxy, the flux from such a region could dominate compared to flux from milder regions, resulting in that region's low metallicity measurement reducing the overall metallicity measurement for the galaxy (Sanders et al. 2017). This could perhaps explain the extremely low metallicities of the lowest-mass objects, but does not account for the lower MZ relation overall.

5.3.4. Selection Effects, Line Emission, and Star Formation

In Figure 7, we have also included the FIGS objects with calibrated R_{23} metallicities (see Section 4), denoted by triangles. There continues to be no significant redshift evolution, as these new, higher-z objects tend to have higher metallicity. This is likely a selection effect: the highest redshift objects are also exclusively R_{23}-calibrated. This means that there cannot have been a detected [O III]λ4363 line, which itself implies possibly lower [O III]λ4363 emission, which in turn implies a higher metallicity for the objects in the calibration sample. Furthermore, these objects are typically fainter, resulting in larger flux errors which contribute to broader error bounds on the metallicities. The high-z objects are still consistent with the possible range of the MZ fit, and there are a few objects in the R_{23} sample where we measure low metallicities comparable to what we measure with [O III]λ4363. This means that [O III]λ4363 selection alone cannot fully account for the metallicity offsets, and so lends support to the findings in Xia et al. (2012), which suggest that emission-line strength itself is an indicator of low metallicity.
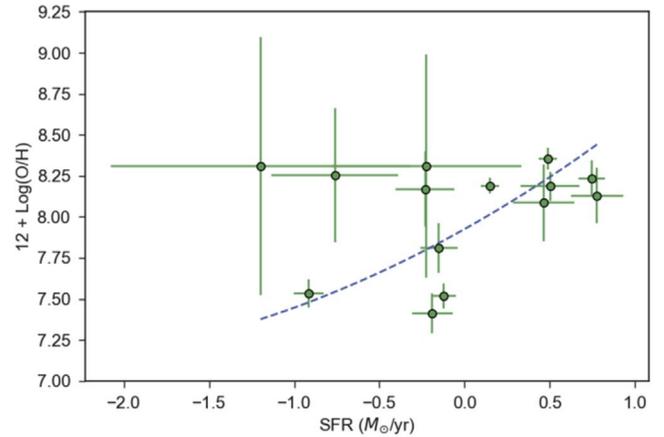
Figure 8. The gas-phase metallicity of the FIGS objects as a function of the SFR. The dashed line shows the nonlinear least-squares fit, which shows a trend of increasing metallicity with increasing SFR.

More recently, Amorín et al. (2017) find a sample of $2.4 < z < 3.5$ galaxies with $7.4 < 12 + \log(O/H) < 7.7$ using strong UV emission lines. They find low metallicities consistently across a broad range of stellar masses, up to $\log(M_\star/M_\odot) = 9.8$. Their sample also shows indicators of recent star formation, suggesting a link between star formation and metallicity somewhat independent of the stellar mass. We explore the effects of star formation for the FIGS objects in the following section.

5.4. SFR and the FMR

We calculate the SFR for the 14 objects with direct-measurement metallicities based on the line flux conversion given in Kennicutt (1998)

$$\text{SFR}(M_\odot \text{ yr}^{-1}) = 7.9 \times 10^{-42} L(\text{H}\alpha)(\text{erg s}^{-1}) \qquad (8)$$

