Morphology Dependence of Stellar Age in Quenched Galaxies at Redshift 1.2:Massive Compact Galaxies Are Older than More Extended Ones

The Astrophysical Journal, 838:94 (13pp), 2017 April 1 https://doi.org/10.3847/1538-4357/aa662f

Morphology Dependence of Stellar Age in Quenched Galaxies at Redshift ∼1.2: Massive Compact Galaxies Are Older than More Extended Ones

Christina C. Williams[1], Mauro Giavalisco[2], Rachel Bezanson[1,9], Nico Cappelluti[3,4], Paolo Cassata[5], Teng Liu[2,6], Bomee Lee[2], Elena Tundo[7], and Eros Vanzella[8]

[1] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA; ccwilliams@email.arizona.edu
[2] Department of Astronomy, University of Massachusetts, 710 North Pleasant Street, Amherst, MA 01003, USA
[3] Department of Physics, Yale University, P.O. Box 208121, New Haven, CT 06520, USA
[4] Yale Center for Astronomy & Astrophysics, Physics Department, P.O. Box 208120, New Haven, CT 06520, USA
[5] Instituto de Física y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Gran Bretaña 1111, Valparaíso, Chile
[6] CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, 230026 Hefei, Anhui, China
[7] Visiting scholar, Department of Astronomy, University of Massachusetts, 710 North Pleasant Street, Amherst, MA 01003, USA
[8] INAF—Osservatorio Astronomico di Bologna, Bologna, Italy
Received 2016 July 20; revised 2017 March 8; accepted 2017 March 9; published 2017 March 30

Abstract

We report the detection of morphology-dependent stellar age in massive quenched galaxies (QGs) at $z \sim 1.2$. The sense of the dependence is that compact QGs are 0.5–2 Gyr older than normal-sized ones. The evidence comes from three different age indicators—D_n4000, H_δ, and fits to spectral synthesis models—applied to their stacked optical spectra. All age indicators consistently show that the stellar populations of compact QGs are older than those of their normal-sized counterparts. We detect weak [O II] emission in a fraction of QGs, and the strength of the line, when present, is similar between the two samples; however, compact galaxies exhibit a significantly lower frequency of [O II] emission than normal ones. Fractions of both samples are individually detected in 7 Ms *Chandra* X-ray images (luminosities $\sim 10^{40}$–10^{41} erg s^{-1}). The 7 Ms stacks of nondetected galaxies show similarly low luminosities in the soft band only, consistent with a hot gas origin for the X-ray emission. While both [O II] emitters and nonemitters are also X-ray sources among normal galaxies, no compact galaxy with [O II] emission is an X-ray source, arguing against an active galactic nucleus (AGN) powering the line in compact galaxies. We interpret the [O II] properties as further evidence that compact galaxies are older and further along in the process of quenching star formation and suppressing gas accretion. Finally, we argue that the older age of compact QGs is evidence of progenitor bias: compact QGs simply reflect the smaller sizes of galaxies at their earlier quenching epoch, with stellar density most likely having nothing directly to do with cessation of star formation.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: high-redshift

1. Introduction

The formation and evolution of massive early-type galaxies remains poorly understood, despite much recent progress. Constraints from the local universe indicate that their stellar ages are very old (>10 Gyr), suggesting that they formed the bulk of their stellar masses at $z > 2$ (Bower et al. 1992; Renzini et al. 1993; van Dokkum & Ellis 2003; Heavens et al. 2004; Renzini 2006) subsequently quenched star formation, and remained quenched until the present. Constraints on stellar abundance ratios (high α/Fe) indicate that their star formation took place on short timescales (Thomas et al. 2005, 2010; Renzini 2006). Additionally, it has been observed that galaxy morphology and star formation properties are highly correlated, such that this quenched nature in massive galaxies appears coincident with a morphological transformation to an ellipsoidal stellar structure (e.g., Strateva et al. 2001; Kauffmann et al. 2003b; Franx et al. 2008). Despite efforts to study this transition from star-forming galaxy to quenched ellipsoid, we have gained very little insight into both the transformational quenching process and the mechanisms preventing further star formation for the majority of the universe's history.

Of particular importance to this effort are constraints from observing the progenitors of these massive early-type galaxies at $z > 1$, shortly after their transformation from star-forming galaxies. Recently quenched galaxies (QGs) begin to appear in large numbers at $z \sim 2$ (Cimatti et al. 2008; van Dokkum et al. 2008; Cassata et al. 2011, 2013) and have even been observed out to $z \sim 3$–4 (Fontana et al. 2009; Gobat et al. 2012; Guo et al. 2012; Muzzin et al. 2013; Stefanon et al. 2013; Straatman et al. 2014). The properties of these quenched high-redshift galaxies provide significant insight into both the formation process of the galaxies during the star formation phase and the quenching mechanisms causing their transformation. The most striking feature of these recently QGs at high redshift is their stellar structure; while already having built up an amount of stellar mass similar to that of their $z \sim 0$ counterparts, they are remarkably compact in stellar density (Daddi et al. 2005; Bundy et al. 2006; Trujillo et al. 2006; Toft et al. 2007; Zirm et al. 2007; Cimatti et al. 2008; van der Wel et al. 2008; van Dokkum et al. 2008; Bezanson et al. 2009; Damjanov et al. 2009; Saracco et al. 2009; Williams et al. 2010). The overwhelming majority of QGs (>80%) at $z > 1.5$ have stellar densities higher than the lower 1σ of passive (early-type) galaxies at $z \sim 0$ at the same stellar mass (Cassata et al. 2013). Additionally, they are much smaller than the majority of massive star-forming galaxies at the same epoch

[9] Hubble Fellow.

(van der Wel et al. 2014), and, in fact, one of the strongest predictors of quiescence among high-redshift galaxies is centrally concentrated light (i.e., a measure of compactness; Franx et al. 2008; Bell et al. 2012; Omand et al. 2014; Teimoorinia et al. 2016; Whitaker et al. 2016). It appears, therefore, that compactness and quenched nature at high redshift are inextricably linked.

However, the physical reason for this correlation is also poorly understood. Does the existence of the compact QGs at high redshift imply something very fundamental about quenching? In other words, does some physical process associated with stellar compactness predispose galaxies to quench? Alternatively, are the earliest galaxies to form and complete their evolution simply the densest because the universe was denser at early times (e.g., Lilly & Carollo 2016), or is it because of some highly dissipative gaseous process that could take place predominantly at high redshift (e.g., Dekel et al. 2009; Johansson et al. 2012; Dekel & Burkert 2014; Ceverino et al. 2015; Zolotov et al. 2015)?

There are some physically motivated reasons to believe that the former may be true. First, high stellar density implies a previous epoch of high surface density of star formation, which would mean a higher energy input into the interstellar medium (ISM) of compact galaxies than might be present in larger, extended galaxies. Hopkins et al. (2010) made this argument based on the observation that there appears to be a maximum stellar density for any structure in the universe ($\Sigma \sim 10^{11}\,M_{\odot}\,\mathrm{kpc}^{-2}$). This limit in stellar density exists despite covering eight orders of magnitude in stellar mass, from star clusters within galaxies to the $z > 2$ compact QGs. This empirical limit argues for some stellar feedback process, such as massive stellar winds, that truncates star formation and prevents further growth beyond this density limit. Studies of objects with high surface density of star formation, where such extreme stellar feedback might be expected, have in fact found evidence of feedback in the form of very fast ($\sim$1000 km s^{-1}) galactic-scale outflows from extremely compact star-forming regions that approach the Eddington limit (Diamond-Stanic et al. 2012; Sell et al. 2014).

An alternative scenario is that this connection is a very simple consequence of the size evolution of star-forming galaxies, whose radii (at fixed mass) are observed to decrease with increasing redshift (e.g., van der Wel et al. 2014). The most massive galaxies in the early universe formed the earliest in cosmic time and therefore evolved to the end of their star-forming lifetimes earliest. In this scenario, the density of the parent halo of the quenched population at any redshift reflects the density of the universe at its formation epoch (e.g., Mo et al. 1998) and therefore will progressively increase in size (and thus stellar density) over time. Such a scenario, known as progenitor bias (as described by Lilly & Carollo 2016), would also contribute to the increasing size evolution of QGs over cosmic time (Valentinuzzi et al. 2010a, 2010b; Carollo et al. 2013; Poggianti et al. 2013; Lilly & Carollo 2016; see also Bezanson et al. 2009). In this scenario, the significance of compactness is irrelevant for quenching; rather, galaxies quench when they have reached a sufficient mass to no longer support star formation. The quenching mechanism, then, may be related to halo mass or some other mass-related mechanism to cut off the gas supply for future star formation (e.g., Birnboim & Dekel 2003; Dekel & Birnboim 2006; Peng et al. 2010). Distinguishing between these scenarios is highly important for understanding the evolution of early-type galaxies.

Each of these two scenarios has empirical predictions for the properties of QGs. In the stellar density–regulated star formation scenario, galaxies with a high enough surface density of star formation will quench and produce remnants with high stellar density. Thus, at any given epoch, the most recently quenched objects should also be the densest (e.g., Whitaker et al. 2012). There is no explicit prediction for a trend of stellar density with stellar age or mass (e.g., Hopkins et al. 2010). However, progenitor bias explicitly predicts that stellar age and stellar density are correlated, with the densest objects also being the oldest at any given epoch (in the absence of size growth via merging; Lilly & Carollo 2016). In this paper, we seek to distinguish between these two scenarios and, in the process, gain insight into why galaxies quench their star formation early in cosmic time. In Section 2, we present the data used in this study. In Section 3, we present our results; in Section 4, we discuss these results in the context of quenching and the formation of QGs. Throughout this paper, we assume a cosmology with $\Omega_{\Lambda} = 0.7$, $\Omega_{M} = 0.3$, and $H_{o} = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2. Data

2.1. Samples

We select QGs from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) data (Grogin et al. 2011; Koekemoer et al. 2011) in GOODS-South according to the selection outlined in Cassata et al. (2011), which identified 179 QGs at $z > 1$ with $M^{*} > 10^{10}\,M_{\odot}$, specific star formation rate (sSFR) $< 10^{-2}\,\mathrm{Gyr}^{-1}$, and centrally concentrated spheroidal morphologies. For this study, we measure the properties of this sample using the CANDELS multiwavelength photometry (Guo et al. 2013) following the spectral energy distribution (SED) fitting procedure outlined in B. Lee et al. (2017, in preparation) using the SpeedyMC Bayesian SED-fitting software (Acquaviva et al. 2011), where star formation history is left as a free parameter. We use the morphologies measured from the *Hubble Space Telescope (HST)*/ Wide Field Camera 3 (WFC3) F160W H-band imaging from CANDELS using GALFIT (Peng et al. 2002) presented in van der Wel et al. (2012). In particular, we define the size to be the circularized half-light radius in kpc, $R_{\mathrm{eff}} = r_{e}\sqrt{b/a}$, where r_{e} is the length of the semimajor axis in arcseconds converted to kpc using the spectroscopic redshift, and b/a is the axis ratio.

