Resolving Quiescent Galaxies at $z \gtrsim 2$. I. Search for Gravitationally Lensed Sources and Characterization of Their Structure, Stellar Populations, and Line Emission

Andrew B. Newman[1] , Sirio Belli[2] , Richard S. Ellis[3,4] , and Shannon G. Patel[1]

[1] The Observatories of the Carnegie Institution for Science, Pasadena, CA, USA; anewman@carnegiescience.edu
[2] Max-Planck-Institut für Extraterrestrische Physik (MPE), Giessenbachstr. 1, D-85748 Garching, Germany
[3] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[4] European Southern Observatory (ESO), Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany

Abstract

Quiescent galaxies at $z \gtrsim 2$ are compact and have weak or absent emission lines, making it difficult to spatially resolve their kinematics and stellar populations using ground-based spectroscopy. Gravitationally lensed examples provide a promising route forward, but such objects are very rare. We describe a search in the fields of 232 galaxy clusters that has uncovered five bright ($H_{AB} < 20$) lensed galaxies with red near-infrared colors. These include MRG-M0138, which is the brightest lensed galaxy known in the near-infrared. Analysis of near-infrared spectra and multiband photometry confirms that all are quiescent galaxies at $z = 1.95$–2.64 with stellar ages of 0.5–1.4 Gyr (corresponding to formation epochs $z_{form} \simeq 3$–4) and stellar masses of $10^{11.6 - 12.8} \mu^{-1} M_\odot$, where μ is the magnification. In three cases, we derive lens models and reconstruct the source structure; these galaxies are massive ($M_* \gtrsim 10^{11.0} M_\odot$) and follow the mass–size relation defined by unlensed samples. In two of these three galaxies, the main structural component is an inclined disk. Weak emission lines are detected in four of five galaxies with high ratios [N II]/H$\alpha \simeq 2$–6 that are inconsistent with a star formation origin. Based on the line ratios, the Hα equivalent widths, and the distribution and kinematics of the gas, we infer that shocks are likely to be present in at least two galaxies and could be present in all of the line emitters. We speculate that these could be analogs of local galaxies in which AGN jet-driven outflows are thought to heat the interstellar medium and suppress star formation. In further papers, we will present spatially resolved measurements of the stellar populations and kinematics of this unique sample.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – gravitational lensing: strong

1. Introduction

Stellar archaeological studies (e.g., Thomas et al. 2010) and the evolution of the fundamental plane (e.g., Kelson et al. 1997; Treu et al. 2005; van der Wel et al. 2005) indicate that the stars in the most massive galaxies were formed at $z \gtrsim 2$. With the advent of deep, wide near-infrared imaging surveys, the emergence of quiescent galaxies is now charted directly. Quiescent galaxies constitute a 10%–20% minority of the population at $z \gtrsim 3$, even among the highest-mass galaxies, but they become the majority of $\gtrsim 10^{11} M_\odot$ galaxies by $z \simeq 1.5$–2 (Muzzin et al. 2013b).

Although the stellar populations of these early quiescent galaxies may have evolved more-or-less passively after quenching, their structures have evolved dramatically in the intervening 10 Gyr. Numerous studies have shown that the typical size of $z \sim 2$ quiescent galaxies is 3–5× smaller than local counterparts of the same stellar mass (e.g., Trujillo et al. 2006; Buitrago et al. 2008; van Dokkum et al. 2008; Toft et al. 2009; Damjanov et al. 2011; Newman et al. 2012). Importantly, the number density of the most compact quiescent galaxies has been declining since $z \sim 1.5$ (van der Wel et al. 2014), showing that part of this remarkable evolution must arise from the continued growth of massive galaxies after the cessation of star formation (Belli et al. 2015).

Observations of quiescent galaxies at $z \gtrsim 2$ are still rather crude and have largely been confined to bulk properties: number densities, sizes, colors, Sérsic indices, and shapes. Spectroscopy of the more massive examples has been enabled by near-infrared (NIR) spectrographs on large telescopes (Keck/MOSFIRE, *Magellan*/FIRE, VLT/X-Shooter) and with grisms on the *Hubble Space Telescope* (*HST*; e.g., Bedregal et al. 2013; Whitaker et al. 2013; Krogager et al. 2014; Newman et al. 2014; Fumagalli et al. 2016; Lee-Brown et al. 2017). These data have been used to measure velocity dispersions, stellar ages, and chemical abundances. These observations are still very demanding due to the faintness of the stellar continuum compared to the NIR background. Within the literature, we find only 10 distinct quiescent galaxies beyond $z = 2$ that have been observed with the spectral resolution and depth needed to measure stellar kinematics (Kriek et al. 2009, 2016; Toft et al. 2012; van de Sande et al. 2013; Newman et al. 2015a; Hill et al. 2016; Belli et al. 2017b; Toft et al. 2017).

More detailed information is needed to address open questions regarding the formation and evolution of the quiescent population. For example, cosmological simulations predict multiple paths to form a compact quiescent galaxy at $z \sim 2$. In these simulations, some compact galaxies formed very early, when the universe was more dense, and remained compact. Others were once more extended and then "shrank" in half-light radius due to centrally concentrated star formation (Wellons et al. 2015; Zolotov et al. 2015). The relative importance of these scenarios could be constrained by spatially resolving the star formation histories in a sample of objects. Another example concerns morphology and dynamics. Observations have indicated a rise in the proportion of flattened quiescent galaxies toward higher redshifts, leading to the inference that quiescent galaxies are more disk-like at early epochs (van der Wel et al. 2011; Chang et al. 2013). Measurements of rotation are needed to observe this directly,

but this requires the stellar kinematics to be spatially resolved. Such observations could also help to link recently quenched galaxies to star-forming progenitor populations, whose gas distribution and kinematics are now being measured in the ionized (Barro et al. 2014; Nelson et al. 2014; van Dokkum et al. 2015) and molecular phases (Barro et al. 2016, 2017; Tadaki et al. 2017).

These examples motivate the need for *spatially resolved* spectroscopy of the stellar continuum for quiescent galaxies at $z \gtrsim 2$. Unfortunately, their small angular sizes (half-light radii $R_e \sim 0\rlap{.}''2$) make this impractical in ground-based seeing. Observations with the *HST* grisms or ground-based adaptive optics instruments have the necessary angular resolution but lack the necessary spectral resolution in the former case and sensitivity in the latter. The best way to resolve the stellar continuum of high-redshift quiescent galaxies using current facilities is to locate gravitationally lensed examples.

Numerous lensed *star-forming* galaxies have been identified and used to spatially resolve the distribution of star formation, gas kinematics, and metals (e.g., Stark et al. 2008; Swinbank et al. 2009; Jones et al. 2010; Leethochawalit et al. 2016a, 2016b). These lensed sources are optically bright and have been identified through either targeted imaging of massive clusters or panoramic optical imaging surveys (Allam et al. 2007; Cabanac et al. 2007; Hennawi et al. 2008; Stark et al. 2013).

The study of lensed *quiescent* galaxies at high redshifts, on the other hand, is a recent development. Oldham et al. (2017) located 14 lensed early-type galaxies at $z \sim 0.6$ whose spectrum was blended with that of the lens galaxy in the Sloan Digital Sky Survey (SDSS). At higher redshifts, only a few examples have been discovered. Geier et al. (2013) identified two quiescent galaxies magnified by foreground clusters: A1413-1 at $z = 1.71$ and M2129-1 at $z = 2.15$. They used the lensing magnification to study the source structure and spectral energy distribution. Hill et al. (2016) studied the intermediate-mass quiescent galaxy COSMOS 0050+4901, which is quadruply imaged by a foreground galaxy (Muzzin et al. 2012). The flux amplification enabled a measurement of velocity dispersion at a remarkably high redshift $z = 2.76$. Ebeling et al. (2018) recently published the discovery of a very highly magnified quiescent galaxy at $z = 1.6$. However, none of these studies spatially resolved the spectrum in the analysis, partly because two of these sources (A1413-1 and COSMOS 0050+4901) are still rather faint ($H_{AB} \gtrsim 21$) and compact, even when magnified.

The paucity of known lensed quiescent galaxies, especially at high redshifts, arises from their relatively low density on the sky combined with their optical faintness. Locating lensed quiescent galaxies at $z \gtrsim 2$ requires NIR imaging (since the Balmer/4000 Å break is redshifted beyond 1.2 μm) that covers a large source plane area.

Motivated by the utility of a such a sample, we embarked on an NIR imaging survey that was designed to locate particularly bright examples that are extended enough to be spatially resolved from the ground during good seeing conditions. In this paper, we describe a search in the fields of 232 massive galaxy clusters using both archival *HST* images and a new imaging survey with the FourStar camera (Persson et al. 2013) at the Magellan Baade telescope. We have located five magnified galaxies that are exceptionally bright ($H_{AB} < 20$) and have colors consistent with $z \gtrsim 2$ quiescent galaxies. These include what we believe to be the NIR-brightest giant arc known ($H_{AB} = 16.5$). Analysis of their rest-frame optical spectra and ultraviolet-to-NIR photometry confirms that these are quiescent galaxies at $z = 1.95$–2.64. Prior to the survey, only one of these five galaxies was known (M2129-1; Geier et al. 2013). In four cases, we are able to spatially resolve the stellar continuum in ground-based NIR spectra. These objects are rare and valuable resources that we will use to investigate the spatially resolved star formation histories and stellar kinematics, which is not currently possible for any other sample.

In Newman et al. (2015a), we presented a pilot study of one galaxy in our sample (RGM0150, named MRG-M0150 in the scheme used in this paper). This was the first galaxy beyond $z \simeq 1.1$ for which spatially resolved stellar kinematics were measured. We showed that MRG-M0150 rotates rapidly compared to its likely descendants and, therefore, must "spin down" between $z = 2.64$ and the present. Toft et al. (2017) showed that the lensed galaxy M2129-1 is a massive quiescent system at $z = 2.15$ that is a nearly pure disk and is also rotating surprisingly rapidly. In this and subsequent papers in the series, we will present new spectroscopic observations of this galaxy (MRG-M2129 in our naming scheme) and we compare our measurements to those of Toft et al. 2017).

We begin this paper by presenting results from the survey. In Section 2, we introduce the imaging program used to locate the five lensed quiescent galaxies. Sections 3 and 4 describe the follow-up imaging and spectroscopic observations. In Section 5, we derive lens models for these three systems and then use them to reconstruct the sources. In Section 6, we analyze the unresolved spectra and photometry of the lensed galaxies to establish their bulk properties, including redshifts, stellar ages, and quiescence. In Section 7, we demonstrate the nearly ubiquitous presence of low-level line emission in our sample and discuss its possible origins. In Section 8, we discuss the representative nature of our lensed galaxy sample, the implications of their structures and emission line properties for their evolutionary histories, and the utility of the sample for future spatially resolved studies.

In the companion Paper II, we measure the resolved stellar kinematics of four galaxies in the sample. Further papers will discuss the resolved stellar ages and chemical abundances.