using the extinction-corrected Hα line flux. The SFR error is estimated based on the line flux error. The metallicity as a function of SFR is shown in Figure 8, along with a nonlinear fit. This shows a trend of metallicity increasing with the SFR. In Figure 9, we plot the gas-phase metallicity versus the Specific Star Formation Rate (SSFR), which is the SFR per stellar mass, as we as the fit of the FIGS objects. This shows a slight trend of declining metallicity with increased SSFR, with the lowest metallicity (and smallest mass) galaxies having SSFR $> 10^{-9}$ yr^{-1}. Ellison et al. (2008) has shown a relation where metallicity is lower for galaxies with higher SSFR at a given stellar mass, with a metallicity offset of up to 0.15 dex at the lowest stellar masses ($M_\star \sim 10^{8.5}\ M_\odot$ in their study). This suggests that the large SSFR we observe in several of the FIGS objects could be a driver for the low-metallicity offset compared to other mass–metallicity relations. If this is the case, it likely has implications for how star formation interacts with the nebular gas. There are two plausible scenarios. First, inflows of circumgalactic gas could both bring lower-metallicity gas into the galaxy and trigger new star formation, producing strong line emission in the metal-deficient medium around the new stars. Alternatively, recent star formation in a galaxy produces strong stellar winds and supernovae, which could cause outflows that push the most metal-enriched gas out
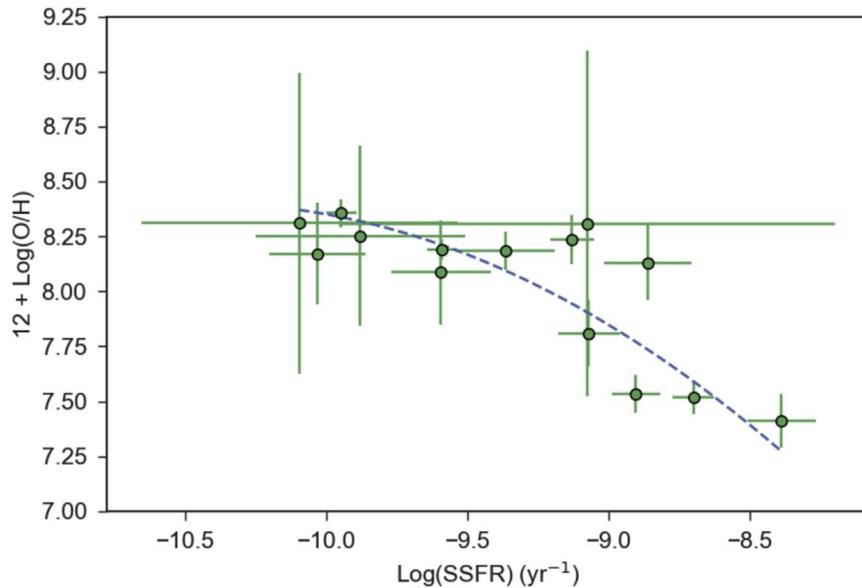
Figure 9. The gas-phase metallicity as a function of the specific star formation rate (SSFR). The dashed line shows the nonlinear least-squares fit of the FIGS points, which shows a possible trend of lower metallicity at increased SSFR.
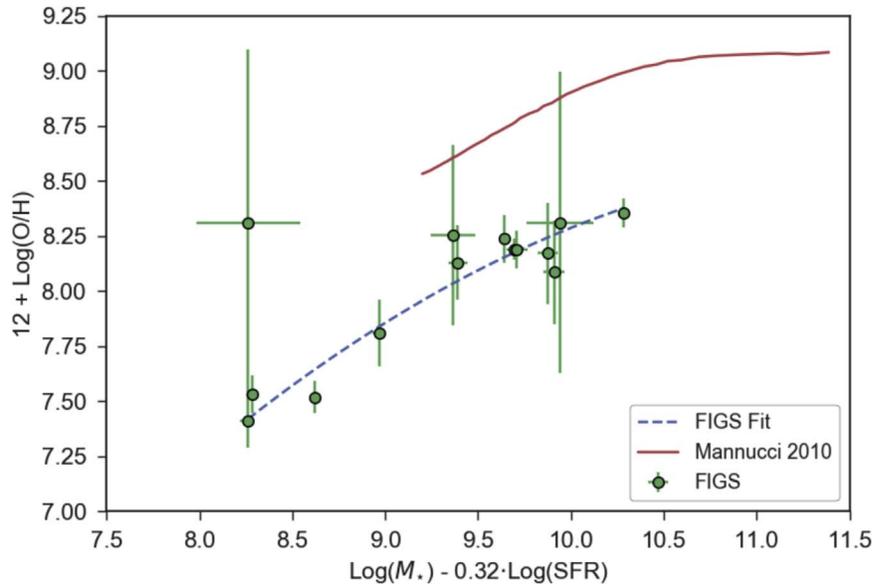
Figure 10. The gas-phase metallicity as a function of the Fundamental Metallicity Relation (FMR), as given by Mannucci et al. (2010). The Mannucci relation is shown by the red solid line, with the FIGS points in green. The blue dashed line gives the nonlinear least-squares fit of the FIGS points.

of the galaxy. In either case, increased star formation would show a clear link with measuring reduced metallicity in a galaxy's nebular gas.