From this parent sample, we identify 61 QGs with complementary publicly available spectra (see next section). We define a "compactness" cut on the QG sample according to the local $z \sim 0$ size–mass relation for local galaxies from the Sloan Digital Sky Survey (SDSS). We define QGs as compact if their size (at a given mass) is smaller than the lower 1σ of the local early-type galaxy size–mass relation (Shen et al. 2003). This roughly corresponds to a stellar mass surface density of $\Sigma \sim 3 \times 10^{9}\,M_{\odot}\,\mathrm{kpc}^{-2}$. We refer to any QG with a stellar density higher than this threshold as compact. QGs that are more extended, and thus similar to the majority of local SDSS early-type galaxies in mass and size, are referred to as normal. The position of the QGs in the size–mass diagram, illustrating the compactness cut from the SDSS, are presented in Figure 1. There are 28 compact galaxies and 33 normal-sized

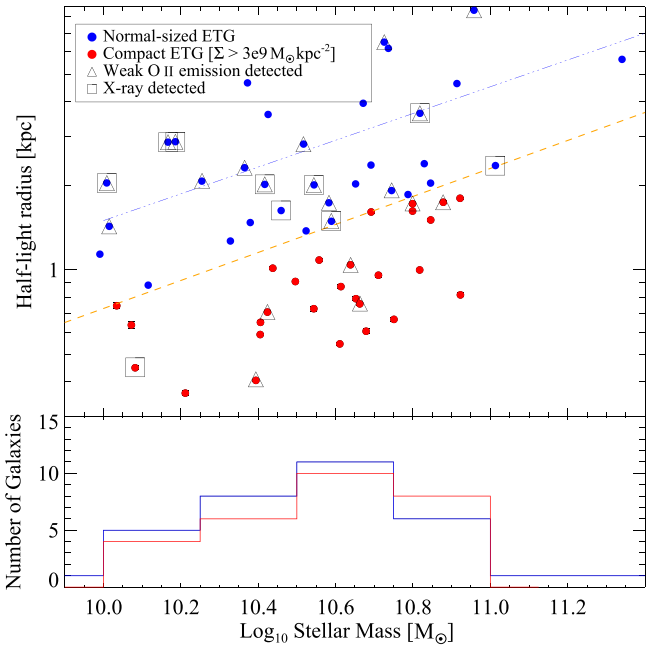

Figure 1. Size–mass relation of $z \sim 1.2$ QGs in this study. In red are the QGs defined as compact according to Cassata et al. (2013) as being below the lower 1σ of the $z \sim 0$ early-type galaxy size–mass relation (orange dashed line; Shen et al. 2003). The $z \sim 0$ mean early-type galaxy size–mass relation is shown by the blue dotted–dashed line (Shen et al. 2003). Blue galaxies are considered normal-sized QGs (relative to early-type galaxies at $z \sim 0$). Triangles designate galaxies in which [O II] λ3727 emission was detected. Squares designate galaxies with X-ray detections. Bottom panel: the mass distributions of each sample are roughly equivalent.

galaxies. Average redshifts for the samples are $\langle z \rangle = 1.22$ and 1.13 for compact and normal, respectively.

2.2. Spectroscopy

Our spectroscopic data for this sample were obtained at the European Southern Observatory's Very Large Telescope as part of the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004) spectroscopic program. In particular, we use spectra from programs with FORS2 (Vanzella et al. 2005, 2006, 2008; Kurk et al. 2009, 2013) and VIMOS (Popesso et al. 2009; Balestra et al. 2010). The FORS2 spectra have an instrument resolution of $R \sim 660$, which is 13 Å at 8600 Å observed (average 5.9 at 3900 Å restframe; all spectra between redshifts 1 and 1.4). The spectra we use from VIMOS in the Balestra et al. (2010) and Popesso et al. (2009) release all use the medium-resolution grism, which has comparable resolution.

2.3. Stacking Procedure

To produce average composite spectra (stacks), we perform the following procedure. We first individually transform each spectrum into the restframe using the published redshifts (Vanzella et al. 2008; Kurk et al. 2009; Popesso et al. 2009). We then flux normalize each spectrum using the median flux value measured between restframe 4000 Å $< \lambda <$ 4050 Å. Finally, we stack the normalized spectra using the *scombine* package in Image Reduction and Analysis Facility (IRAF) and rebin to a common dispersion ($d\lambda = 1.4$ Å pixel^{-1}, restframe), performing a 3σ clipping during the stack. Our stacked composite spectra of the two QG samples are presented in Figure 2. There are six out of the 28 compact QGs that do not have spectral coverage redward of restframe 4000 Å; thus, we exclude them from the stack and its analysis presented in Section 3.2. However, their spectra are suitable for the analyses presented in Sections 3.3 and 3.4, which include the full sample of 28 galaxies.

We estimate the errors on the stacks as the sample standard deviations of each spectroscopic sample. To measure the standard deviations, we repeat the stacks of each sample using jackknife resampling, each time removing one spectrum and stacking the rest of the spectra following the identical procedure described above. The final sample error is the standard deviation of the jackknifed stacks at each spectral point.

3. Results

3.1. Composite Spectra

The stacks presented in Figure 2 exhibit many features typical of old stellar populations, namely, strong Balmer absorption typical of post-starburst galaxies, strong G-band absorption, and a prominent 4000 Å break. Additionally, very weak [O II] emission is occasionally present (as we discuss later), indicating little (if any) star formation in some cases. We discuss the [O II] properties of the galaxies further in Section 3.3. In Section 3.2, we show the properties of the stellar populations, including the average stellar population ages, in these two samples.

3.2. Stellar Ages

3.2.1. Estimates from Lick Indices

We measure the average age of the stellar populations of the two QG samples from their stacks using two age diagnostics: the 4000 Å break (D_n4000) and Hδ absorption. For the D_n4000 diagnostic, we use the age calibration presented in Balogh et al. (1999) and Kauffmann et al. (2003a), which takes the ratio of the mean flux between 4000 and 4100 Å to the mean flux between 3850 and 3950 Å in the stack. We measure the error on the D_n4000 diagnostic by generating 1000 Gaussian deviates of each spectral point in the stack using the observed sample error described in Section 2.3 and repeating the D_n4000 measurement each time. The error is then the standard deviation of the sample of Gaussian-deviated D_n4000 measurements. We find that the compact sample has a $D_n4000 = 1.45 \pm 0.03$, a larger measure (i.e., older age) than for the normal sample, from which we measure $D_n4000 = 1.398 \pm 0.002$ (although the difference is marginal, $\sim 1.7\sigma$). We compare these D_n4000 measurements to those made with single-age stellar population models using the same procedure (Figure 3; Kauffmann et al. 2003a). Adapted from Kauffmann et al. (2003a), the left panels of the figure present the evolution of this diagnostic as a function of stellar age at solar metallicity for a single instantaneous burst of star formation according to the STELIB library (solid lines; Le Borgne et al. 2003), the Pickles (1998) library (dotted lines), and Jacoby et al. (1984; dashed lines). In the right panels, the age evolution of the diagnostic is presented from the STELIB library for bursts of solar metallicity (solid lines), 20% solar (dotted lines), and 2.5 times solar (dashed lines). For all models, the D_n4000 of the compact sample implies an older stellar age than that of the normal sample.

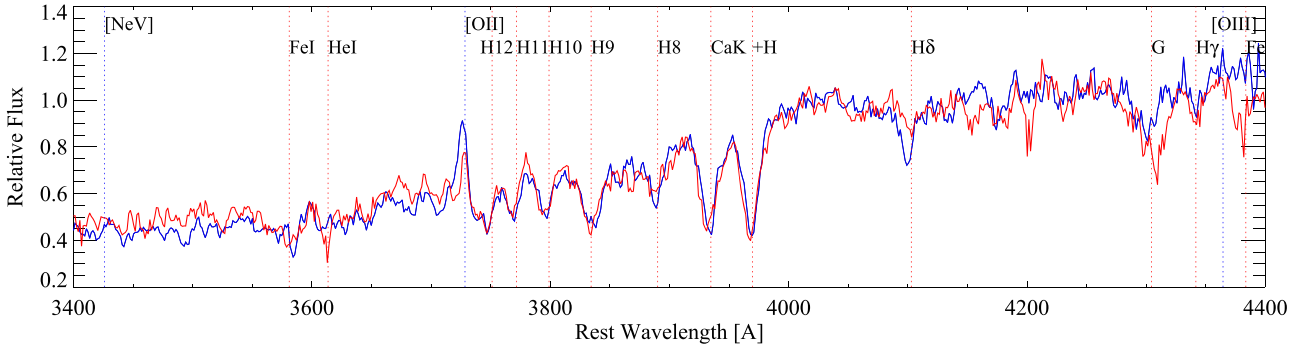

Figure 2. Stacks of QGs, separated by stellar density relative to the size–mass relation: compact galaxies (red) exhibit older spectral features than more extended galaxies (blue). Absorption features are identified by red dotted lines, and emission features are identified by blue dotted lines. The [O II] emission and Hδ absorption are stronger and the G-band is weaker in normal QGs.