Throughout, we refer to magnitudes on the AB system and assume a flat ΛCDM cosmology with $\Omega_m = 0.3$ and $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$.

2. Search for Gravitationally Lensed Quiescent Galaxies at $z \gtrsim 2$

We now describe our search for lensed quiescent galaxies at $z \gtrsim 2$ in the fields of 232 massive galaxy clusters. After reviewing the motivation to image a large number of clusters, we derive an efficient color selection, outline its application to archival *HST* data and new *Magellan*/FourStar imaging, and then present our yield of five lensed quiescent galaxies.

2.1. Basic Considerations

Massive quiescent galaxies at high redshifts are relatively rare. The probability to find such a galaxy magnified by a given cluster is the product of the population's surface density and the source plane area magnified by the cluster above some minimum magnification of interest. The typical value of this area depends on the cluster sample, so only rough estimates are possible. An upper limit can be estimated from the Hubble Frontier Fields, which were chosen to be among the most
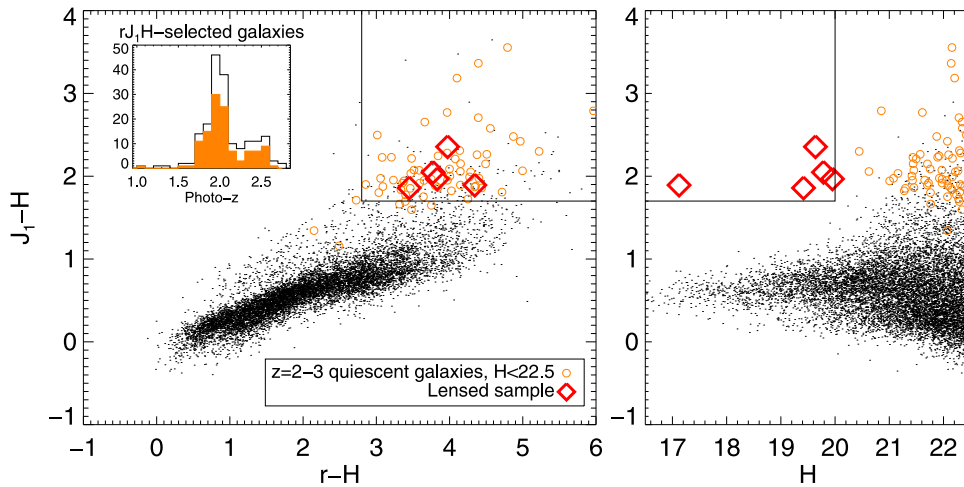
Figure 1. Color–color (left panel) and color–magnitude (right panel) diagrams showing all $H < 22.5$ galaxies selected from the NMBS/COSMOS catalog as black points. The targeted subpopulation, consisting of UVJ-quiescent galaxies with $2 < z_{phot} < 3$, are shown as orange circles. Black lines enclose the selection boxes used for our lensed galaxy search. The inset in the left panel shows the photometric redshift distribution of NMBS galaxies falling in the color selection box, with the UVJ-quiescent subset filled. In both panels, the five lensed galaxies identified in our search (two from the FourStar and three from the *HST* components) are plotted as red diamonds. The Subaru/SuprimeCam r^+ filter was used for the NMBS galaxies. For the lensed galaxies, which were observed with various filters, we integrated the fitted SED to synthesize r^+J_lH magnitudes.

powerful cluster lenses. For these clusters, Johnson et al. (2014) derive a typical source plane area of $\simeq 0.3$ arcmin^2 magnified by $\mu > \mu_{min} = 3$. This area approximately scales as μ_{min}^{-2}. These figures enable a rough estimate of the frequency of lensed sources magnified above some limiting flux. We consider a source population of quiescent galaxies at $z = 1.8$–3, whose magnitude distribution we estimated using the NEWFIRM Medium Band Survey (NMBS) catalogs (Whitaker et al. 2011). Here quiescence is defined using the UVJ criterion in the form presented by Whitaker et al. We find a probability of 0.02 per cluster to find such a galaxy magnified to $H < 20$. As we will show, this flux limit is of interest because it is brighter than the unlensed examples that are found in the widest extragalactic deep fields. Although the estimated probability is uncertain and should be considered to be an upper limit, it shows that one must search hundreds of clusters to find a few very bright, lensed quiescent galaxies at these redshifts.

2.2. Color Selection

Although they are rare, massive quiescent galaxies at $z \simeq 2$–3 can be efficiently identified with two or three filter imaging, and their high NIR surface brightnesses, imply that relatively shallow exposures are sufficient. This makes it feasible to search large numbers of clusters.

Quiescent galaxies at high redshifts can be identified on the basis of their red colors at NIR wavelengths, which are produced by the redshifted Balmer/4000 Å break (Franx et al. 2003). The addition of an optical magnitude probing the rest-frame UV further helps to eliminate dusty star-forming galaxies (Daddi et al. 2004). We have applied this basic strategy to two imaging surveys of massive clusters: (1) a dedicated campaign with the FourStar camera at the Magellan Baade telescope, and (2) a search of archival data from Wide Field Camera 3 (WFC3) onboard *HST*. These surveys will be described in Sections 2.3 and 2.4.

Using the NMBS catalogs, we experimented with color cuts that efficiently identify a target population of UVJ-quiescent galaxies at $z = 2$–3 with $H < 22.5$, which corresponds to the

desired lensed magnitude limit of $H < 20$ for a maximum expected magnification of $\mu = 10$. (The precise limit is not relevant since the colors do not vary drastically with magnitude.) The goal was to balance a high completeness for selecting this target population (orange circles in Figure 1) with minimal contamination from star-forming or lower-redshift galaxies (black points). The FourStar filter set includes the broadband J, H, and K_s filters and the medium-band J_1, J_2, J_3, H_s, and H_l filters (Persson et al. 2013). We found that the combination $J_1 - H > 1.7$ is nearly optimal, as shown in Figure 1. A second cut of $r - H > 2.8$ is also plotted. For this filter combination, the optical-IR cut removes only a few galaxies, so it is not necessary to identify bright quiescent candidates in our FourStar imaging: $J_1 - H$ is sufficient. For our *HST* archival search, we adapted these color cuts based on the available filters, as described in Section 2.4. The optical-IR cut plays a greater role for some of these filter combinations.

The color cuts $J_1 - H > 1.7$ and $r - H > 2.8$ select 85% of the target population, so our selection is reasonably complete. Furthermore, 31% of the color-selected galaxies are quiescent galaxies at $z > 2$, and of the remainder, an additional 33% are quiescent galaxies at slightly lower redshifts $z = 1.6$–2, which are still of interest. This level of purity makes it possible to pursue efficient spectroscopic follow-up to confirm the redshifts and quiescent nature of the sources. As outlined later, we have examined the fields of 232 clusters and located five color-selected galaxies that are magnified above the $H < 20$ flux limit. Fainter color-selected galaxies were also identified, but in this paper we confine ourselves to $H < 20$ sample, for which our spectroscopic follow-up is complete.

2.3. Magellan/FourStar Search

We imaged 131 clusters through the J_1 and H filters with FourStar over five semesters from 2014A to 2016A. Targets were drawn from several sources. Approximately 40% were X-ray–selected clusters from the Massive Cluster Survey (MACS; Ebeling et al. 2001; Repp & Ebeling 2018). Although these were included in the target lists of the *HST* snapshot programs described in Section 2.4, they had not yet been

observed with WFC3-IR. Another $\simeq 40\%$ were selected from the first *Planck* catalog of Sunyaev–Zel'dovich sources (Planck Collaboration et al. 2014). To maximize the lensing efficiency, we only considered those sources that had been confirmed as clusters at redshifts $0.3 < z < 0.8$, and we gave priority to those with higher signal-to-noise ratios in the *Planck* maps. Finally, $\simeq 20\%$ of our FourStar targets were optically selected clusters that were drawn primarily from the redMaPPer DR8 catalog (Rykoff et al. 2014) and prioritized by richness. A handful were drawn from Wong et al.'s (2013) catalog of fields suggested to be powerful lenses.

The typical observing sequence was to move to the target, correct the focus and mirror figure with the facility Shack–Hartmann system and obtain 5×5.8 s unguided exposures in the H band at each of 11 random dither positions within a $90'' \times 90''$ box, followed by 2×32 s exposures at 13 random dither positions through the J_1 filter. The total wall clock time per cluster was approximately 30 minutes.

The data were reduced using the automated system described by Kelson et al. (2014), which produces stacked images for each cluster and filter with astrometry tied to 2MASS. Photometric calibration was obtained using stars in the 2MASS point source catalog (Skrutskie et al. 2006). This is straightforward for the H and K_s filters, which are present in the 2MASS catalog. For the medium-band filters, we used the Pickles (1998) stellar library to derive mean relations that relate the 2MASS magnitudes of a star to its J_1, J_2, and J_3 magnitudes.

Applying the cuts $J_1 - H > 1.7$ and $H < 20$ to this sample of 131 clusters yielded five sources. Two of these sources are clearly magnified by the clusters PSZ1-G295.24–21.55 and MACSJ0138.0–2155. As the right panel of Figure 1 shows, these objects occupy a region of color–magnitude space that is virtually empty in extragalactic field surveys. The other three sources are compact and are located at least 4 arcmin from the cluster center. The FourStar field of view is much larger than the high-magnification region of a cluster, so these three sources are likely to be unmagnified systems on the very bright tail of the luminosity function. Their presence is not unexpected given that the FourStar survey encompasses 4.3 deg^{-2}; however, since they are not magnified, we will not discuss these sources further. The yield of $2/131 \simeq 0.02$ lensed quiescent galaxies per cluster is consistent with the upper limits roughly estimated in Section 2.1.

2.4. HST Archival Search

Many galaxy clusters have been observed with *HST*, but a much smaller subset has been observed with WFC3-IR through two filters and with ACS or WFC3-UVIS through at least one optical filter, as required to implement our color criteria. At the time of our search (2014–2016), the vast majority of such observations had been undertaken through one of three programs: (1) a snapshot imaging campaign based on the Massive Cluster Survey (MACS), led by PIH. Ebeling in Proposal IDs 10491, 10875, 12166, and 12884; (2) multiband imaging of lensing clusters identified in the Sloan Digital Sky Survey, led by PIM. Gladders in Proposal ID 13003, and (3) the Cluster Lensing and Supernova survey with Hubble (CLASH), a multi-cycle treasury program, led by P.I. M. Postman in Proposal ID 12065.

For all of the clusters included in these programs that, at the time of the analysis, had been observed through two WFC3-IR filters and one or more ACS or WFC3-UVIS filters, we produced multiband photometric catalogs and then applied color and magnitude criteria to search for lensed quiescent galaxies. Since the surveys used a variety of filters, we adapted the color criteria shown in Figure 1 for each survey. For all of the galaxies in the NMBS catalogs that were used to construct Figure 1, we integrated the spectral energy distribution fit to produce synthetic magnitudes in the relevant *HST* filters. We then reproduced Figure 1 with the appropriate filter combinations and adjusted the color thresholds to match the balance of completeness and contamination shown in the figure.