Further investigation of the relationship between these parameters is needed. Mannucci et al. (2010) describes the dependence of the gas-phase metallicity on stellar mass and the SFR as the FMR, for which they obtain the projection

$$12 + \log\left(\frac{O}{H}\right) = \log\left(\frac{M_\star}{M_\odot}\right) - 0.32 \cdot \log(SFR). \quad (9)$$

This projection, derived from a sample of SDSS $z = 0.07$–0.30 ELGs which had an Hα S/N > 25, minimizes the scatter in metallicity around the relation. Mannucci et al. also find good agreement with the FMR and this projection for

galaxies up to $z = 2.5$. We calculated this projected FMR for the FIGS galaxies with SFR, plotted in Figure 10, to see how well our results match this relation. The FIGS 14 objects follow the trend of the lower FMR but sit lower on the plot due to their lower metallicities. This is perhaps partially accounted for by the differences between direct and R_{23} metallicity measurements as described in Section 5.1, though the metallicities used in Mannucci et al. were derived from either R_{23} or from the [N II]$\lambda6584$/Hα ratio. Mannucci et al. estimates a 0.09 dex offset in metallicity between these two methods, making the magnitude of the offset from R_{23} and direct measurements difficult to determine. Figure 10 also shows a difference in the range of values for the $M_\star$–SFR axis, with the FIGS sample probing much lower stellar masses than (and thus also lower SSFR than) the Mannucci sample. While the FIGS objects span
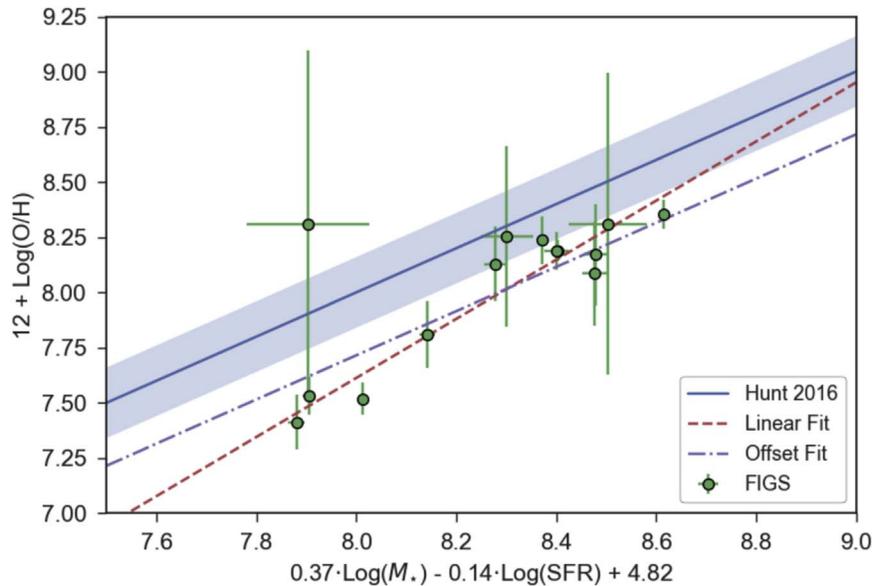
Figure 11. This plot shows the 14 direct-method FIGS objects plotted using the Fundamental Plane of Metallicity formulation from Hunt et al. (2016). The blue line shows the one-to-one correlation of metallicity to combined mass and SFR around which the formulation was based. The FIGS points are in green, and the dashed red line represents the simple linear fit to the FIGS data. The dotted–dashed purple line shows the linear fit to the FIGS data if the slope is fixed to match the Hunt correlation. This results in a metallicity offset of 0.28.

a range of SFR similar to that seen in Mannucci et al., the $M_\star$ values are lower, and we do not know how well Mannucci's projection reduces scatter at lower stellar mass.

We also tried comparing our results to the Fundamental Plane of Metallicity (FPZ) derived by Hunt et al. (2016) using the Metallicity Evolution and Galaxy Assembly (MEGA) data set. Hunt et al. attempted to derive a fundamental relation between metallicity, mass, and SFR from a large set of galaxies with a wide range of properties and redshifts, including a variety of strong-line methods for measuring the metallicity (the direct method among them, but not predominantly so). With this data set, Hunt et al. performed a Principal Component Analysis to derive a plane relating the three variables:

$$12 + \log\left(\frac{O}{H}\right) = -0.14 \log(\text{SFR}) + 0.37 \log\left(\frac{M_\star}{M_\odot}\right) + 4.82. \tag{10}$$

In Figure 11, we plotted the FIGS objects on this plane. The blue line gives the one-to-one correspondence given by Equation (10), with the shaded region providing the $\sigma = 0.16$ scatter from Hunt et al.'s narrowest residual distribution. The FIGS points lie systematically below this, though within the total scatter of MEGA objects around it. The dashed and dotted–dashed lines show linear fits to the FIGS points: the red (dashed) line allows both parameters of the linear fit to move freely, while the purple (dotted–dashed) line assumes the same slope as the one-to-one correspondence and only lets the y-intercept vary. This produces an overall metallicity offset of ~ 0.3 dex. The linear fit demonstrates that the higher-metallicity FIGS ELGs are actually fairly consistent with 1σ range of the FPZ measure, and that the lowest-metallicity objects are the ones driving the offset.