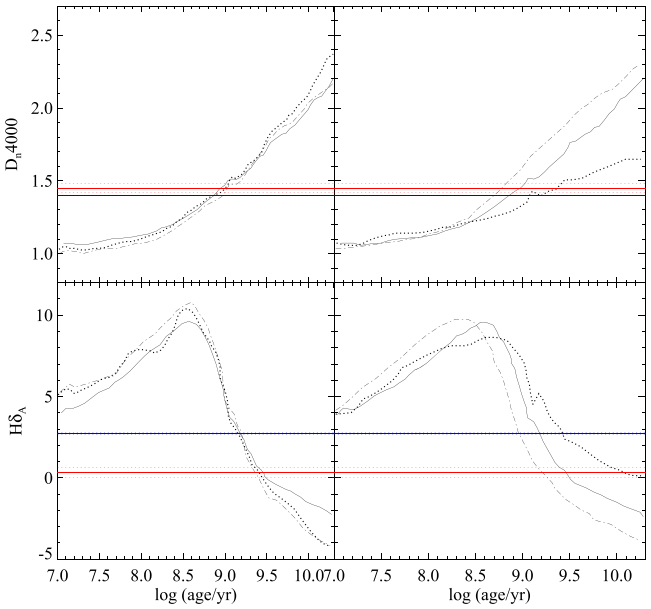

Figure 3. Evolution of both age diagnostics (D_n4000 and Hδ_A) as a function of stellar age following a single instantaneous burst of star formation using stellar population models of solar metallicity. Left: evolution of solar metallicity single-age stellar populations from the STELIB library (solid lines; Le Borgne et al. 2003), the Pickles (1998) library (dotted lines), and Jacoby et al. (1984; dashed lines). Right: evolution of age diagnostics in the STELIB library for bursts of solar metallicity (solid lines), 20% solar (dotted lines), and 2.5 times solar (dashed lines). The observed age diagnostics of the two populations studied here are in red (compact) and blue (normal). The figure is adapted from Kauffmann et al. (2003a).

Similarly, we make the comparison using the age-sensitive Hδ absorption feature. Unlike the D_n4000 diagnostic, which grows monotonically with age, the Hδ (as well as other Balmer) absorption feature peaks for stellar populations of age $\sim$1 Gyr (A-type stars). As the stellar population ages, the Balmer absorption starts to decrease in strength. We measure Hδ_A using the Lick absorption line index (Worthey & Ottaviani 1997). To estimate the error on our measurement of Hδ_A, we follow the same procedure as for D_n4000 by remeasuring the index from 1000 Gaussian deviates of each spectral point and taking the standard deviation of these Hδ_A measurements. We find that the compact sample has H$\delta_A = 0.33 \pm 0.31$, a smaller value (i.e., older age) than for the normal sample, from which we measure H$\delta_A = 2.74 \pm 0.06$. The bottom panels of Figure 3 show how these measured values compare to those from the evolutionary models explored in Kauffmann et al. (2003a).

Although the trends between the age diagnostics and the stellar population age are obvious for each of the models shown in Figure 3, it is also obvious that the diagnostics depend on other features of the models, such as metallicity. Therefore, we do not attempt to use these measurements as average age measurements of the galaxy populations. Rather, we seek to gain insight into relative age differences between the two samples using the few models presented in Figure 3. At face value, the average age implied by the six models from D_n4000 suggests an age difference between the two galaxy samples of roughly $\sim$0.3 Gyr. The Hδ_A index suggests a larger age difference, with the compact sample $\sim$2.5 Gyr older than the normal QG sample (we discuss in Section 3.2.2 why the difference implied by Hδ_A is likely overestimated). However, it is clear that both age indicators imply that the compact sample, on average, has older stellar ages than the normal sample.

We verify that the qualitative results are robust to the compactness definition by increasing it to 1.2σ below the $z \sim 0$ mean, which corresponds to roughly $5.8 \times 10^9\, M_\odot\, \mathrm{kpc}^{-2}$, a factor of $\sim$2 denser than the definition outlined above. This splits the compact sample defined by the 1σ line roughly in half, resulting in 16 compact galaxies and 48 extended ones. With the smaller sample, there is a significant decrease in the signal to noise of the compact sample with respect to the extended one, but we are able to measure age diagnostics. We find that the D_n4000 of the more stringently selected compact galaxies increases on average, suggesting an older age, although within the errors of the previous measurement. The Hδ_A measurement in the more stringently selected compact sample increases slightly, in the sense of younger age; however, it is again within the errors. We note that the error on the compact Hδ_A measurement is larger with the more stringent selection, likely because Hδ_A is typically not individually detected in the spectra, and this sample is small. Both age diagnostics of the extended sample from the more stringent selection change toward older stellar age; this makes sense in the context of our interpretation, because we have essentially added 12 (formerly) compact galaxies whose average age is older to a younger sample, and the expected effect would be to increase the age. This is, in fact, what we see. We interpret the changes in diagnostics from the stringently selected compact sample as consistent with this picture: the D_n4000 shows an increase in age (although consistent within the errors of the 1σ selected compact sample), and the noise in the Hδ_A

measurement increases due to decreased signal to noise from the small sample.

3.2.2. Other Spectral Features

In the literature, it has been extensively discussed that other spectral features adjacent to the Hδ absorption line can affect the measurement of the Hδ_A Lick index (e.g., Dressler et al. 2004; Prochaska et al. 2007). In particular, stellar continuum absorption by molecular CH and CN lines to the blue and red of Hδ can depress the pseudocontinuum regions used to measure the index, resulting in an underestimated value for the index. The net consequence is an overestimated age based on the line. Although the nature of this continuum absorption is not well understood, it likely originates in old stellar populations from metal-enriched cool stars (Schiavon et al. 2002; Dressler et al. 2004).

As can be seen in Figure 2, we observe that the compact QG stack exhibits a prominent peak immediately blueward of the Hδ line, followed by a band of continuum absorption between 4100 and 4200 Å. The blue peak is an indication of the true continuum level, unaffected by the CN absorption. Such prominent features are not seen in the normal QG stack. This feature is undoubtedly affecting the Hδ_A measurement in the compact sample, such that the implied age is overestimated (and, therefore, the relative age difference as well).

However, this evidence of old, metal-enriched stellar populations in the compact sample is in general agreement with the evidence for an older average stellar age than that in the normal QG sample. We discuss further constraints on this age differential between the two samples in the following section.

3.2.3. Stellar Population Synthesis Modeling

An alternative estimate for the age differential between the two populations can be obtained by fitting stellar population templates to the stacks. In this section, we use the Penalized Pixel-Fitting (pPXF) software (Cappellari & Emsellem 2004) to explore the average ages of the two galaxy samples, again with the goal of gaining insight into relative age differences. As templates for the fitting, we use the Vazdekis et al. (2010) Simple Stellar Population (SSP) models based on the Medium resolution INT Library of Empirical Spectra (MILES; Sánchez-Blázquez 2006). Due to the poorly understood nature of the CN and CH stellar continuum absorption features, we choose to use these templates because they are composed of empirical stellar spectra where these features are observed (e.g., Vazdekis 1999).

We allow pPXF to choose the "optimal" combination of templates, with multiplicative and additive polynomials, to fit the continuum shape, of each stacked spectrum, where the initial mass function is fixed with a slope of 1.3 (Salpeter 1955), and we constrain the maximum age of any template to be the age of the universe at $z = 1.2$. We allow pPXF to choose from a range of templates of varying ages, which are weighted and summed together to construct the best-fitting model. The signal to noise of our stacks starts to deteriorate at wavelengths longer than restframe 4800 Å from a combination of decreased number of spectra with wavelength coverage in that region and possibly poor subtraction of telluric features, and so we constrain the spectral region provided to pPXF to 3500–4800 Å, masking out 40 Å surrounding the [O II] λ3727 emission.

To choose the appropriate metallicity range for the fitting, we use constraints from the mass–stellar metallicity relation out to $z \sim 0.7$ (Gallazzi et al. 2005, 2014) at our average stellar masses ($\langle M_* \rangle \sim 10.7$ for both samples). Only the solar metallicity models in Vazdekis et al. (2010, using the Padova +00 isochrones of Girardi et al. 2000) fall within the confidence intervals of the Gallazzi et al. (2014) mass–metallicity relation at this average mass. Therefore, we limit our fits to these models. The pPXF software measures the luminosity-weighted age based on the sum of two templates, which assumes that galaxies may be composed of several stellar populations. Those individual populations may not necessarily follow the mass–stellar metallicity relation. However, we note that our results do not strongly depend on either the wavelength range or metallicity assumed by the fit.

We estimate errors on the age of the stellar populations given by the fit as the variance of a series of bootstrap resampled templates added to the fit residuals. We find that the compact stack has an average age of 2.54 ± 0.63 Gyr (reduced $\chi^2 = 15.3$), while the normal sample is 1.87 ± 0.65 Gyr (reduced $\chi^2 = 30.2$), an age differential that is in general agreement with the findings using the Lick indices in Section 3.2.1 in that the age measured from the compact sample is older. These results are shown in Figure 4.

To conclude this section, our results from both the spectral indices and the full spectral fitting consistently indicate that there is evidence that the average age of the compact QG sample may be older than that of the normal QG sample.

3.3. [O II] Emission

In the stacked spectra presented in Figure 2, we find that each QG sample exhibits weak [O II] λ3727 emission. We visually inspect each of the individual spectra in each sample, finding that only a fraction of them exhibit detected [O II] emission and that, in general, it always appears weak. The right panel of Figure 5 shows the estimated star formation rate (SFR) from the [O II] flux, assuming that the emission is produced entirely by star formation (Kennicutt 1998). The [O II] flux was measured from the individual spectra using the IRAF routine *splot*. The [O II] is always weak (corresponding to $\lesssim 2\,M_\odot\,\mathrm{yr}^{-1}$; for the majority of the sample, this is less than predicted from the photometry from the best-fitting SED). It is therefore unlikely that this [O II] emission is contributed to significantly by AGNs (see Section 3.4 for additional constraints). There is no obvious difference in the line luminosities or SFRs between the two samples (Figure 5).