Within the MACS cluster sample, we processed 46 clusters with WFC3-IR and ACS imaging. The color criteria became (F110W $-$ F140W > 0.7) and (F814W $-$ F140W > 2 or F606W $-$ F140W > 2.5). We also imposed a flux limit of F140W < 20.7, which reflects our fiducial $H < 20$ cut adjusted by the mean F140W $- H$ color of the target population. We identified one lensed galaxy behind cluster MACSJ0150.3–1005.

For the Gladders sample, we processed 30 clusters that had been imaged with WFC3-IR and WFC3-UVIS. The color criteria became (F125W $-$ F160W > 0.7 or F110W $-$ F160W > 1.0 or F105W $-$ F160W > 1.4) and (F606W $-$ F160W > 3) and F160W < 20.3. We identified one source magnified by the cluster SDSSJ1522 $+$ 2535.

For the CLASH sample, 16 bands of imaging are available for 25 clusters. The color criteria became F110W $-$ F160W > 0.9 and F814W $-$ F160W > 2.5 and F160W < 20.3. We identified one source magnified by MACSJ2129.4–0741.

2.5. The Lensed Quiescent Galaxy Sample

The coordinates and basic properties of the five lensed color-selected galaxies that were located in our survey are listed in Table 1. We name these galaxies MRG-M0138, MRG-M0150, MRG-P0918, MRG-S1522, and MRG-M2129 based on the name of the lensing cluster, where the prefix MRG denotes Magnified Red Galaxy. Of these objects, MRG-M2129 was previously identified by Geier et al. (2013) and studied by Toft et al. (2017), while the other four were discovered in this survey. A pilot study of MRG-M0150 was presented by Newman et al. (2015a).

Figure 2 shows *HST* images of the five lensed galaxies. We will now briefly describe each of these galaxies, their lensing configurations, and the lensing clusters. We note that the distinctive colors and high surface brightnesses of the lensed quiescent galaxies makes it straightforward to identify multiple images, which are confirmed by lens modeling in Section 5.

MRG-M0138 is a remarkable system that presents five multiple images: two merging images forming a giant arc southward of the cluster center, a single image on the western side, a radial arc near the brightest cluster galaxy (BCG), and a central image (not visible in Figure 2). The brightest image (Image 1) has $H = 17.1$, which is astoundingly bright for this redshift: Figure 3 shows that this image is ~ 3 mag brighter than the brightest unlensed quiescent galaxies at similar redshifts found in the 1.6 deg^{-2} UltraVISTA field (Muzzin et al. 2013a)! Summing the two merging images 1 and 2, the total AB magnitudes of the giant arc are $H = 16.5$ and $K_s = 16.1$. We believe this is the brightest giant arc known at NIR wavelengths, at least among $z \gtrsim 1$ sources. It is 1.1 mag brighter at K_s than the giant arc discovered by Dahle et al. (2016) and 2.3 mag brighter than the source described by Wuyts et al. (2010). For the lens MACSJ0138.0-2155, we

Table 1
Sample of Lensed Quiescent Galaxies

Lensed Galaxy	R.A. (hr)	Decl. (deg)	z_{spec}	H (mag)	Lensing Cluster	z_{lens}	Source of z_{lens}
MRG-M0138	01:38:03.9	−21:55:49	1.95	17.1	MACSJ0138.0-2155	0.338	This paper
MRG-M0150	01:50:21.0	−10:05:14	2.64	19.6	MACSJ0150.3-1005	0.365	SDSS DR12 (Alam et al. 2015)
MRG-P0918	09:18:34.1	−81:03:08	2.36	19.4	PSZ1 G295.24-21.55	0.61	Planck Collaboration et al. (2014)
MRG-S1522	15:22:53.6	+25:35:49	2.45	19.8	SDSSJ1522 + 2535	0.58	SDSS DR12 (photometric)
MRG-M2129	21:29:22.3	−07:41:31	2.15	20.0	MACSJ2129.4-0741	0.589	Ebeling et al. (2007)

Note. In cases of multiple images, the coordinates and magnitudes listed are those of the brightest image.

measured a redshift of 0.338 based on a Magellan/LDSS-3 spectrum of the brightest cluster galaxy.

MRG-M0150 is also a multiple image system, presenting three multiple images in a "naked cusp" configuration analyzed by Newman et al. (2015a). MRG-P0918, MRG-S1522, and MRG-M2129 are all singly imaged. The images of MRG-P0918 and MRG-M2129 are both highly elongated, whereas that of MRG-S1522 is bright but only modestly elliptical. The clusters lensing MRG-P0918 and MRG-S1522 are not well-known and, as we will discuss below, we do not have the constraints needed to determine the magnification factor or reconstruct the source in these cases. MACSJ2129.4-0741, on the other hand, is one of the most X-ray-luminous clusters (Ebeling et al. 2007; $z = 0.589$) and is a well-studied lens with multiple public lens models.

Figure 3 compares the observed (lensed) magnitudes of this sample to unlensed quiescent galaxies in the UltraVISTA field (Muzzin et al. 2013a). All of the sources in our lensed sample are substantially brighter than even the brightest quiescent galaxies located at comparable redshifts, even within the largest deep fields. This makes them premier targets for absorption line spectroscopy in the near-infrared and enables a number of new observations, including their spatially resolved stellar kinematics (Newman et al. 2018, hereafter Paper II) and populations. In the following sections, we will describe follow-up observations to verify the quiescent nature and redshifts $z = 1.95$–2.64 of these sources (Sections 3 and 4) before deriving the lensing magnifications (Section 5) and the source intrinsic properties (Sections 6–7).

3. Imaging Data and Photometry

We observed our lensed galaxy sample using various ground- and spaced-based telescopes to characterize their spectral energy distributions from optical to near-infrared wavelengths. The data described in this section will be used to construct lens models and to study the stellar populations and structures of our sample in Sections 5–6.

3.1. HST Observations

A variety of archival data exists for the lensed galaxies that we located in our *HST* search. The sources are listed in Section 2.4. For MRG-M0150, we used archival WFC3-IR/F140W, WFC3-IR/F110W, ACS/F814W, and WFPC2/F606W images, as described in Newman et al. (2015a). In order to obtain deeper images with improved sub-pixel sampling and a filter combination that better straddles the Balmer break, we obtained additional WFC3-IR images through the F160W and F125W filters, each at half-orbit depth (Proposal ID 14205, PI: A. Newman). For MRG-S1522, we used WFC3-IR/F160W, WFC3-IR/F105W, and WFC3-UVIS/F606W archival images from the Gladders program (see

Section 2.4). For MRG-M2129, we used archival CLASH imaging through 13 ACS and WFC3 filters (see Table 2).

No archival data existed for the two lensed galaxies that were located in our FourStar survey. Therefore, for MRG-P0918 we obtained WFC3-IR/F160W and F105W images, each with half-orbit depth, and ACS/F814W and F555W images, each with one-orbit depth, through the aforementioned program. For MRG-M0138, observations were undertaken through a mid-cycle program (Proposal ID 14496, PI: A. Newman). Two orbits were split between WFC3-IR/F160W (1.6 ks) and F105W (3.6 ks) exposures, and two further orbits were devoted to an F555W exposure.

All of the new *HST* observations employed standard sub-pixel dither patterns and were reduced using the Multi-Drizzle package (Koekemoer et al. 2003). For the source reconstructions that we will perform in Section 5, we require an estimate of the WFC3 F160W point-spread function (PSF). For MRG-M0150 and MRG-M2129, we constructed PSFs from stars in the mosaic. Suitably bright and isolated stars were not present for MRG-M0138, so we instead generated a synthetic TinyTim (Krist et al. 2011) PSF and convolved it with a circular Gaussian to best match a (slightly broadened) star in the mosaic. As expected, the curves of growth for all three PSFs are very similar.

3.2. Spitzer/IRAC Observations

We analyzed archival images from the Infrared Array Camera (IRAC) onboard the *Spitzer* Space Telescope for three galaxies: MRG-P0918 was observed in Program ID 90233 by PI: C. Lawrence, MRG-S1522 was observed in Program ID 70154 by PI M. Gladders, and MRG-M2129 was observed in Program ID 90009 by PI M. Bradač. We obtained new IRAC images of MRG-M0138 via a Director's Discretionary Time program (Program ID 12127) and of MRG-M0150 via a joint *HST* program in cycle 23 (Program ID 12003, PI: A. Newman).

The IRAC images cover the 3.6 μm and 4.5 μm channels, with the exception of MRG-P0918. This galaxy falls near the edge of the 4.5 μm mosaic and suffers from various artifacts. Therefore, we excluded this image from our analysis. We performed photometry using the standard calibrated mosaics produced by the IRAC pipeline with a 0″.6 pixel scale.

3.3. Ground-based Observations

We imaged the entire sample through various filters using FourStar at the Magellan Baade telescope. For MRG-M0138, MRG-P0918, and MRG-S1522, we obtained images through the J_1, J_2, J_3, H, and K_s filters, using the medium bands to improve the sampling around the Balmer/4000 Å break. For MRG-M0150 and MRG-M2129, due to the
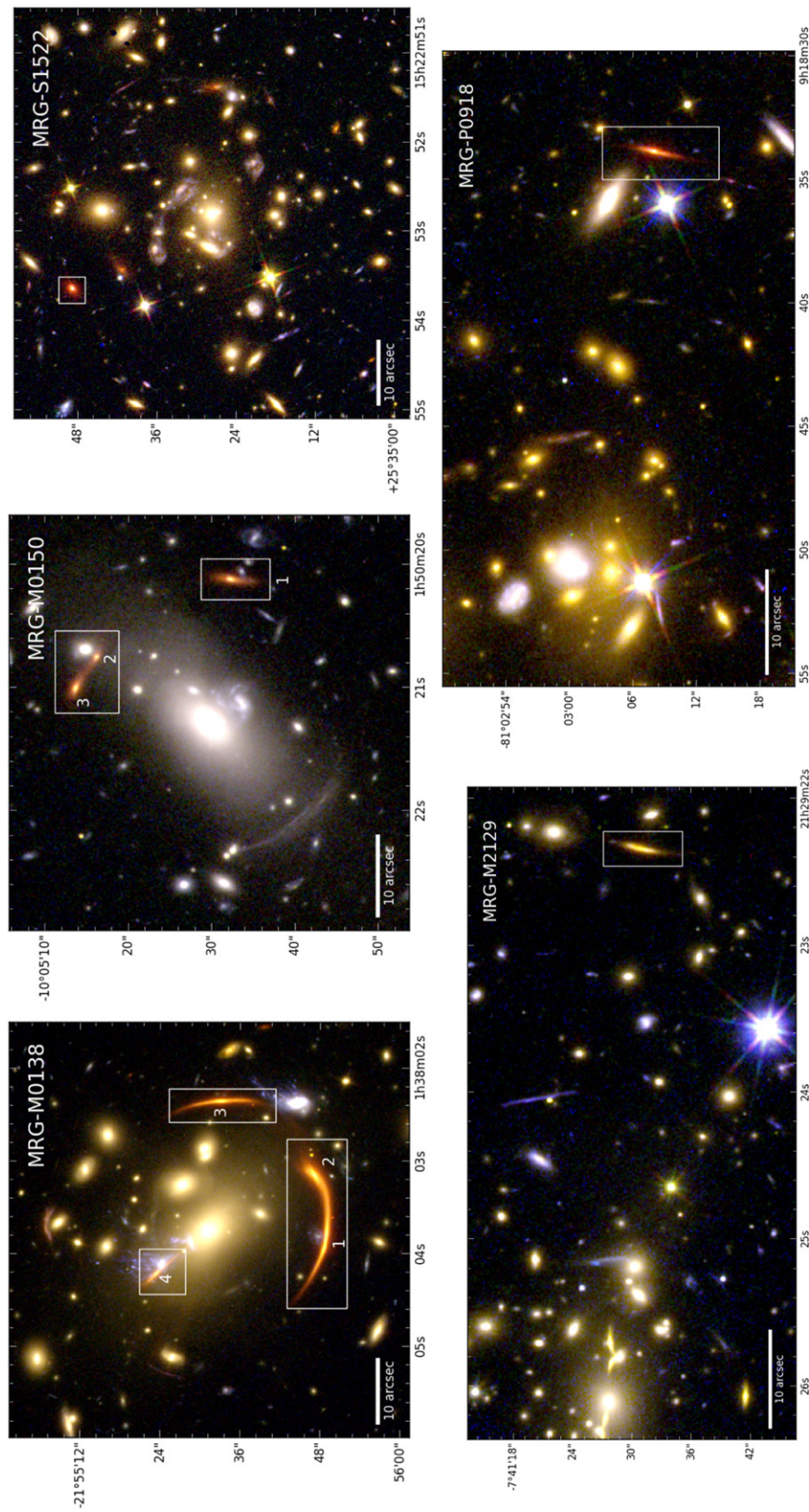
Figure 2. Composite *HST* images of the sample displayed with a logarithmic stretch. The lensed quiescent galaxies are identified by boxes. Multiple images are numbered in the cases of MRG-M0150 and MRG-M0138. (The central image of MRG-M0138 is not labeled.) Depending on the available filters, one of F555W, F606W, or F814W was used for the blue channel; F105W or F125W for the green; and F160W for the red.