This all suggests that our sample of galaxies with direct metallicity measurements includes some uniquely low-mass, low-metallicity objects.

6. Conclusions

By using near-infrared spectroscopy from FIGS, we were able to identify 71 ELGs in the GS1/HUDF field, primarily through the identification of Hα, [O III]λ5007, and [O II]λ3727 emitters in the redshift range of $0.3 < z < 2.0$. We were able to confirm 41 out of the 71 ($\sim$58%) by identifying complementary lines in matching optical data, either with ACS grism spectroscopy from the previous GRAPES survey, or from new MUSE-VLT optical spectroscopy. We measure line fluxes down to a sensitivity of 10^{-17} erg cm^{-2} s^{-1} in FIGS and $\sim 3 \times 10^{-19}$ erg cm^{-2} s^{-1} from MUSE-VLT.

Out of these objects, we found 14 for which we were able to measure the auroral [O III]λ4363 emission line in MUSE optical spectra with an S/N of at least three, with a redshift range of $0.3 < z < 0.8$ and with stellar masses down to $10^{7.9} M_\odot$. We used these measurements to calculate the gas-phase metallicity via the electron temperature, and from this we produced a mass–metallicity relation. When compared to MZ relations at similar redshifts, we find a significant offset to lower metallicity. We examined several possible causes for the offset and find that redshift evolution does not account for the difference in metallicity. The offset can be only partially explained by differences with this metallicity derivation method compared to the more common R_{23} method, as previously seen in Andrews & Martini (2013) and Sanders et al. (2016). Selecting [O III]λ4363 line emitters does select for lower metallicity in general, but with a new R_{23} calibration, we found other galaxies at similarly low metallicity, so selection effects alone cannot account for the difference. To further explore the metallicity offset, we determined the SFR, SSFR, and FMR for the sample. We find a trend between metallicity and SSFR, showing that the low-mass, low-metallicity FIGS objects have a large SSFR and are low-metallicity outliers in attempts to find a fundamental relationship between these parameters. This suggests that recent star formation is connected to inflows or outflows of nebular gas, leading to the measurement of low-metallicity gas in the galaxy. The existence

of such outliers shows the need for further spectroscopic analysis of low-mass galaxies, which may be host to significant activity well after the universal peak of star formation at $z \simeq 2$.

ORCID iDs

John Pharo ⓘ https://orcid.org/0000-0003-2249-2539
Sangeeta Malhotra ⓘ https://orcid.org/0000-0002-9226-5350
James Rhoads ⓘ https://orcid.org/0000-0002-1501-454X
Lise Christensen ⓘ https://orcid.org/0000-0001-8415-7547
Steven L. Finkelstein ⓘ https://orcid.org/0000-0001-8519-1130
Norman Grogin ⓘ https://orcid.org/0000-0001-9440-8872
Tianxing Jiang ⓘ https://orcid.org/0000-0002-2222-6129
Keunho Kim ⓘ https://orcid.org/0000-0001-6505-0293
Anton Koekemoer ⓘ https://orcid.org/0000-0002-6610-2048
Andrea Cimatti ⓘ https://orcid.org/0000-0002-4409-5633
Nimish Hathi ⓘ https://orcid.org/0000-0001-6145-5090
Pascale Hibon ⓘ https://orcid.org/0000-0003-3726-5494
Rogier Windhorst ⓘ https://orcid.org/0000-0001-8156-6281