However, the occurrence of [O II] emission is much more frequent among the normal QG sample than among the compact sample (top panel of Figure 1 and left panel of Figure 5). This is the primary reason for the stronger average emission in the normal QG stack. Out of 33 normal QGs, roughly half (15) exhibit detected [O II] emission. Among the 28 compact QGs, only six have detected [O II] (21%).

We conduct the following analyses to assess the significance of the [O II] detection rate among QGs of different stellar density. A priori, we do not have reason to believe [O II] emission should depend on compactness, and so we test the hypothesis that the chance of [O II] emission is random; i.e., there is an equal chance (50%) that any galaxy emits [O II] as that it does not. To test this hypothesis, we conduct a Monte Carlo simulation where we randomly draw samples equal to the number in each QG sample from a binomial distribution. We

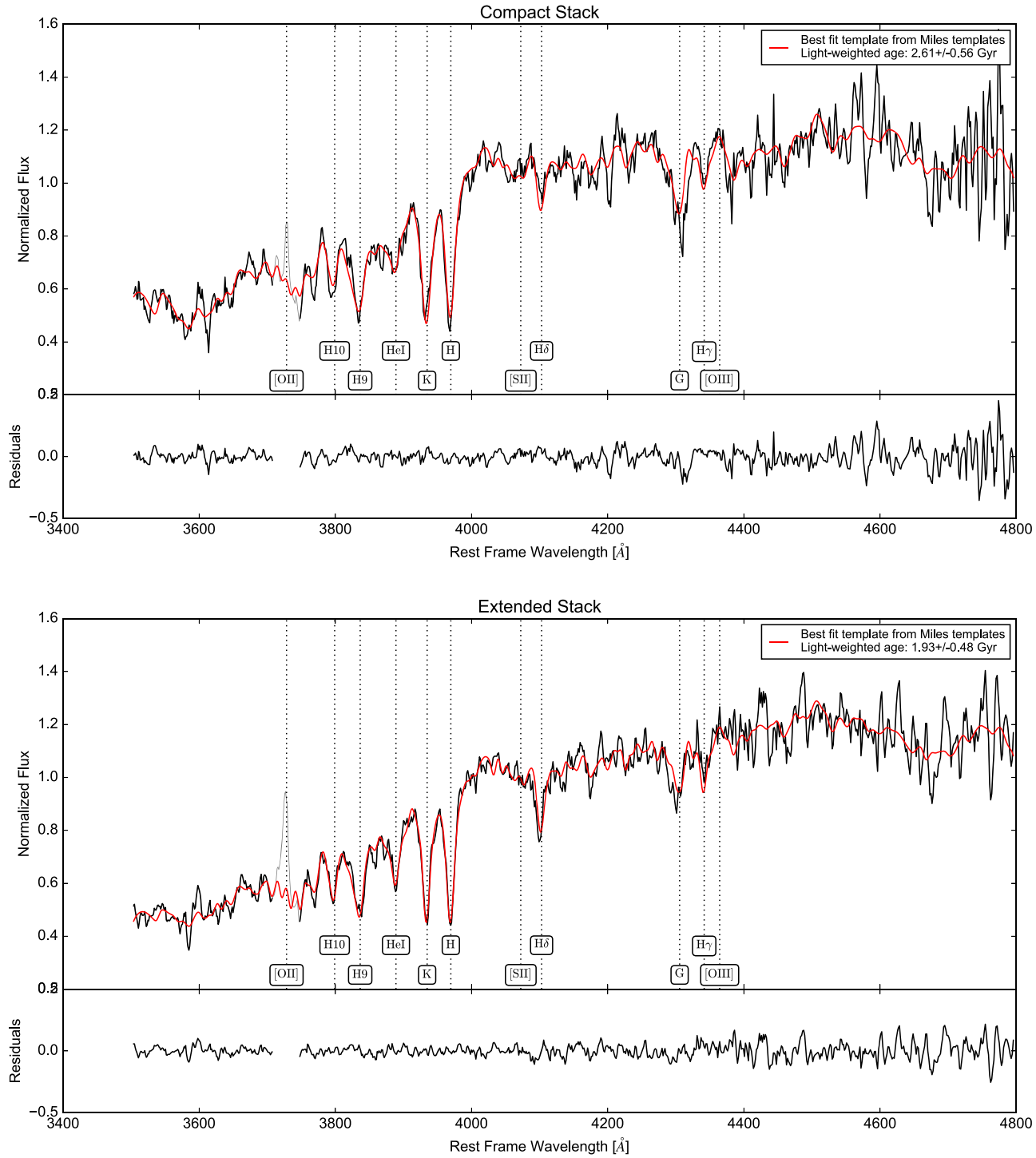

Figure 4. Results from the pPXF fitting of the two stacks. Top panels illustrate the best-fit model to the compact QG stack. The region around the [O II] emission line was excluded from the fit. Bottom panels illustrate the fit for the stack of normal-sized QGs. The light-weighted best fit ages between the two samples for the variety of fits indicate an older average age than that of the compact sample.

create 10,000 realizations and compare the success rate (i.e., an [O II] detection) to the observed [O II] detection rate in each QG sample. The results of this test are presented in Figure 6. The histograms in the top panel show the frequency of [O II] detection from a sample size of 33 (normal QGs; blue) and from a sample size of 28 (compact QGs; red). The dotted lines show the standard deviations of the realizations. The dashed lines show the observed frequency. It is clear from the figure that the normal QGs show an [O II] frequency that is consistent with the hypothesis that [O II] detection is random (50% detection rate). In contrast, the compact QGs show an [O II] frequency that is significantly lower, inconsistent at the 3σ level.

We also investigate the significance of the differing [O II] detection rate between the two samples using the Fisher exact test to measure the probability that the two QG samples are

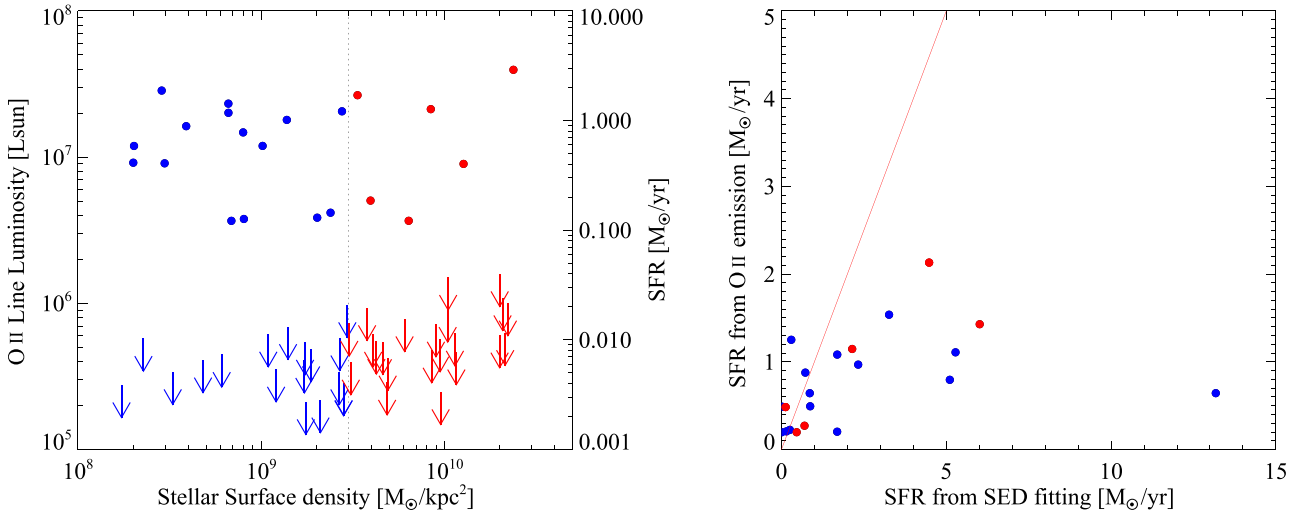

Figure 5. Left: the [O II] emission line fluxes (when detected) plotted vs. the stellar surface density (compactness). There is no trend of flux with compactness. Upper limits for spectra without detected [O II] emission are indicated by the top of the downward arrows. Right: SFR predicted from the [O II] line luminosities (in galaxies with [O II] detections only; Kewley et al. 2004) compared to the SFR from SED fitting (B. Lee et al. 2017, in preparation). Few galaxies show [O II]-derived SFRs in excess of that estimated from broadband photometry; therefore, it is unlikely that AGNs contribute significantly to the [O II] luminosity.

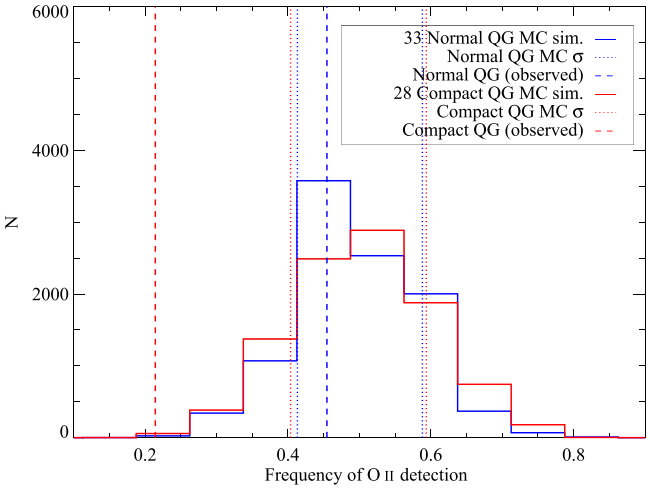

Figure 6. Results of a Monte Carlo simulation to test the hypothesis that [O II] detection in a sample is random. Histograms are the distribution of frequencies expected using 10,000 simulated QG samples if the intrinsic detection rate is 50% for a sample that is the same size as the compact ones (red) and the normal ones (blue). Each distribution has a mean of roughly 50%. Standard deviations of each Monte Carlo distribution are indicated by the dotted lines. The observed frequency in the real data is indicated by the dashed lines. Normal QGs show a detection rate consistent with a random [O II] occurrence, but the compact QGs show an [O II] frequency that is significantly lower.

drawn from the same distribution of [O II] emitters. We construct a 2×2 contingency table for each QG sample and calculate the associated p-value from the Fisher exact test. We use the function *fisher.test* in the *R* statistical software environment. We calculate a p-value of 0.06, indicating that we can reject the null hypothesis that the two samples come from the same parent sample of [O II] emitters. The occurrence rate of [O II] is significantly different between the two samples at the $\sim 2\sigma$ level, according to this test. We conclude that the frequency of [O II] detection among compact QGs differs significantly from that among the more extended sample. Whatever produces the [O II] emission in QGs, whether it is warm gas or an energy source such as residual or rejuvenated star formation (SF), it appears to be significantly less frequent in compact galaxies than in the normal sample.