greater number of *HST*/WFC3-IR observations, we obtained FourStar images in *J*, *H*, and K_s for the former and only in K_s for the latter.

At optical wavelengths, we observed MRG-M0138 through the *g*, *r*, and *z* filters using the LDSS-3 imaging spectrograph at the Magellan Clay telescope. We also

Table 2
Photometry

Instrument	Filter	AB mag.
	MRG-M0138	
LDSS	g	22.61 ± 0.19
ACS	F555W	22.26 ± 0.20
LDSS	r	21.49 ± 0.15
LDSS	i	20.80 ± 0.13
LDSS	z	19.64 ± 0.12
FourStar	J_1	19.06 ± 0.07
WFC3-IR	F105W	18.93 ± 0.04
FourStar	J_2	18.40 ± 0.07
FourStar	J_3	17.76 ± 0.06
WFC3-IR	F160W	17.28 ± 0.03
FourStar	H	17.08 ± 0.06
FourStar	K_s	16.67 ± 0.06
IRAC	Ch. 1	16.21 ± 0.05
IRAC	Ch. 2	16.00 ± 0.05
	MRG-M0150	
WFPC2	F606W	23.89 ± 0.16
ACS	F814W	22.87 ± 0.06
WFC3-IR	F110W	21.52 ± 0.04
FourStar	J	21.27 ± 0.08
WFC3-IR	F125W	21.17 ± 0.03
WFC3-IR	F140W	20.43 ± 0.03
WFC3-IR	F160W	19.90 ± 0.03
FourStar	H	19.63 ± 0.06
FourStar	K_s	19.16 ± 0.06
IRAC	Ch. 1	18.79 ± 0.07
IRAC	Ch. 2	18.33 ± 0.06
	MRG-P0918	
ACS	F555W	23.59 ± 0.14
ACS	F814W	22.36 ± 0.05
FourStar	J_1	21.22 ± 0.09
WFC3-IR	F105W	21.41 ± 0.04
FourStar	J_2	20.87 ± 0.08
FourStar	J_3	20.38 ± 0.08
WFC3-IR	F160W	19.54 ± 0.03
FourStar	H	19.36 ± 0.06
FourStar	K_s	19.17 ± 0.06
IRAC	Ch. 1	19.05 ± 0.07
	MRG-S1522	
WFC3-UVIS	F606W	23.91 ± 0.05
LDSS	z	22.65 ± 0.12
FourStar	J_1	21.80 ± 0.07
WFC3-IR	F105W	21.95 ± 0.04
FourStar	J_2	21.64 ± 0.08
FourStar	J_3	20.90 ± 0.09
WFC3-IR	F160W	19.96 ± 0.03
FourStar	H	19.75 ± 0.06
FourStar	K_s	19.41 ± 0.06
IRAC	Ch. 1	19.14 ± 0.06
IRAC	Ch. 2	18.99 ± 0.06
	MRG-M2129	
ACS	F435W	25.13 ± 0.74
ACS	F475W	24.53 ± 0.29
ACS	F555W	24.53 ± 0.21
ACS	F606W	24.16 ± 0.22
ACS	F625W	23.91 ± 0.23
ACS	F775W	23.22 ± 0.15
ACS	F814W	23.16 ± 0.08
ACS	F850LP	22.55 ± 0.13
WFC3-IR	F105W	21.90 ± 0.05
WFC3-IR	F110W	21.14 ± 0.04
WFC3-IR	F125W	20.78 ± 0.04
WFC3-IR	F140W	20.31 ± 0.03
WFC3-IR	F160W	20.06 ± 0.03
FourStar	K_s	19.59 ± 0.06

Table 2
(Continued)

Instrument	Filter	AB mag.
IRAC	Ch. 1	19.11 ± 0.05
IRAC	Ch. 2	19.02 ± 0.05

Note. Total fluxes for MRG-M0138 and MRG-M0150 are normalized to Image 1. The giant arc in MRG-M0138 (merging Images 1 and 2) is 0.5 mag brighter.
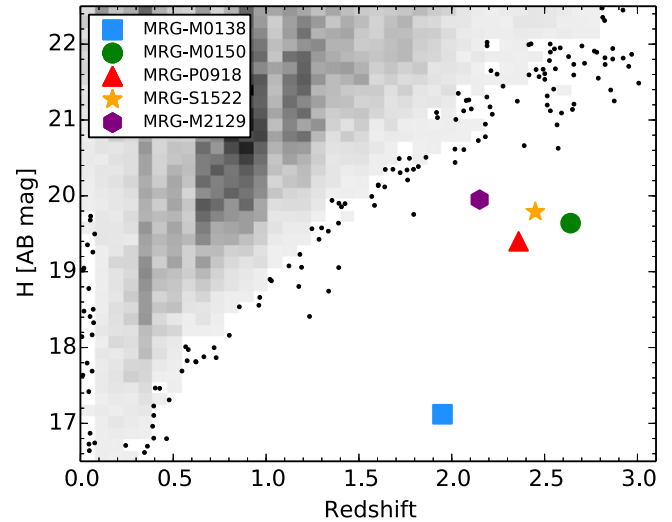
Figure 3. Magnitude distribution of *UVJ*-quiescent galaxies drawn from the UltraVISTA survey (Muzzin et al. 2013a) is indicated by the gray histogram, with individual black points shown where the density is low. The magnified red galaxy (MRG) sample is significantly brighter than the brightest unlensed galaxies found over the 1.6 deg^{-2} field.

observed MRG-S1522 through the *z* filter. Photometric calibration was obtained using images of stellar fields in the Sloan Digital Sky Survey.

3.4. Photometry

In each field, we produced pixel registered images with matching PSFs following the procedures described by Newman et al. (2012). Briefly, the PSF in each image was constructed by combining suitable bright stars. The ground-based and *HST* images were convolved to match the image with the worst seeing, which was always ⩽1″. Since the IRAC PSFs are considerably broader, we did not convolve all images to match the IRAC resolution. The arc flux measured in the IRAC images was instead scaled by the fraction of light lost from the photometric aperture when the F160W image was convolved to match the IRAC PSF (see Newman et al. 2012).

We measured colors in rectangular apertures aligned with the direction of magnification that had the following dimensions: 4″ × 1″.5 on MRG-M0138 Image 3, 1″.6 × 1″.2 on MRG-M0150 Image 1, 2″ × 1″.5 on MRG-P0918, 1″.5 × 1″.5 on MRG-S1522, and 3″.2 × 1″.5 on MRG-M2129. The lengths of these apertures (in the direction of maximum magnification) approximately match the spectroscopic apertures described in Section 6.1 and so are appropriate for a joint analysis of the

Table 3
Spectroscopic Observing Log

Target	Instrument	Dates	Exposure Time (hr)	Seeing	Slit PA (deg)
MRG-M0138 Image 1	MOSFIRE J	2015 Nov 3	1.0	0″73	82
MRG-M0138 Image 1	MOSFIRE H	2015 Nov 3, 6	2.0	0″79	82
MRG-M0138 Image 2	FIRE	2016 Sep 8–9	5.5	0″48	−60
MRG-M0150 Image 1	MOSFIRE H	2014 Nov 26–27	4.3	0″67	−7.5
MRG-M0150 Image 1	FIRE	2014 Nov 1, 3	6.5	0″55	−7.5
MRG-P0918	FIRE	2014 Apr 14–15	7.0	0″42	176
MRG-S1522	FIRE	2014 Feb 28–29, 2014 Apr 13–15	9.0	0″57	136
MRG-M2129	FIRE	2015 Sep 27–28, 2016 Sep 8–9	16.7	0″49	−13

spectra and photometry.[5] The rectangular apertures are wider (1″5) than the spectrograph slit (0″6–0″75), because we preferred to avoid selecting an aperture much smaller than the IRAC PSF. However, this does not introduce a significant mismatch between the spectroscopic and photometric apertures because the images are narrow in the direction of minimum magnification.

For each galaxy, the fluxes were then uniformly scaled to match the total F160W flux within a large aperture and corrected for Galactic extinction following Schlafly & Finkbeiner (2011). Uncertainties of 5% (3% for *HST* measurements) were added in quadrature to account for uncertainties in the PSF matching and photometric calibration. The resulting photometric measurements are listed in Table 2.

For MRG-M0138, colors were measured on Image 3, which is the most isolated and affords the cleanest photometry. In Table 2, we have scaled these fluxes to match the total observed F160W flux of the brightest image and our spectroscopic target, which is Image 1. (Although there is some ambiguity in separating the merging Images 1 and 2, this only affects the total fluxes in Table 2 and has no consequence for any of the inferred source properties, which will ultimately be scaled based on a model of the source.)