References

Aller, L. H. 1984, Astrophysics and Space Science Library, Vol. 112, Physics of Thermal Gaseous Nebulae (Dordrecht: Reidel), 360
Amorín, R., Fontana, A., Pérez-Montero, E., et al. 2017, NatAs, 1, 0052
Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
Bacon, R., Accardo, M., Adjali, L., et al. 2010, Proc. SPIE, 7735, 773508
Bacon, R., Conseil, S., Mary, D., et al. 2017, A&A, 608, A1
Baldry, I. K., & Glazebrook, K. 2003, ApJ, 593, 258
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Christensen, L., Richard, J., Hjorth, J., et al. 2012, MNRAS, 427, 1953
Dopita, M. A., & Evans, I. N. 1986, ApJ, 307, 431
Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, ApJL, 672, L107
Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813
Finkelstein, S. L., Cohen, S. H., Moustakas, J., et al. 2011, ApJ, 733, 117
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Horne, K. 1986, PASP, 98, 609
Hunt, L., Dayal, P., Magrini, L., & Ferrara, A. 2016, MNRAS, 463, 2002
Izotov, Y. I., Stasinska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, A&A, 448, 955
Jiang, T., Malhotra, S., Rhoads, J. E., & Yang, H. 2019, ApJ, 872, 145
Jones, E., Oliphant, E., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, v1.10, https://www.scipy.org/
Jones, T., Martin, C., & Cooper, M. C. 2015, ApJ, 813, 126
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35

Kewley, L. J., Dopita, M. A., Leitherer, C., et al. 2013, ApJ, 774, 100
Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183
Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961
Kobulnicky, H. A., & Kewley, L. J. 2004, ApJ, 617, 240
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, PASP, 121, 59
Kuntschner, H., Kümmel, M., Walsh, J. R., & Bushouse, H. 2011, Space Telescope WFC Instrument Science Rep., 5
Lara-López, M. A., Cepa, J., Bongiovanni, A., et al. 2010, A&A, 521, L53
Lee, J. C., Pirzkal, N., & Hilbert, B. 2014, Instrument Science Report, WFC3, 2014-01
Lilly, S. J., Carollo, C. M., & Stockton, A. N. 2003, ApJ, 597, 730
Ly, C., Malkan, M. A., Nagao, T., et al. 2014, ApJ, 780, 122
Maier, C., Lilly, S. J., Carollo, C. M., Stockton, A., & Brodwin, M. 2005, ApJ, 634, 849
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Malhotra, S., Rhoads, J. E., Pirzkal, N., et al. 2005, ApJ, 626, 666
Mannucci, F., Cresci, G., Maiolino, R., et al. 2009, MNRAS, 398, 1915
Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., & Smith, G. 1979, MNRAS, 189, 95
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pharo, J., Malhotra, S., Rhoads, J. E., et al. 2018, ApJ, 856, 116
Pilyugin, L. S., & Thuan, T. X. 2005, ApJ, 631, 231
Pirzkal, N., Collodel, L., Erben, T., et al. 2001, A&A, 375, 351
Pirzkal, N., Malhotra, S., Ryan, R. E., et al. 2017, ApJ, 846, 84
Pirzkal, N., Rothberg, B., Ly, C., et al. 2013, ApJ, 772, 48
Pirzkal, N., Rothberg, B., Ryan, R. E., et al. 2018, ApJ, 868, 61
Pirzkal, N., Xu, C., Malhotra, S., et al. 2004, ApJS, 154, 501
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16
Rhoads, J., Malhotra, S., Pirzkal, N., et al. 2009, ApJ, 697, 942
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, ApJ, 799, 138
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2016, ApJL, 825, L23
Sanders, R. L., Shapley, A. E., Zhang, K., & Yan, R. 2017, ApJ, 850, 136
Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, ApJ, 801, 97
Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, ApJ, 635, 260
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165
Straughn, A. N., Pirzkal, N., Meurer, G. R., et al. 2009, AJ, 138, 1022
Terlouw, J. P., & Vogelaar, M. G. R. 2015, Kapteyn Package, v2.3, http://www.astro.rug.nl/software/kapteyn/
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Xia, L., Malhotra, S., Rhoads, J., et al. 2011, AJ, 141, 64
Xia, L., Malhotra, S., Rhoads, J., et al. 2012, AJ, 144, 28
Xu, C., Pirzkal, N., Malhotra, S., et al. 2007, AJ, 134, 169
Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017, ApJ, 847, 38
Zahid, H. J., Dima, G. I., Kewley, L. J., Erb, D. K., & Dáve, R. 2012, ApJ, 757, 54
Zahid, H. J., Geller, M. J., Kewley, L. J., et al. 2013, ApJL, 771, L19
Zahid, H. J., Kewley, L. J., & Bresolin, F. 2011, ApJ, 730, 137