3.4. X-Ray Properties

The X-ray properties of the two QG samples provide further insight into residual energy sources in these galaxies, which may or may not be related to the quenching process that shut down the star formation or the origin of the [O II] emission when present. The majority of our QGs are undetected (27 out of 28 compact QGs; 24 out of 33 normal QGs) in the *Chandra* 7 Ms data in GOODS-South (Luo et al. 2017). For the X-ray counterpart identification, we have made use of two catalogs. First, we have used the 4 Ms catalog by Cappelluti et al. (2016), which is able to include significantly fainter X-ray sources than blind detections (e.g., Xue et al. 2011) due to a novel technique based on prior information on the positions of optical-near-infrared (NIR) sources. Second, we have used the 7 Ms source catalog presented in Luo et al. (2017). There is excellent correspondence between the two catalogs. All galaxies detected in the 4 Ms catalog are also detected in the 7 Ms data, with the 7 Ms data providing one extra detection not present in the 4 Ms data. The 7 Ms X-ray–detected galaxies are indicated by squares in Figure 1, where we also identify those with significantly detected [O II] emission. We note that, among the compact sample, none of the [O II] detections come from an X-ray–detected galaxy. The one X-ray detection exhibits no [O II] emission. In the normal sample, a large fraction (seven out of 15) of the [O II] detections come from an X-ray–detected galaxy. The majority of X-ray detections (seven out of nine) are [O II] emitters. We present the 7 Ms X-ray fluxes, hardness ratios, and [O II] luminosities for all X-ray–detected sources in Table 1. In general, it does not appear that [O II]-detected galaxies exhibit any obvious difference in their X-ray properties from those without [O II] emission. Similarly, there does not appear to be an obvious distinction in X-ray properties between the one detected compact QG and the normal QG detections (see also Figure 7). To estimate the upper limits to the [O II] line flux, we first measure the rms in the continuum in the vicinity of the line and multiply by the

Table 1
Properties of X-Ray–detected Galaxies in 7 Ms *Chandra* Data

Galaxy Sample	ID[a]	Soft Flux[b]	Soft Luminosity[c]	Hard Flux	Hard Luminosity	Hardness Ratio	[O II] Luminosity[d]
Compact	559	$2.16^{+0.68}_{-0.55}$	9.07E+40	$3.46^{+1.98}_{-1.58}$	1.45E+41	−0.27	−99
Normal	515	1.74	3.28E+40	$8.60^{+4.12}_{-3.54}$	1.62E+41	0.10	1.19E+07
	555	$3.21^{+1.06}_{-0.88}$	4.37E+40	11.56	1.57E+41	0.10	1.63E+07
	616	$2.69^{+0.78}_{-0.63}$	3.73E+40	$9.00^{+2.93}_{-2.40}$	1.25E+41	0.02	−99
	745	$3.64^{+0.87}_{-0.73}$	7.13E+40	$4.55^{+2.24}_{-1.84}$	8.91E+40	−0.36	1.48E+07
	861	$3.95^{+1.00}_{-0.87}$	5.66E+40	$13.49^{+4.33}_{-3.88}$	1.93E+41	0.02	1.80E+07
	881	1.69	2.91E+40	$17.86^{+6.29}_{-5.73}$	3.09E+41	0.39	2.85E+07
	448	$2.88^{+1.18}_{-1.06}$	7.18E+40	$166.27^{+14.24}_{-13.58}$	4.15E+42	0.82	−99
	574	$2.36^{+0.76}_{-0.61}$	3.24E+40	3.97	5.45E+40	−0.19	2.06E+07
	594	$5.14^{+1.15}_{-1.02}$	7.42E+40	$16.30^{+4.76}_{-4.30}$	2.35E+41	−0.01	9.08E+06

Notes.
[a] X-ray properties from catalog published in Luo et al. (2017).
[b] Soft and hard fluxes in 1e-17 [erg s^{-1} cm^{-2}]. No uncertainties indicate an upper limit on the flux.
[c] X-ray luminosities in erg s^{-1}. No flux uncertainties indicate the luminosity is an upper limit.
[d] [$L_{\odot}$].

square root of the number of pixels per resolution element (4 and 5 pixels for FORS2 and VIMOS, respectively).

We measure hardness ratios from hard- and soft-band fluxes for the detections as (Hard − Soft)/(Hard + Soft), where the fluxes are in counts s^{-1}. In general, AGNs, depending on type, can span a range of hardness ratios (e.g., Szokoly et al. 2004), as do our detected QGs, although the luminosities of our QGs are relatively weak for AGNs (Mainieri et al. 2002; Szokoly et al. 2004; Hasinger 2008). The left panel of Figure 7 shows that the hardness ratio of our sample increases with hard X-ray luminosity, indicating that, in general, an increase in hardness ratio in these samples is driven by an increase in hard-band luminosity. For comparison, we include all X-ray detections in the catalog of Luo et al. (2017), where detections have been grayscaled according to their photometric redshift (or spectroscopic redshift, if available). By number, the X-ray detections from Luo et al. (2017) are dominated by AGNs, but at the low-luminosity end, star-forming galaxies are abundant.

The significance of the X-ray detection rate among the two samples can again be assessed using the Fisher exact test, as in Section 3.3. Using a 2 × 2 contingency table for the X-ray detection rate among the two QG samples, we find a p-value of 0.015, corresponding to a significant difference between the two samples at the $\sim 2.4\sigma$ level. We also compute the significance jointly with [O II] using a 4 × 2 contingency table, where the rows represent the number of galaxies in each sample that have both X-ray and [O II] detections, [O II] detections only, X-ray detections only, and no detections. We find a p-value of 0.02 from the Fisher exact test, indicating that the samples differ at the $\sim 2.3\sigma$ level, consistent with the findings from assessing the [O II] and X-ray detection rates individually.

For the nondetected sources, we stack the 7 Ms X-ray images at the position of the optical-NIR sources to gain insight into the average X-ray properties of the two samples using the *Chandra* stacking analysis tool CSTACK[10] v4.3. Detected galaxies are excluded from the stack. We remove one compact QG from the stack due to its proximity to a very off-axis, bright X-ray source, where we suspect that leaking flux after source removal using the point spread function (PSF) has affected the counts in our stack. The results of the stacking analysis are presented in Table 2. We find significant stacked flux in the soft band for both samples, but both samples are essentially undetected in the hard band (the normal sample has a marginal detection with a signal to noise of ∼1.4). The stacked soft fluxes for both samples do not differ significantly from each other. The luminosities, assuming the average redshifts in each sample, are very low level, inconsistent with the presence of powerful AGNs. The stacked soft-band flux (and lack of hard-band detection), which represents the average of the majority of our samples, seems to indicate that the X-ray emission is soft. Sources of soft X-ray emission in non-star-forming galaxies include low ionization nebular emission regions (LINERs; Heckman 1980) and bremsstrahlung emission from hot (∼1 keV) ISM or halo gas. Such hot gas emission at the luminosities we observe here has been observed in local QGs (e.g., Fabbiano et al. 1992; Boroson et al. 2011; Kim & Fabbiano 2013). We discuss this possibility in more depth in Section 4.3.

4. Discussion and Conclusion

There are three main results in this study: (1) in the redshift range that we have considered, $1 \leqslant z \leqslant 1.4$, massive compact QGs have stellar populations that are, on average, older than those of normal-sized ones; (2) the frequency of [O II] and X-ray detection is significantly lower among the compact galaxies; and (3) the X-ray properties generally disfavor the presence of strong AGNs in both samples of recently QGs (low luminosities $\approx 10^{40}$–10^{41} erg s^{-1} from both the few X-ray detections and average stacked emission). While X-ray–detected normal QGs often also have [O II] emission, not a single compact galaxy with [O II] emission is individually detected in the *Chandra* images. This strongly argues against AGNs as the power source of the [O II] emission in compact galaxies and favors instead warm gas, stellar remnants, residual star formation, minor merging with a gas-rich companion, or LINER emission, possibly powered by stellar sources (e.g., Yan & Blanton 2012; Singh et al. 2013). These mechanisms may also be active in the normal galaxies. Although it is uncertain what the sources of the [O II] emission are, it is clear