4. Spectroscopic Data

Each of the five lensed galaxies in our sample was observed using FIRE—a near-infrared echellette spectrograph that is mounted on the Magellan Baade telescope (Simcoe et al. 2013). MRG-M0138 and MRG-M0150 were also observed with the near-infrared spectrograph MOSFIRE at the Keck I telescope (McLean et al. 2012). Here we describe the observing strategy and reduction procedures for the spectroscopic data.

4.1. FIRE Observations

We generally used the 0″75-wide slit, which provides a spectral resolution of $\sigma_{\rm inst} = 33 \; {\rm km \, s}^{-1}$. For a portion of the MRG-M2129 observations that were conducted in excellent seeing, we used the 0″60-wide slit. To minimize read noise, we operated the detector in the up-the-ramp sampling mode for integration times of 20–30 minutes.

During twilight, we obtained sky exposures to measure the illumination of the slit. Although some of the target arcs could

be acquired directly on the slit-viewing acquisition camera, we usually acquired a nearby star and then offset the telescope to the target. After the offset, we compared the position of the offset star and other sources to their expected pixel coordinates in images taken with the acquisition camera. In some cases, these differed by up to 0″4, and we offset the telescope to place the sources at the expected positions. We monitored these positions throughout the exposure sequence and corrected gradual drift when it occurred.

The 6″-long FIRE slit was oriented in the direction of elongation of the target lensed galaxy (see P.A. in Table 3). Observations were made in an AB pattern with short dithers of 0″8–2″5 depending on the angular size of the target. A simple A-B subtraction is not feasible due to the resulting overlap of the extended target in the two dither positions and also to the long exposures needed to minimize the read noise. The spectra of the sky and target must instead be modeled in each exposure, as we describe below. Exposures of the internal quartz and ThAr lamps were interspersed throughout the FIRE observations and used for flat fielding and wavelength calibrations. To remove telluric absorption, A0V stars were observed both before and after each target and usually in the middle of longer exposure sequences.

The seeing is an important ingredient in our dynamical modeling. To measure the seeing, we monitored stars on the acquisition camera and on the facility guide camera. The seeing was also estimated from the science spectra themselves through a comparison with *HST* images (see Section 4.2). By comparing these methods, we estimate that the uncertainty in the seeing is ≲0″1. The mean seeing during the observations of each galaxy is listed in Table 3 and ranges from 0″42 to 0″57.

4.2. FIRE Data Reduction

The FIREHOSE pipeline[6] was used to flat field the data and to provide an initial wavelength solution and initial rectification of each spectral order. Since FIREHOSE was designed primarily for the reduction of point sources, in the subsequent steps we relied on custom IDL routines that were based on the FIREHOSE code.

First, traces of bright stars were used to make small corrections to the rectifications in each order. For each science exposure, we then masked all orders to isolate the inter-order background. We fit and subtracted a smooth variation with column within each of the four amplifier regions, which was

---

[5] Where slight aperture differences are present, these are used to avoid contamination from foreground objects that is negligible in the NIR spectra but may be significant in bluer filters.

[6] http://www.firespectrograph.org

necessary to remove discontinuities at the boundaries. We then modeled and subtracted a smooth scattered light background. Low-order corrections to the initial FIREHOSE wavelength solution, which was derived from ThAr lamp exposures, were then made using the OH lines. Once these low-order corrections were derived for one science exposure in a series, cross-correlations were used to correct for instrumental flexure.

Each spectral order of each exposure was then modeled as the sum of sky and galaxy emission using iterative bspline techniques as employed by FIREHOSE (see, e.g., Kelson 2003). This method requires that the spatial distribution of the emission be specified. For the sky background, we found that the residual intensity variation along the slit can be modeled as a quadratic polynomial. For the targeted galaxy, we first measured its spatial profile in the H band using the initial FIREHOSE sky model. However, this estimate is imprecise due to the difficulty of separating an extended source from the sky background over a short slit with no prior information on the source structure. Therefore, we used the HST WFC3/F160W image to measure the expected galaxy flux profile along the slit, taking into account its width and orientation. This HST-based flux profile was then shifted and convolved by a Gaussian PSF to best match the profile in the spectrum. This procedure provides an estimate of the seeing and the position of the target in each exposure, and it produces a galaxy profile suitable for accurate sky modeling.

Observations of A0V stars were analyzed with the xtellcorr package (Vacca et al. 2003), as implemented within FIREHOSE, to provide flux calibration and removal of telluric absorption. While the A0V observations are needed to track the temporal variation of telluric features, we found that the relative flux calibration could be improved through observations of white dwarf standards. These were reduced in the same way as the galaxy spectra and were then used to derive low-order corrections to the flux calibration in each order. Additionally, in a few cases, we made low-order corrections to the continuum shape based on comparisons to the fitted stellar population synthesis models (Section 6.2).

For each science exposure, two-dimensional rectified spectra were then produced from each order. The exposures were then normalized to a common flux level, measured in H band. All of the exposures in each order were then spatially registered and averaged using inverse variance weighting. Residual outlier pixels were then identified and interpolated over. One-dimensional spectra were then extracted in each order in a specified aperture. Small multiplicative offsets were applied to ensure that spectra extracted in adjacent orders have consistent fluxes within the wavelength regions of overlap. Finally, the orders were combined into a single one-dimensional spectrum with a scale of $12.5 \, \mathrm{km \, s^{-1} \, pixel^{-1}}$.

4.3. MOSFIRE Observations and Reduction

We observed two targets with MOSFIRE: Image 1 of MRG-M0150 in H band and Image 1 of MRG-M0138 in both J and H bands. We formed a $0\overset{''}{.}7$-wide long slit on each target and on several stars in the field, which were used to align the mask and to measure the seeing. The observations were conducted with an AB dither pattern. Since subtraction of consecutive dithered exposures is fundamental to the MOSFIRE Data Reduction Pipeline (DRP), we took care to ensure that the dither distance

($5\overset{''}{.}4$ for MRG-M0150 and $9\overset{''}{.}2$ for MRG-M0138) was sufficiently wide to avoid self-subtraction of the extended arcs.

The observations and reduction of MRG-M0150 were described by Newman et al. (2015a). To reduce the MRG-M0138 observations, we used the DRP to produce coadded 2D spectra for each night of observations (see Table 3). Relative flux calibration was then performed using twilight observations of the white dwarf GD71. Given the very high signal-to-noise ratio of the MRG-M0138 spectrum, we performed telluric absorption corrections differently from the other observations. In the H band, we iteratively modeled the galaxy stellar continuum and the telluric absorption using the radiative transfer code molecfit (Smette et al. 2015). A similar procedure was described by Newman et al. (2017). In the J band, where the signal-to-noise ratio is lower, we instead used molecfit to model the telluric absorption in observations of GD71. We then divided this synthetic absorption spectrum from the galaxy observations. Finally, observations made on different nights were registered, scaled to a common flux level, and averaged with inverse variance weighting.

5. Lens Models and Source Structures

For three of the galaxies in our sample, we are able to construct lens models to measure the magnification and reconstruct the source. This is necessary to estimate the stellar masses of the galaxies (Section 6) and to compare their sizes, ellipticities, and Sérsic indices with those of unlensed samples and thereby evaluate the representativeness of our lensed galaxy sample (Section 8.1). We will also compare the galaxies' structures to those of their likely $z \sim 0$ descendants to constrain their future evolution (Section 8.2) and use the lens models to interpret our stellar kinematic data (Paper II).

For MRG-M0150 and MRG-M0138, we constrained the lens mass distribution and the source light distribution using the detailed structure of the multiple images of the quiescent galaxies themselves. MRG-M2129 is singly imaged; however, since the lens is a well-studied cluster, we can rely on published lens models that are constrained by many multiple image systems. In this section, we describe the construction of these models. For the other two galaxies in our sample, the singly imaged MRG-S1522 and MRG-P0918, we cannot construct a meaningful lens model. MRG-S1522 lies at radii beyond the known multiple images, and consequently the mass distribution is not well constrained. Although the cluster magnifying MRG-P0918 produces several multiple image systems, their redshifts are not yet known. For these two systems, we do not estimate the magnification and will confine our analysis to magnification-independent quantities (e.g., σ, V/σ, age, specific star formation rate, and emission line ratios).

5.1. MRG-M0138

5.1.1. Lens and Source Modeling Methods

We used the ray tracing code introduced by Newman et al. (2015b) and employed by Newman et al. (2015a) to fit a simply parameterized model of the lens mass distribution and the source light distribution to the pixel-level data. The regions of the WFC3/F160W image that were used to constrain the model are outlined in blue in Figure 4(a). The BCG light has been modeled and subtracted in this image.
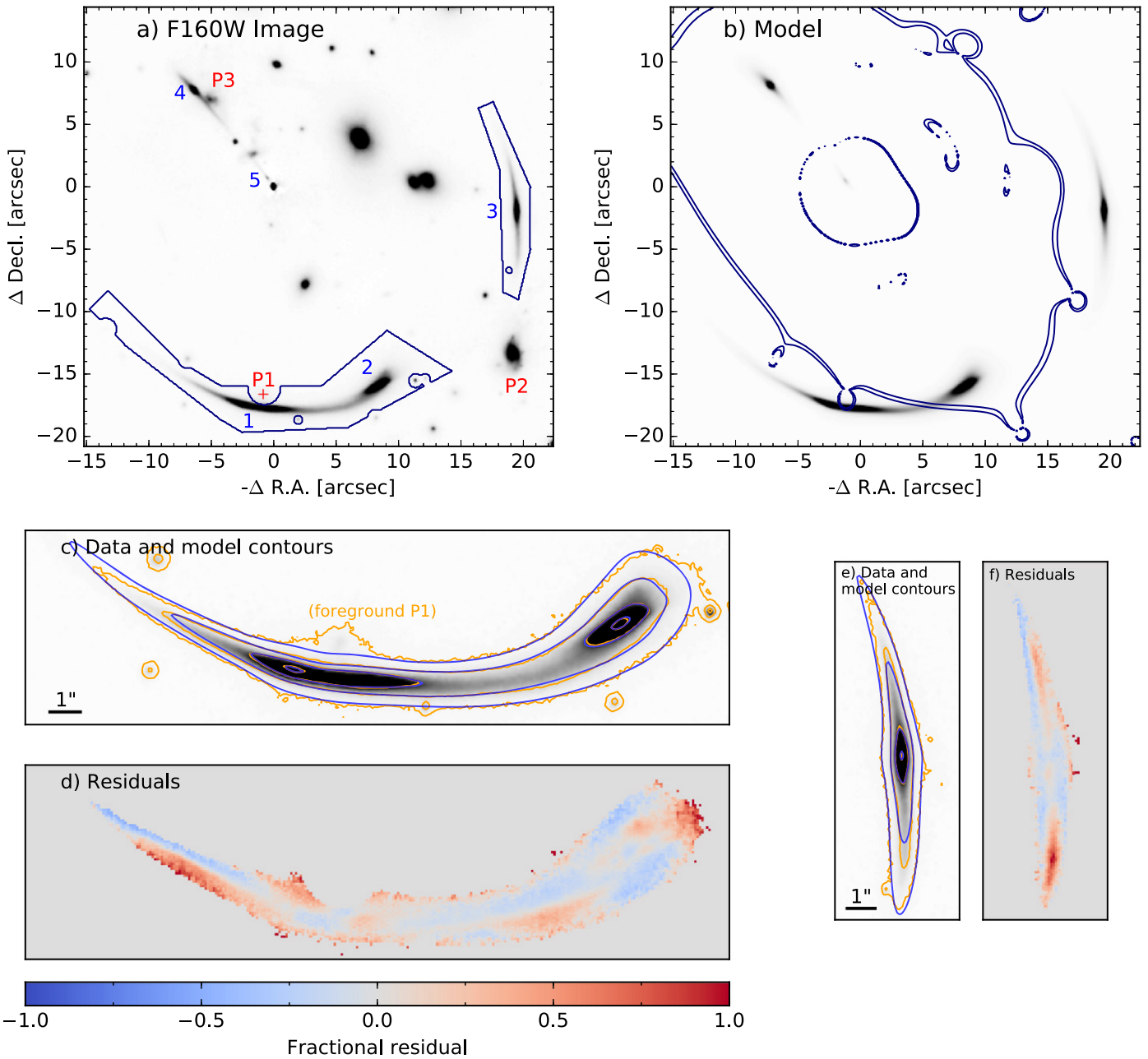
Figure 4. Lens model of MRG-M0138. Panel (a) shows the *HST*/WFC3 F160W image with the BCG subtracted, displayed with a linear stretch. The regions used to constrain the lens and source model are outlined. Blue labels number the images. Coordinates are relative to the BCG center. Panel (b) is the model of the image plane produced by a double Sérsic model of the source traced through the lensing potential and convolved by the PSF. Colored curves enclose the critical line. Panels (c) and (e) show zooms of panel (a) with orange and blue contours of the data and model image, respectively. Note that the orange contours include flux from several foreground galaxies, including P1, which were masked during the fit. Panels (d) and (f) show the fractional residuals.

The lensing cluster was modeled as a superposition of dual pseudo-isothermal elliptical (dPIE) mass distributions (see Elíasdóttir et al.'s 2007 appendix). Each dPIE was described by a center, PA, ellipticity, two characteristic radii r_{core} and r_{cut}, and a normalization σ_0. Since the lensing cluster appears to be a simple, relaxed system, we found that a single dPIE component is adequate to model the cluster dark matter halo. We left all of its parameters free except for r_{cut}, which lies well beyond the strong lensing zone and can be fixed to 1 Mpc.

We modeled the stellar mass in the BCG by fitting a dPIE profile to the surface brightness distribution to set

r_{core} and r_{cut}. We used Gaussian priors of PA $= 46° \pm 10°$ and $b/a = 0.42 \pm 0.05$ that were informed by the photometry. The normalization $\sigma_{0,BCG}$ was allowed to vary freely.

Other cluster galaxies were generally included in the mass model using scaling relations with luminosity: $\sigma_0 = \sigma_0^*(L/L_*)^{1/4}$ and $r_{cut} = r_{cut}^*(L/L_*)^{1/2}$, where σ_0^* and r_{cut}^* are free parameters with observationally motivated priors (see Newman et al. 2013). This approach ties the center, PA, and ellipticity of the mass distribution to the galaxy light and imposes a constant mass-to-light ratio. The radial distribution of the mass, encoded by r_{cut}, can differ from the light. Similar approaches have been widely used in
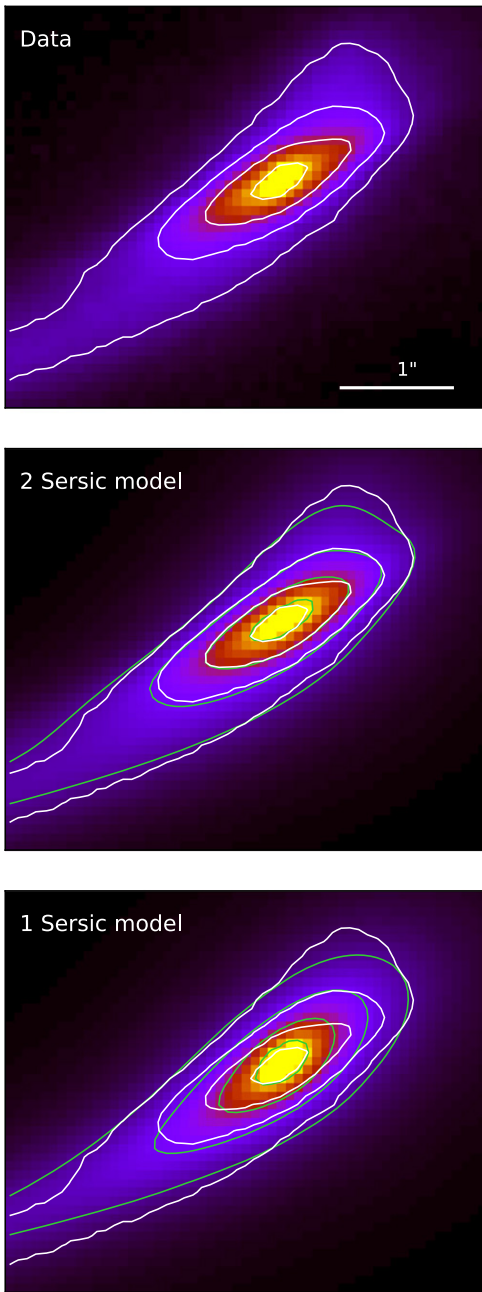
Figure 5. Comparison of Image 2 of MRG-M0138 to two models. The two-component source model (middle panel) matches the data (top panel) much better than a single-component source model (bottom panel), particularly the flattened and tilted inner structure. White contours in the upper panel are repeated in the middle and lower panels, where the green contours trace the models.

other parameterized lens models (e.g., Jullo et al. 2007; Richard et al. 2010; Newman et al. 2013).

Three galaxies either significantly perturb the critical lines or are likely to deviate from these general scaling relations and, therefore, were modeled independently of them. The galaxy labeled P1 in Figure 4(a) is located very close to Image 1, and it significantly alters the critical line and increases the magnification. This perturbing galaxy is faint and not easily visible in Figure 4, but it can be seen in Figure 2. We allowed all of the parameters describing the P1 mass distribution to vary freely. Galaxies P2 and P3 are blue cluster galaxies that show spectacular examples of ram pressure stripping in Figure 2.

Given their blue colors, we do not expect them to share a common mass-to-light ratio with the other cluster members. We allowed σ_0 to vary freely for both P2 and P3, as well as r_{cut} for P2. The final ingredient in the lens model is an external shear that is uniform across the image, which we ultimately found to be small, $|\gamma| = 0.05$.

We initially modeled the source light distribution using a single elliptical Sérsic profile (although, as described below, we ultimately adopted a two-component model). We denote the effective radius as $R_{e,\mathrm{maj}}$ to emphasize that it is the semimajor axis of the ellipse containing half of the light. For a given set of parameters that describe the lens model and source, we first cast the centroids of Images 1–4 back to the source plane and then computed the dispersion in their positions. Models with a large dispersion were immediately discarded. For the remaining models, we traced pixels in the image plane back to the source to compute the surface brightness. Care was taken to integrate the surface brightness within the pixels near the center of the source. The model image plane was then convolved by the PSF and compared to the data. A Markov Chain Monte Carlo (MCMC) sampler, MultiNest (Feroz et al. 2009), was used to explore the 29-dimensional parameter space.

Although this model was able to reproduce the overall lens configuration, we found that it could not match the internal structure of all of the images in detail. The clearest example of this can be found in Image 2, which we show in Figure 5. This image presents a flattened inner component that is misaligned from the overall direction of the arc (upper panel). A model with a single-component Sérsic source (lower panel) is not able to reproduce this. However, this deficiency can be eliminated by introducing a two-component source: this model matches the morphology of the image much better in the inner parts (middle panel). Specifically, we modeled the source as the sum of two Sérsic components that share a common center and PA, but which have different axis ratios b/a, effective radii $R_{e,\mathrm{maj}}$, Sérsic indices n, and magnitudes.

To assess the fit quality, we compare the observed image to the model image plane in panels (c)–(f) of Figure 4. Although some areas of mismatch are visible, the model contours (blue) generally follow the data (orange) well. (The contours near the "foreground P1" label include flux from the foreground galaxy P1, which was masked during the fit.) We note that the radial images 4 and 5 were not used to constrain the model except via their approximate positions, so a close match to their detailed structure is not expected.

The magnification factors[7] for images 1, 2 and 3 are $\mu_1 = 12.5 \pm 5.4$, $\mu_2 = 10.3 \pm 3.1$, and $\mu_3 = 4.9 \pm 1.6$, whose uncertainties are described in the next section.

5.1.2. Magnification Uncertainties

The information in the pixel-level *HST* data over-constrains the lens model and results in minuscule formal uncertainties. However, it is known that different lens modeling assumptions and procedures can lead to different estimates of the magnification. Although these uncertainties are difficult to quantify, we attempted to estimate them by constructing a set of Lenstool models (Kneib et al. 1993; Jullo et al. 2007). These models are constrained by the positions and ellipticities of the images rather than the pixel-level data; however, the