[10] http://lambic.astrosen.unam.mx/cstack_v4.3

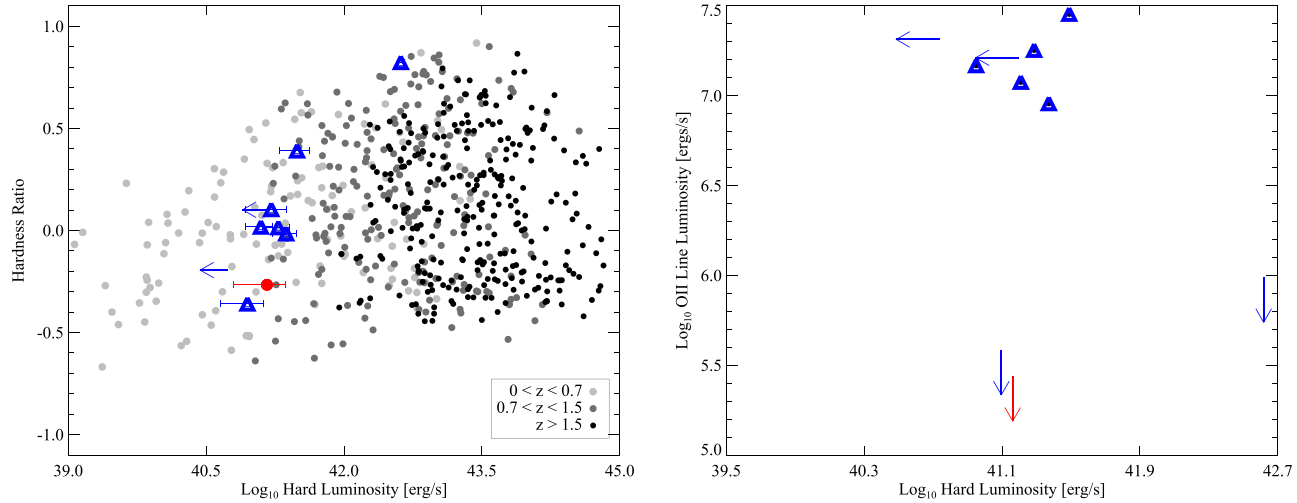

Figure 7. Left: hardness ratios vs. hard-band X-ray luminosities of the detected galaxies among the two samples. Red circles are compact QG detections, and blue triangles are normal QG detections. Gray points are X-ray detections at all redshifts in GOODS-South from Cappelluti et al. (2016). All but one X-ray–detected QG are too weak to be considered AGNs; their emission is likely due to some other source. Right: the [O II] line luminosity vs. hard X-ray luminosity for X-ray–detected galaxies (triangles) and upper limits to [O II] luminosity when it is not detected (arrows).

that they are less active among the compact sample. Taken all together, these lines of evidence paint a picture in which compact galaxies formed and evolved earlier than normal ones and, consequently, quenched star formation earlier.

4.1. Age Constraints: Evidence for Progenitor Bias

The evidence for the age difference between normal and compact QGs that we present here comes from two independent age diagnostics, D_n4000 and $H\delta_A$, as well as stellar population synthesis modeling. The D_n4000 is larger in compact galaxies than in normal ones, corresponding to an age difference of $\sim$0.3 Gyr using the calibrations by Kauffmann et al. (2003a). The difference in $H\delta_A$ implies a larger age differential ($\sim$2.5 Gyr) but, as discussed in Section 3, may be somewhat overestimated due to continuum absorption in old, metal-rich stars. Although these age conversions are model-dependent, the evidence for an age difference between samples is independent of the adopted stellar libraries and assumed metallicity for the range of values found here. Stellar population synthesis modeling with pPXF provides results consistent with those of the Lick indices. All age diagnostics considered here imply age differentials in the same direction, i.e., more compact passive galaxies are older.

The age differential between normal and compact QGs that we discuss here is at redshift $z \sim 1.2$. However, the result is in qualitative and quantitative agreement with the observations by Belli et al. (2015) at $z \sim 2$ that, at a given redshift, the largest galaxies (in radius, although related to stellar density) are among the youngest, suggesting that the property is a general feature of passive, massive galaxies at high redshift. Saracco et al. (2009) arrived at the same result via the opposite analysis: by separating the oldest QGs during the epoch $1 < z < 2$ from the youngest (δage $\sim$ 1.5–2 Gyr), they found that the youngest reside on the local $z \sim 0$ early-type galaxy mass–size relation (like our normal QG sample), whereas the old QGs are denser, with a radius that is a factor of 2.5–3 smaller than that of local early-type galaxies. They also noted that their QG samples differed in mass, with the younger sample being less massive (see also Thomas et al. 2010; Fagioli et al. 2016). The fact that we observe the same trend with essentially an identically mass-matched sample indicates that mass is not the primary factor related to the age differential in the population; rather, the stellar density or size may be the primary factor (Saracco et al. 2011 found a similar result). In a complementary study, Fagioli et al. (2016) found that, at lower redshifts than those of our sample ($0.2 < z < 0.8$), this trend of age and compactness in QGs persists in the stellar mass range explored here (however, Trujillo et al. 2011 did not find evidence for such a trend to $z \sim 0$).

At high redshift, compact galaxies dominate the population of QGs at the high-mass end (Cimatti et al. 2008; van Dokkum et al. 2010; Cassata et al. 2011, 2013; van der Wel et al. 2014). This has prompted investigations of scenarios in which -quenching is more efficient in galaxies with high stellar density because of increased stellar feedback (e.g., Hopkins et al. 2010). However, a causal relationship between high stellar density (compactness) and likelihood of quenching does not directly predict an age–density correlation; rather, at any given epoch, the densest but not necessarily oldest galaxies should be the most likely to quench (see, e.g., Whitaker et al. 2012; Yano et al. 2016).

The presence of a relationship between age and stellar density in which denser galaxies are older is precisely the prediction of the progenitor bias scenario (López-Sanjuan et al. 2012; Carollo et al. 2013; Poggianti et al. 2013; Belli et al. 2015; Keating et al. 2015; Wellons et al. 2015, 2016; Lilly & Carollo 2016). That is, due to the observed size evolution of star-forming galaxies (e.g., van der Wel et al. 2014), the density of a galaxy reflects the density of the universe when the galaxy formed (assuming very little or average structural disruption); therefore, older galaxies should be denser. Thus, at face value, our results support the idea that galaxies that form and complete their evolution earlier are simply denser than larger galaxies that form and evolve later, without the density necessarily having anything directly to do with the cessation of star formation activity. Star formation may then be affected by some other quenching agent (e.g., halo or mass quenching; Birnboim & Dekel 2003; Dekel & Birnboim 2006; Peng et al. 2010). This idea is extensively discussed in Lilly & Carollo (2016), who show by means of a

Table 2
Stacked 7 Ms X-Ray Fluxes

Flux (erg s^{-1} cm^{-2})[a]	Galaxy Sample	Soft Band	Hard Band
	Compact	$3.46 \pm 1.37 \times 10^{-18}$	$-0.30 \pm 1.73 \times 10^{-17}$
	Normal	$5.99 \pm 1.53 \times 10^{-18}$	$2.48 \pm 1.68 \times 10^{-17}$
Luminosity (erg s^{-1})[b]	Galaxy Sample	Soft Band	Hard Band
	Compact	$2.96 \pm 1.17 \times 10^{40}$	$-0.26 \pm 1.48 \times 10^{41}$
	Normal	$4.23 \pm 1.08 \times 10^{40}$	$1.75 \pm 1.19 \times 10^{41}$

Notes.
[a] Energy conversion factors evaluated using the *Chandra* PIMMS tool (http://cxc.harvard.edu/toolkit/pimms.jsp).
[b] Assuming the average redshift of each sample.

simple toy model that this scenario will naturally explain the correlations between galaxy structure and star formation properties without the need of a stellar density–related quenching mechanism (see also Abramson & Morishita 2016).

4.2. Energy Sources: Quenching Agents in QGs

An independent but complementary piece of information comes from the X-ray and [O II] properties of our two samples. In themselves, the [O II] emission line and the X-ray data do not provide any firm indication as to the causes of quenching. The average X-ray luminosity for the majority of our sample does not show evidence of any powerful AGNs, but one could have been present prior to reaching the current very low level of star formation activity.

There is evidence that the sources of ionizing radiation or warm gas that are still present in the two samples are different at the time of observation; namely, the compact galaxies show a significantly lower detection rate of [O II] and X-ray emission than the normal QGs. This difference is fully consistent with a scenario in which the quenching occurred earlier in the compact sample and has therefore had a longer time to fade. Larger QGs from the normal sample are more likely to exhibit emission from energizing sources simply because, on average, quenching in larger galaxies was initiated more recently. Alternatively, the presence of [O II] emission may be evidence of rejuvenated star formation due to gas from minor merging (e.g., Treu et al. 2002). In general, our data do not provide any conclusive constraints on the quenching mechanisms that truncated the star formation in these galaxies, nor do they determine whether the quenching mechanisms differ with stellar density.

It remains an important goal of galaxy evolution to understand the quenching mechanisms in massive galaxies. Over the last several years, efforts have been made to identify the star-forming progenitors of soon-to-be QGs at $z > 2$. These efforts have relied on the fortuitous observation that the first galaxies to quench are compact, making it relatively easy to identify their immediate star-forming progenitors among compact star-forming galaxies (Barro et al. 2013; Patel et al. 2013; Stefanon et al. 2013; Nelson et al. 2014; Williams et al. 2014; van Dokkum et al. 2015). It is interesting to note that various studies have come to very different conclusions about the nature of feedback in compact star-forming galaxies. Barro et al. (2013) claimed that the AGN fraction in compact star-forming galaxies may be up to 50%, suggesting that the presence of AGNs may truncate the star formation on short timescales. However, Spilker et al. (2016) discovered that a subset of this sample, despite being on the star-forming main sequence, has much lower molecular gas masses than normal main-sequence galaxies, suggesting that star formation will be quenched on short timescales due to simple gas exhaustion if the influx of gas has been suppressed. Additionally, Williams et al. (2015) found no evidence for AGNs among compact star-forming galaxies but instead detected both faster outflow velocities in the ISM and extreme redshifted Lyα emission among compact star-forming progenitors. (Similar Lyα signatures were identified among quenching galaxies by Taniguchi et al. 2015.) They interpreted these observational signatures as related to the compact galaxies having enhanced feedback in the ISM due to a higher surface density of star formation than their more extended counterparts (see also Diamond-Stanic et al. 2012; Sell et al. 2014; Alexandroff et al. 2015; Heckman & Borthakur 2016), plausibly leading to the truncation of future star formation.

From our data it is clear, however, that in the time since quenching was initiated in these QG samples, major energizing sources have already dissipated. Future detailed studies of galaxies closer to their quenching epoch may identify the physical processes that shut down star formation at high redshift.