---

[7] We define the mean magnification over an image as $\langle \mu \rangle = \sum \mathrm{Image}(x, y) / \sum [\mathrm{Image}(x, y)/\mu(x, y)]$, where the sum is over the pixels in the mask. This is equivalent to the ratio of the image and source fluxes in the limit of an infinite aperture.
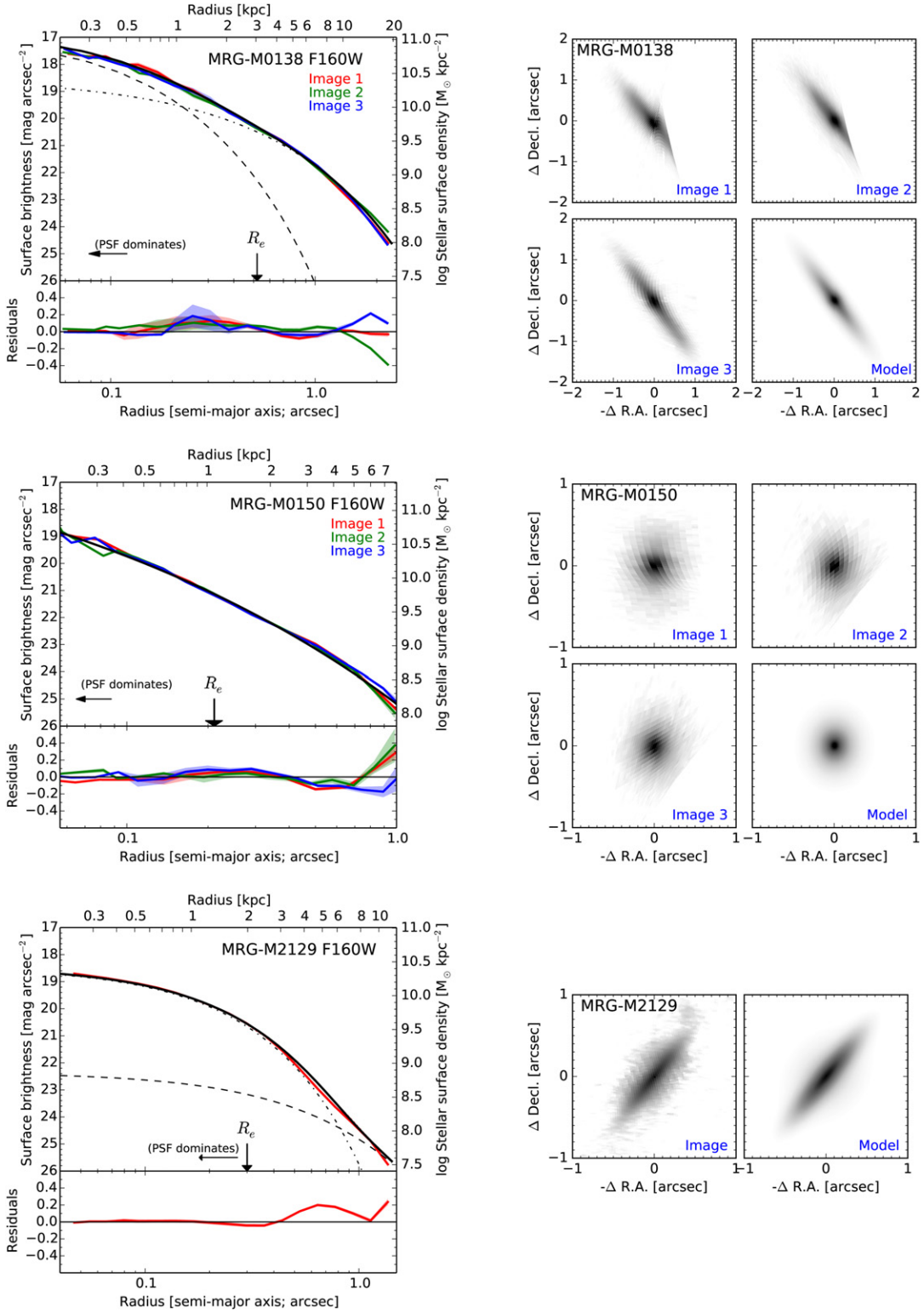
Figure 6. Surface brightness profiles (left panels) and source plane reconstructions (right) of the three galaxies with lens models. These are constructed using the deconvolved images Data − Model$_{convolved}$ + Model$_{unconvolved}$, as described in Section 5.1.3. Left panels: the solid colored lines show the profiles constructed from each multiple image. Solid back lines show the analytic profile of the best-fit single- or double-Sérsic model, with dashed and dot-dashed curves showing the individual contributions of the two components for MRG-M0138 and MRG-M2129. Residuals between the image and model are plotted in mag arcsec^{-2}. For multicomponent models, the surface brightness has been averaged within elliptical annuli whose axis ratio is that of the single-Sérsic model in Table 4. The right axes show the corresponding stellar surface mass density based on the global mass-to-light ratios inferred in Section 6. Right panels: for each image, the corners of each pixel were cast back to the source plane, conserving the surface brightness, to produce these reconstructions. Images are displayed using an arcsinh stretch. Note that the axis ranges are not the same for every galaxy.

Table 4
Sérsic Model Parameters

Quantity	Units	Single Sérsic Model	Double Sérsic Model	
MRG-M0138				
$m_{\rm F160W}$	mag	$20.0 \pm 0.1 \pm 0.4$	$20.4 \pm 0.1 \pm 0.4$	$21.5 \pm 0.1 \pm 0.4$
$R_{e,\rm maj}$	arcsec	$0.''57 \pm 0.''07 \pm 0.''12$	$0.''85 \pm 0.''07 \pm 0.''18$	$0.''10 \pm 0.''02 \pm 0.''02$
$R_{e,\rm maj}$	kpc	$4.8 \pm 0.6 \pm 1.0$	$7.1 \pm 0.6 \pm 1.5$	$0.8 \pm 0.2 \pm 0.2$
n	…	2.9 ± 0.7	1.3 ± 0.3	1[a]
b/a	…	0.26 ± 0.04	0.19 ± 0.03	0.67 ± 0.16
PA	deg	36 ± 2	36 ± 1	
MRG-M0150				
$m_{\rm F160W}$	mag	$21.3 \pm 0.1 \pm 0.4$		
$R_{e,\rm maj}$	arcsec	$0.''21 \pm 0.''03 \pm 0.''04$		
$R_{e,\rm maj}$	kpc	$1.7 \pm 0.2 \pm 0.3$		
n	…	3.5 ± 0.4		
b/a	…	0.87 ± 0.05		
PA	deg	-6 ± 10		
MRG-M2129				
$m_{\rm F160W}$	mag	21.8 ± 0.2	22.0 ± 0.2	23.4 ± 0.1
$R_{e,\rm maj}$	arcsec	$0.''29 \pm 0.''02$	$0.''27 \pm 0.02$	$0.''47 \pm 0.05$
$R_{e,\rm maj}$	kpc	2.4 ± 0.2	2.2 ± 0.2	3.9 ± 0.4
n	…	1[a]	1[a]	1[a]
b/a	…	0.29 ± 0.03	0.24 ± 0.03	0.80 ± 0.07
PA	deg	-39 ± 5	-39 ± 5	
Gaussian $m_{\rm F160W}$	mag	25.9 ± 0.5	26.1 ± 0.3	
Gaussian R_e	arcsec	$<0.''007$ (95%)	$<0.''008$ (95%)	

Notes. For the extrinsic parameters (magnitudes and R_e) of the multiply imaged systems, we first list the uncertainty derived from the scatter among the multiple images (Section 5.1.4) and then list the systematic uncertainty arising from the overall magnification uncertainties (Section 5.1.2).
[a] Edge of prior.

cruder constraints make it feasible to vary several assumptions. In particular, we varied the type of constraint (image positions, positions and fluxes, source versus image plane fits), the radial density profile of the cluster (dPIE versus generalized Navarro-Frenk-White), and the inclusion of the BCG or some perturbing galaxies as separate mass components. For each of multiple images 1–3, we evaluated the maximum difference in the magnification between the fiducial model and the set of `Lenstool` models. We find that the uncertainties range from 30% to 43%. These results are comparable to the differences between lens models of Frontier Fields clusters discussed by Priewe et al. (2017) and Meneghetti et al. (2017). When describing the source properties, we will conservatively use the maximum uncertainty among these multiple images as a systematic uncertainty in the luminosity and stellar mass, i.e., 43% or 0.19 dex. Since this factor applies to the areal magnification, we approximate the fractional error in R_e as half of this.

5.1.3. Reconstructing the Source Plane

Our reconstruction of MRG-M0138 in the source plane is shown in Figure 6. Given the nonlinear nature of the lens mapping, convolution by the PSF in the image plane can have very complex effects in the source plane. This makes it difficult to compare the consistency of reconstructions derived from different images. To address this issue, we do not directly cast the observed pixels back to the source plane. Instead, we take the unconvolved image plane model, add the residuals, and then cast this image (i.e., Data − Model$_{\rm convolved}$ + Model$_{\rm unconvolved}$) back to the source plane. This technique effectively deconvolves the inner regions of the images under the assumption of a particular source model, while still allowing for deviations from the model at larger radii

(see Szomoru et al. 2012 for an application in a non-lensing context). All three images consistently show a highly flattened, disk-dominated source. We note that since Images 1 and 2 merge into a giant arc, the southwestern portion of the source is not present in these images.

In Figure 6, we also compare the surface brightness profiles derived from these three source plane reconstructions. The three multiple images reproduce a single surface brightness profile (solid colored lines) with impressive consistency. This profile is well fit by the two-component Sérsic model (solid black line), leaving azimuthally averaged residuals of $\lesssim 10\%$ out to $R = 10$ kpc. The leftward arrow indicates the radius at which the effects of PSF convolution are significant; at smaller radii, the shape of the plotted profile is dominated by the assumption of the double Sérsic form.

5.1.4. Source Structure

In the best-fitting two-component model, 74% of the stellar light is in an extended ($R_{e,\rm maj} = 7.1$ kpc), flattened ($b/a = 0.19$), disk-like component with a Sérsic index of $n = 1.3$. A fainter (26% of total flux), rounder ($b/a = 0.67$), and much more compact ($R_{e,\rm maj} = 0.8$ kpc) component is also present. It has a Sérsic index $n = 1$, the smallest value allowed in the fits, which could indicate a structure analogous to a pseudobulge. However, we caution that the images are significantly affected by the PSF within the effective radius of the bulge-like component (Figure 6).

Table 4 lists the Sérsic parameters and uncertainties. The uncertainties were derived from the image-to-image scatter: specifically, we fixed the lens mapping, fit the source to each of the three images individually, and then measured the standard deviation of each parameter. For the total flux and R_e, we
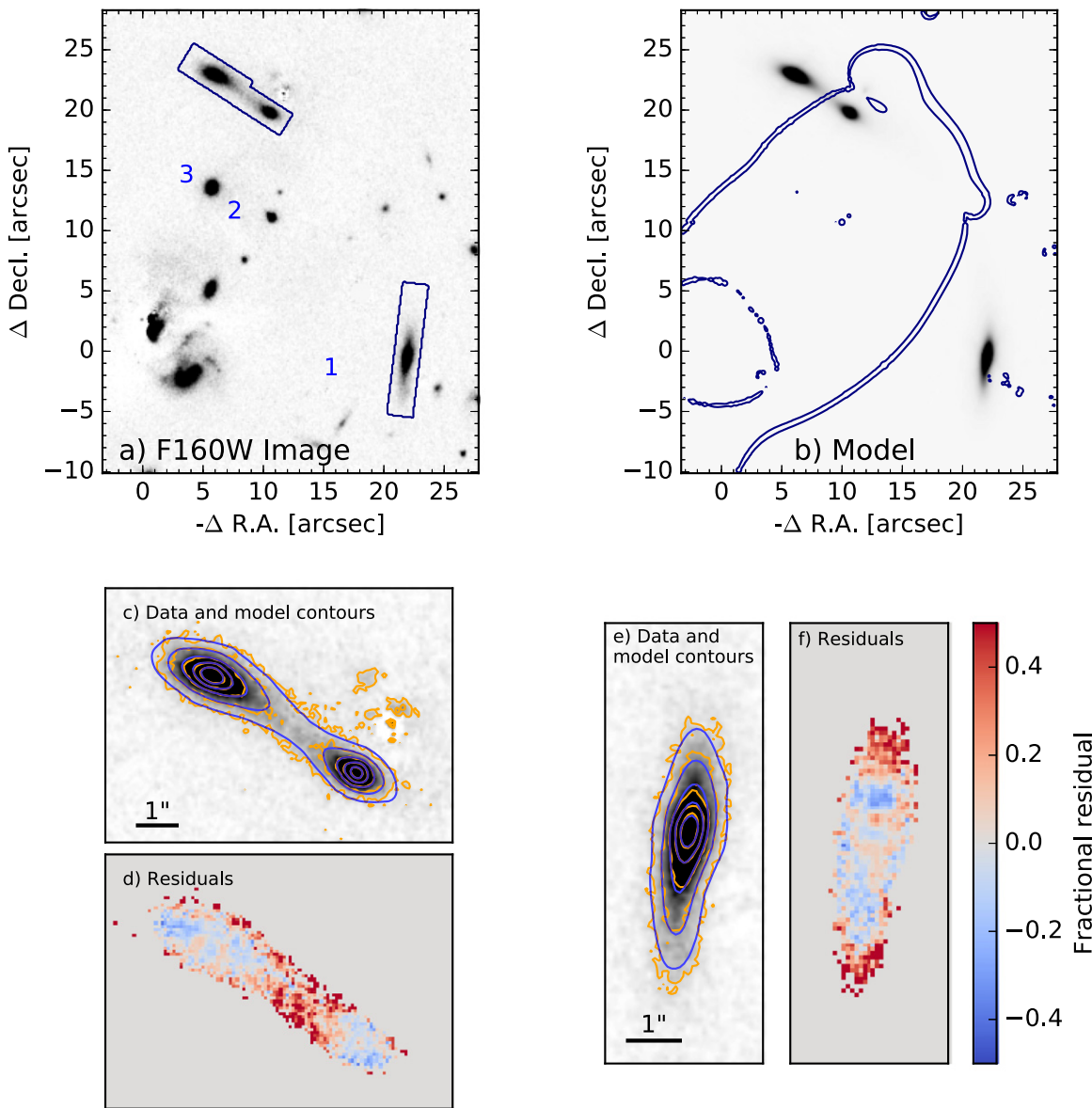
Figure 7. Updated lens model of MRG-M0150. Panel (a) shows the *HST*/WFC3 F160W image with the BCG subtracted, displayed with a linear stretch. The regions that are used to constrain the lens and source model are outlined. Blue labels number the images. Coordinates are relative to the BCG center. Panel (b) is the model of the image plane produced by the single-component Sérsic model of the source traced through the lensing potential and convolved by the PSF. Colored curves enclose the critical line. Panels (c) and (e) show zooms of panel (a) with orange and blue contours of the data and model image, respectively. Panels (d) and (f) show the fractional residuals.