4.3. X-Ray Emission from Passive Galaxies: AGNs, Binaries, or Hot Gas?

An important feature of both samples is that some of the QGs have X-ray emission detected in the deep *Chandra* 7 Ms data of GOODS-South. AGNs have long been suspected to be the key agent behind the quenching of star formation (e.g., Granato et al. 2004), as well as the prevention of future star formation in galaxies (e.g., Ciotti & Ostriker 1997; Fabian 2012). Thus, an obvious question is whether or not the X-ray properties of our QGs are consistent with this idea. Only a minority of the galaxies are individually detected (see Table 1); stacking the images at the position of the nondetected galaxies, however, yields measurable flux in the soft band but not in the hard band. While the range of hardness ratios spanned by our galaxies is the same as that of narrow-line AGNs and normal galaxies (e.g., Szokoly et al. 2004; Hasinger 2008), the distribution of hardness ratio and X-ray luminosity shown by our galaxies qualitatively looks very different from the distribution of the parent population of X-ray detections (see gray points in Figure 7; Luo et al. 2017). In number, these X-ray sources are dominated by AGNs with luminosities $>10^{42}$ erg s^{-1} (see Wilkes et al. 2013 for the distribution among more powerful obscured AGNs). The distribution exhibits a large scatter in hardness ratio and X-ray luminosity, most likely a reflection of the diversity of

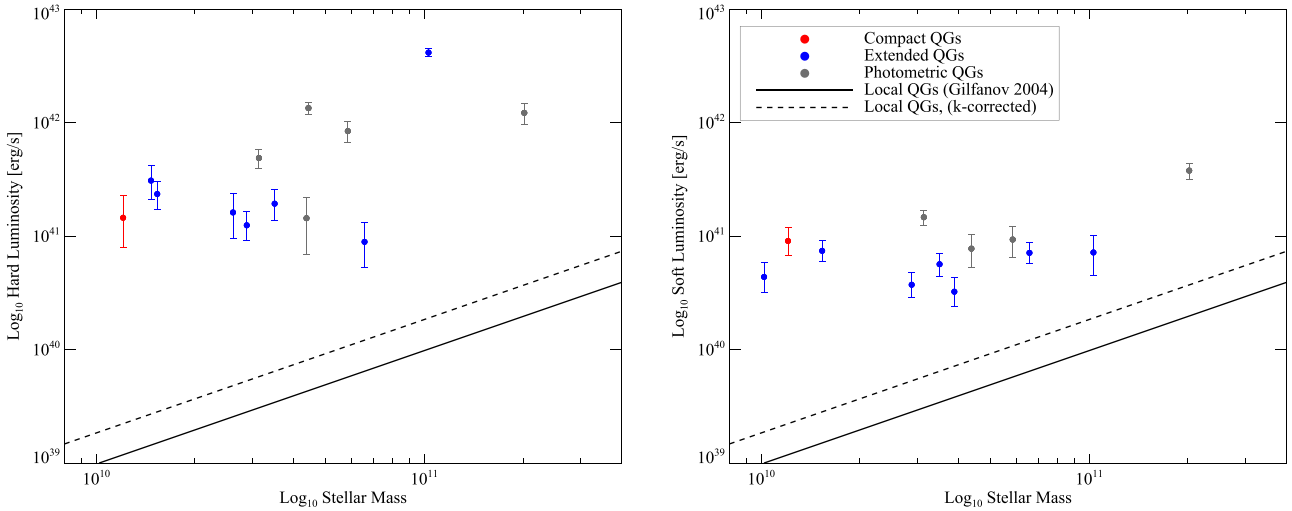

Figure 8. Left: the 7 Ms hard-band-detected QGs, including photometric QGs without optical spectroscopy (gray points), compared to the local QG relation for total X-ray luminosity (0.3–10 keV) from LMXBs (solid line; Gilfanov 2004). We also *k*-correct the local QG relationship to $z \sim 1.2$ (dashed line) assuming a spectral slope as described in the text. Right: same relation for the 7 Ms soft-band-detected QGs. The X-ray luminosity from the $z \sim 1.2$ QG sample is larger than that emitted purely by LMXBs in local QGs at a given stellar mass, plausibly suggesting additional sources of X-ray emission.

obscuration, orientation, spectral slope, power, and selection effects of the AGN population. Our X-ray–detected sample has significantly lower luminosities than the majority of AGNs at comparable redshifts, suggesting a different origin for the X-ray emission (with the exception of the brightest QG, whose high X-ray luminosity and extreme hardness ratio suggest that it may host an obscured AGN).

Figure 7 also shows a plot of the X-ray luminosity versus the [O II] luminosity (right panel). We do not find a correlation between X-ray and [O II] luminosities, which, if present, may have suggested an AGN origin for both (however, we note that the sample has a limited dynamic range in [O II] luminosity). These arguments, taken together with the fact that we observe stacked X-ray emission in the soft band but not the hard band (i.e., on average, our galaxies are soft X-ray sources), suggest that the X-ray emission is unlikely to be powered by an AGN.

Alternatively, the X-ray luminosity of our QGs could be primarily powered by emission from hot gas (e.g., coronal gas that formed as a combination of outflows, stellar winds, or gravitational heating) and/or low-mass X-ray binaries (LMXBs). To investigate the contribution from hot gas and LMXBs, we plot the relationship of both the hard and the soft X-ray luminosities with stellar mass for our X-ray–detected samples in Figure 8. At $z \sim 0$, it is well established that X-ray luminosity from LMXBs scales linearly with stellar mass (e.g., Colbert et al. 2004; Gilfanov 2004; Kim & Fabbiano 2004). We find that our spectroscopic sample (colored points) does not show any correlation between X-ray luminosity and stellar mass, as would be expected if the X-ray luminosity was primarily from LMXBs. Since our X-ray–detected sample is small, we additionally plot photometrically selected QGs from the parent sample selected in Cassata et al. (2013) at $1 < z < 1.5$ that are X-ray detected in the *Chandra* 7 Ms imaging (i.e., those that were excluded from this study due to a lack of optical spectroscopy). We limit the photometric QG sample to this narrow redshift range to mitigate the uncertain effects of comparing the luminosities without the *k*-correction, which becomes increasingly large as the redshift range within the sample increases. The X-ray luminosities we observe in our QG sample (both hard and soft) are well in excess of those seen from LMXBs in local QGs (solid line in Figure 8; Gilfanov 2004). The local scaling relation appears an order of magnitude lower than our data, despite the luminosity being the total sum of the photons with energies spanning both the hard and the soft bands (0.3–10 keV; Gilfanov 2004). We *k*-correct the local QG relation to the observed frame of the QG sample (with mean redshift $z \sim 1.2$) assuming a conservative (steep) but typical power-law spectral shape defined by a photon index of $\Gamma \sim 1.8$, typical for low-redshift LMXBs (Lehmer et al. 2007; Boroson et al. 2011). The *k*-correction increases the expected luminosity at $z \sim 1.2$; however, it is still well below our observations.

Observationally, it is unknown whether or not the X-ray luminosity from LMXBs evolves with redshift; however, theoretical scaling relations indicate the possibility that LMXB luminosity may increase with redshift at a given stellar mass based on metallicity and star formation history evolution in the universe (Fragos et al. 2013). We are unable to place further constraints on the contribution of LMXBs. We conclude by simply noting the following: (1) our QGs are an order of magnitude more luminous in the X-ray than the LMXB contribution in local QGs of the same mass; however, (2) their observed X-ray luminosities are comparable to the total emission of local massive QGs, which includes the luminosity emitted from hot gas (e.g., Kim & Fabbiano 2013; Goulding et al. 2016). Locally, it is known that the contribution from ∼1 keV gas dominates the soft-band emission (Matsumoto et al. 1997; Gilfanov 2004; Sivakoff et al. 2004), which is also consistent with the luminosity seen on average in the stacked $z \sim 1.2$ QGs.

This leaves open the likely possibility that some X-ray emission comes from hot gas in these QGs that is heated by either gravitational or feedback processes. However, the dominant X-ray source in these $z \sim 1.2$ QGs remains unknown. An in-depth study of the X-ray emission of our sample and the comparison with local counterparts is beyond the scope of this paper, which is devoted to the galaxies' stellar age. We simply note the possibility that we are observing hot gas emission from individual QGs at $z \sim 1.2$, an observation that is very important for understanding the evolution of the ISM and circumgalactic medium in early-type galaxies. In a forthcoming

paper, we plan to accurately quantify selection effects and further investigate the X-ray properties of QGs. If confirmed, it will provide a powerful diagnostic of the feedback that took place during the star formation phase (e.g., van de Voort et al. 2016) and that is preventing the galaxies from forming stars again. We also note that, even if the observed X-ray emission does indeed turn out to be from hot gas, this still does not rule out AGNs as the cause of quenching, since the AGN activity could have been quenched together with the star formation (e.g., as a result of the cessation of gas accretion into the galaxies). The question of whether or not AGNs are responsible for star formation quenching in galaxies thus remains an open one. However, if hot gas is indeed a major component of the X-ray emission, this suggests that the quenching may be driven more by heating the gas than expelling it.

5. Summary

We present analysis based on optical spectroscopy of two samples of $z \sim 1.2$ QGs: those that are compact relative to local early-type galaxies and larger, more extended "normal" QGs. We find evidence that compact QGs have, on average, older stellar ages than normal-sized QGs. This observed density-age correlation is an empirical prediction of the progenitor bias scenario, whereby the first galaxies to evolve and quench are the most compact simply due to their early formation time, and extended QGs form later and quench later. We also study the [O II] emission and X-ray properties of the two samples, finding evidence for a lower incidence rate of residual energy sources in the compact sample compared to normal QGs. This is consistent with the idea that the compact sample quenched earlier, having had a longer time to fade. We do not find explicit evidence for compactness-driven quenching, and we suggest that future studies of galaxies at their quenching epoch may help illuminate whether the feedback or regulation of star formation differs with stellar density.

We thank the anonymous referee whose valuable suggestions have improved the paper significantly. C.C.W. acknowledges support from the *JWST*/NIRCam contract to the University of Arizona, NAS5-02015. R.B. gratefully acknowledges support by NASA through Hubble Fellowship grants #HF-51318 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. N.C. acknowledges Yale University's YCAA Prize Postdoctoral fellowship. P.C. acknowledges support from CONICYT through the project FONDECYT regular 1150216. T.L. thanks support from the National Science Foundation of China No. 11403021. This work is based on observations taken by the CANDELS Multi-cycle Treasury Program with the NASA/ESA *HST*. This research made use of Astropy, a community-developed core Python package for astronomy (Astropy Collaboration et al. 2013), and CSTACK (http://lambic.astrosen.unam.mx/cstack/), developed by Takamitsu Miyaji.

References

Abramson, L. E., & Morishita, T. 2016, ApJL, submitted (arXiv:1608.07577)
Acquaviva, V., Gawiser, E., & Guaita, L. 2011, ApJ, 737, 47
Alexandroff, R. M., Heckman, T. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 810, 104
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Balestra, I., Mainieri, V., Popesso, P., et al. 2010, A&A, 512, A12
Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, ApJ, 527, 54
Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, ApJ, 765, 104
Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, ApJ, 753, 167
Belli, S., Newman, A. B., & Ellis, R. S. 2015, ApJ, 799, 206
Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Boroson, B., Kim, D.-W., & Fabbiano, G. 2011, ApJ, 729, 12
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Bundy, K., Ellis, R. S., Conselice, C. J., et al. 2006, ApJ, 651, 120
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Cappelluti, N., Comastri, A., Fontana, A., et al. 2016, ApJ, 823, 95
Carollo, C. M., Bschorr, T. J., Renzini, A., et al. 2013, ApJ, 773, 112
Cassata, P., Giavalisco, M., Guo, Y., et al. 2011, ApJ, 743, 96
Cassata, P., Giavalisco, M., Williams, C. C., et al. 2013, ApJ, 775, 106
Ceverino, D., Dekel, A., Tweed, D., & Primack, J. 2015, MNRAS, 447, 3291
Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A, 482, 21
Ciotti, L., & Ostriker, J. P. 1997, ApJL, 487, L105
Colbert, E. J. M., Heckman, T. M., Ptak, A. F., Strickland, D. K., & Weaver, K. A. 2004, ApJ, 602, 231
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., & Burkert, A. 2014, MNRAS, 438, 1870
Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785
Diamond-Stanic, A. M., Moustakas, J., Tremonti, C. A., et al. 2012, ApJL, 755, L26
Dressler, A., Oemler, A., Jr., Poggianti, B. M., et al. 2004, ApJ, 617, 867
Fabbiano, G., Kim, D.-W., & Trinchieri, G. 1992, ApJS, 80, 531
Fabian, A. C. 2012, ARA&A, 50, 455
Fagioli, M., Carollo, C. M., Renzini, A., et al. 2016, ApJ, 831, 173
Fontana, A., Santini, P., Grazian, A., et al. 2009, A&A, 501, 15
Fragos, T., Lehmer, B., Tremmel, M., et al. 2013, ApJ, 764, 41
Franx, M., van Dokkum, P. G., Schreiber, N. M. F., et al. 2008, ApJ, 688, 770
Gallazzi, A., Bell, E. F., Zibetti, S., Brinchmann, J., & Kelson, D. D. 2014, ApJ, 788, 72
Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJL, 600, L93
Gilfanov, M. 2004, MNRAS, 349, 146
Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, A&AS, 141, 371
Gobat, R., Strazzullo, V., Daddi, E., et al. 2012, ApJL, 759, L44
Goulding, A. D., Greene, J. E., Ma, C.-P., et al. 2016, ApJ, 826, 167
Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, ApJ, 600, 580
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24
Guo, Y., Giavalisco, M., Cassata, P., et al. 2012, ApJ, 749, 149
Hasinger, G. 2008, A&A, 490, 905
Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, Natur, 428, 625
Heckman, T. M. 1980, A&A, 87, 152
Heckman, T. M., & Borthakur, S. 2016, ApJ, 822, 9
Hopkins, P. F., Murray, N., Quataert, E., & Thompson, T. A. 2010, MNRAS, 401, L19
Jacoby, G. H., Hunter, D. A., & Christian, C. A. 1984, ApJS, 56, 257
Johansson, P. H., Naab, T., & Ostriker, J. P. 2012, ApJ, 754, 115
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003a, MNRAS, 341, 33
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003b, MNRAS, 341, 54
Keating, S. K., Abraham, R. G., Schiavon, R., et al. 2015, ApJ, 798, 26
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kewley, L. J., Geller, M. J., & Jansen, R. A. 2004, AJ, 127, 2002
Kim, D.-W., & Fabbiano, G. 2004, ApJ, 611, 846
Kim, D.-W., & Fabbiano, G. 2013, ApJ, 776, 116
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kurk, J., Cimatti, A., Daddi, E., et al. 2009, Msngr, 135, 40
Kurk, J., Cimatti, A., Daddi, E., et al. 2013, A&A, 549, A63
Le Borgne, J.-F., Bruzual, G., Pelló, R., et al. 2003, A&A, 402, 433
Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al. 2007, ApJ, 657, 681
Lilly, S. J., & Carollo, C. M. 2016, ApJ, 833, 1
López-Sanjuan, C., Le Fèvre, O., Ilbert, O., et al. 2012, A&A, 548, A7
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2
Mainieri, V., Bergeron, J., Hasinger, G., et al. 2002, A&A, 393, 425

Matsumoto, H., Koyama, K., Awaki, H., et al. 1997, ApJ, 482, 133
Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Nelson, E., van Dokkum, P., Franx, M., et al. 2014, Natur, 513, 394
Omand, C. M. B., Balogh, M. L., & Poggianti, B. M. 2014, MNRAS, 440, 843
Patel, S. G., van Dokkum, P. G., Franx, M., et al. 2013, ApJ, 766, 15
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193
Pickles, A. J. 1998, PASP, 110, 863
Poggianti, B. M., Calvi, R., Bindoni, D., et al. 2013, ApJ, 762, 77
Popesso, P., Dickinson, M., Nonino, M., et al. 2009, A&A, 494, 443
Prochaska, L. C., Rose, J. A., Caldwell, N., et al. 2007, AJ, 134, 321
Renzini, A. 2006, ARA&A, 44, 141
Renzini, A., Ciotti, L., D'Ercole, A., & Pellegrini, S. 1993, ApJ, 419, 52
Salpeter, E. E. 1955, ApJ, 121, 161
Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703
Saracco, P., Longhetti, M., & Andreon, S. 2009, MNRAS, 392, 718
Saracco, P., Longhetti, M., & Gargiulo, A. 2011, MNRAS, 412, 2707
Schiavon, R. P., Faber, S. M., Castilho, B. V., & Rose, J. A. 2002, ApJ, 580, 850
Sell, P. H., Tremonti, C. A., Hickox, R. C., et al. 2014, MNRAS, 441, 3417
Shen, S., Mo, H. J., White, S. D. M., et al. 2003, MNRAS, 343, 978
Singh, R., van de Ven, G., Jahnke, K., et al. 2013, A&A, 558, A43
Sivakoff, G. R., Sarazin, C. L., & Carlin, J. L. 2004, ApJ, 617, 262
Spilker, J. S., Bezanson, R., Marrone, D. P., et al. 2016, ApJ, 832, 19
Stefanon, M., Marchesini, D., Rudnick, G. H., Brammer, G. B., & Whitaker, K. E. 2013, ApJ, 768, 92
Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, ApJL, 783, L14
Strateva, I., Ivezić, Ž, Knapp, G. R., et al. 2001, AJ, 122, 1861
Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, ApJS, 155, 271
Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., et al. 2015, ApJL, 809, L7
Teimoorinia, H., Bluck, A. F. L., & Ellison, S. L. 2016, MNRAS, 457, 2086
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Toft, S., van Dokkum, P., Franx, M., et al. 2007, ApJ, 671, 285
Treu, T., Stiavelli, M., Casertano, S., Møller, P., & Bertin, G. 2002, ApJL, 564, L13
Trujillo, I., Ferreras, I., & de La Rosa, I. G. 2011, MNRAS, 415, 3903
Trujillo, I., Feulner, G., Goranova, Y., et al. 2006, MNRAS, 373, L36
Valentinuzzi, T., Fritz, J., Poggianti, B. M., et al. 2010a, ApJ, 712, 226
Valentinuzzi, T., Poggianti, B. M., Saglia, R. P., et al. 2010b, ApJL, 721, L19
van de Voort, F., Quataert, E., Hopkins, P. F., et al. 2016, MNRAS, 463, 4533
van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van der Wel, A., Holden, B. P., Zirm, A. W., et al. 2008, ApJ, 688, 48
van Dokkum, P. G., & Ellis, R. S. 2003, ApJL, 592, L53
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 709, 1018
Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, A&A, 434, 53
Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, A&A, 454, 423
Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, A&A, 478, 83
Vazdekis, A. 1999, ApJ, 513, 224
Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, MNRAS, 404, 1639
Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, MNRAS, 449, 361
Wellons, S., Torrey, P., Ma, C.-P., et al. 2016, MNRAS, 456, 1030
Whitaker, K. E., Bezanson, R., van Dokkum, P. G., et al. 2016, arXiv:1607.03107
Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, ApJ, 745, 179
Wilkes, B. J., Kuraszkiewicz, J., Haas, M., et al. 2013, ApJ, 773, 15
Williams, C. C., Giavalisco, M., Cassata, P., et al. 2014, ApJ, 780, 1
Williams, C. C., Giavalisco, M., Lee, B., et al. 2015, ApJ, 800, 21
Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, ApJ, 713, 738
Worthey, G., & Ottaviani, D. L. 1997, ApJS, 111, 377
Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, ApJS, 195, 10
Yan, R., & Blanton, M. R. 2012, ApJ, 747, 61
Yano, M., Kriek, M., van der Wel, A., & Whitaker, K. E. 2016, ApJL, 817, L21
Zirm, A. W., van der Wel, A., Franx, M., et al. 2007, ApJ, 656, 66
Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327