additionally list the systematic lens model uncertainties discussed in Section 5.1.2.

5.2. MRG-M0150

In Newman et al. (2015a) we constructed a parametrized model of the lensing cluster and source that was constrained by the archival *HST* WFC3/F140W image of MACSJ0150.03−1005 that was then available. This technique was very similar to that described in the previous subsection for MRG-M0138, and its application to MRG-M0150 was detailed by Newman et al. (2015a). Subsequently, we obtained a deeper WFC3 image at the slightly redder wavelength of F160W with improved sub-pixel sampling (Section 3.1). Here we briefly describe our improved analysis using this new image.

Figure 7 shows the three multiple images 1–3 (panel (a)), the single-component Sérsic source traced through the lensing

potential and convolved by the PSF (panel (b)), and zooms on the image regions to compare the model contours and residuals (panels (c)–(f)). The model successfully reproduces the structure of the images in detail. The magnifications of the images are $\mu_1 = 4.4 \pm 1.1$, $\mu_2 = 2.6 \pm 1.0$, and $\mu_3 = 4.6 \pm 1.3$. These uncertainties are estimated by varying the parameterization of the mass model and then comparing them to the results from an independent set of lens models constructed with the Lenstool code, as described in Section 5.1.2 and Newman et al. (2015a).

Source reconstructions and surface brightness profiles from each multiple image are shown in Figure 6.[8] The lens model produces three consistent reconstructions of the source. Likewise, the surface

---

[8] The surface brightness profile differs superficially from Figure 2 of Newman et al. (2015a) because we now cast the PSF-deconvolved image back to the source plane, as described in Section 5.1.3. Jaggedness in the innermost regions is due to pixelization.
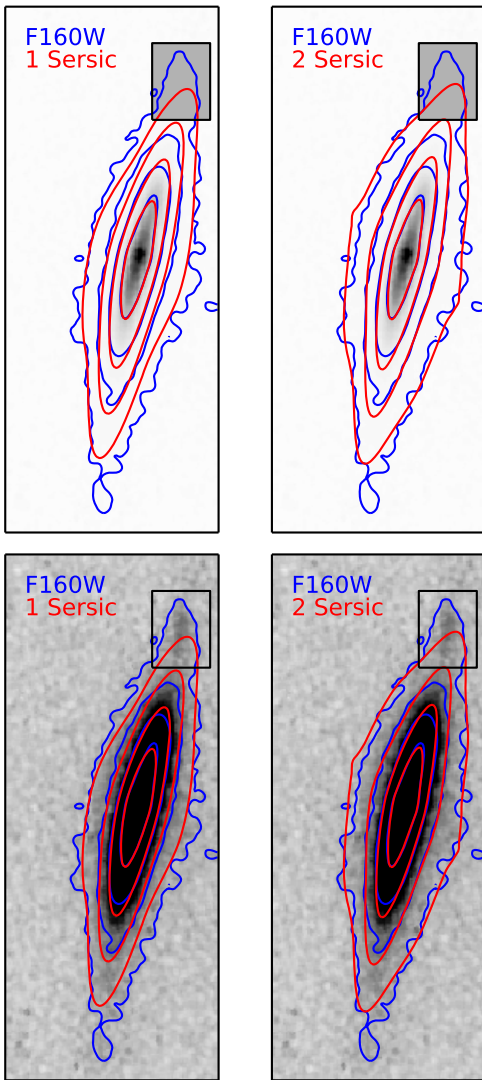
Figure 8. Model of MRG-M2129. The WFC3/F160W image is shown in each panel with logarithmically spaced contours in blue. Red contours in the left and right panels show the models based on a single- and double-Sérsic source, respectively. The top and bottom rows differ only in the stretch of the image. The unresolved nuclear component is more easily seen in the top row, while a detached component (potentially a satellite galaxy) is more easily seen in the bottom row. The black box encloses the region masked during the fit due to the detached component. The double Sérsic model fits the outer contour better than the single-Sérsic model.

brightness profiles of the source from the observed multiple images are mutually consistent at the $\simeq$5% level out to $4R_e$.

The Sérsic parameters are listed in Table 4. The source is a compact galaxy ($R_{e,\mathrm{maj}} = 1.7$ kpc) with a nearly de Vaucoulers profile ($n = 3.5$) and nearly round isophotes ($b/a = 0.86$). Unlike MRG-M0138, we find that a single Sérsic component is adequate to describe the source, but we note that it would be much more difficult to discern the presence of multiple components in this case since MRG-M0150 is nearly round in projection. These parameters are consistent with those measured using the shallower F140W image by Newman et al. (2015a) but the uncertainties are reduced using the new deeper data.

5.3. MRG-M2129

MRG-M2129 is singly imaged, so unlike MRG-M0138 and MRG-M0150, we require other multiple images to constrain the source structure. The lensing cluster MACSJ2129.4–0741

presents a large number of multiple images that are available to constrain the mass distribution, and several authors have published models. We used the deflection angle maps produced by several published lens models to trace a model source through the lens mapping. The model image plane can then be convolved by the PSF and fit to the *HST*/WFC3 F160W image. Since this system was discovered by Geier et al. (2013) and studied by Toft et al. (2017), we also compare our results to theirs in the Appendix.

We considered three lens models. Monna et al. (2017) kindly supplied maps of the deflection angle for their parametric model. We also obtained two mass models constructed by A. Zitrin (Zitrin et al. 2009, 2013) as high-level science products of the CLASH program (Postman et al. 2012). We found that the Monna et al. (2017) lens model produced the lowest χ^2. Therefore, we used the parameters derived from it and estimated the uncertainties in the galaxy structural parameters by comparing to those obtained using the two Zitrin et al. lens models.

Figure 8 shows the region of the *HST*/WFC3 F160W image that we fit. We first modeled and subtracted a foreground cluster member that is located to the west of the arc. To avoid any residual contamination from this galaxy, the fit region is not centered on the arc. We then subtracted the sky background measured in an empty region to the east of the arc. In this case, the lensed galaxy is sufficiently distant from the BCG so that it was unnecessary to model the latter. A second source is visible in the upper right corner, which is disconnected from the main galaxy (see also Geier et al. 2013). We masked this object, which is potentially a satellite galaxy.

We found that a single Sérsic model is inadequate to model MRG-M2129, both in the central and outer regions. First, there is a centrally located, point-like source visible in the top row of Figure 8. We modeled this by adding a second component to source model: a circular Gaussian at the galaxy center with a free effective radius R_e and flux. The inclusion of the central source improved the fit by $\Delta\chi^2 = 759$ and, therefore, is clearly justified. Its magnitude is $m_{\mathrm{F160W}} = 25.9 \pm 0.5$ (demagnified), or $\sim$1% of the galaxy flux, and it is very compact: $R_e < 0\rlap{.}''007 = 60$ pc (95% confidence). The compact central source could be continuum emission from a Seyfert nucleus, which is supported by the emission line ratios (Section 7.2). The properties of the single Sérsic+Gaussian source model are listed in Table 4.

Second, we found that a single-component Sérsic model does not reproduce the shape of the outer isophotes well, which is evident in the upper-left and lower-right regions of the images in Figure 8. The addition of a second Sérsic component with a center and PA tied to those of the first component improved the fit. This can be seen qualitatively in the right panels of Figure 8 and is supported quantitatively by an improvement of $\Delta\chi^2 = 1287$. Like MRG-M0138, MRG-M2129 shows evidence for two components with different ellipticities.

The source model consists of: (1) the central source described above, (2) a flattened ($q = 0.24$) component containing 79% of the flux in a compact ($R_{e,\mathrm{maj}} = 0\rlap{.}''27 = 2.2$ kpc) exponential disk ($n = 1$), and (3) a rounder ($q = 0.80$) component containing 21% of the flux that is more extended ($R_{e,\mathrm{maj}} = 0\rlap{.}''47 = 3.9$ kpc) and also nearly exponential ($n = 1$). The parameters are summarized in Table 4. We restricted the Sérsic index to values $n > 1$ in the fits. Since both components hit this limit, their Sérsic indices

should be regarded with caution. However, the presence of two components—one compact and flattened, one fainter, rounder, and more extended—seems robust and holds for all three lens models. The uncertainties in Table 4 are derived from the standard deviation of values obtained using the three lens models.

5.4. Summary of Galaxy Structures

The structures of MRG-M0138 and MRG-M2129 are both dominated by a highly flattened exponential component, i.e., a nearly edge-on disk. Both galaxies also contain additional components. In MRG-M0138, we find a compact rounder component emitting 26% of the luminosity, which is potentially a nascent bulge that is embedded in a very massive and extended ($R_e = 7$ kpc) disk. In MRG-M2129 we find an extended rounder component and, remarkably, an unresolved central point source that could be Seyfert nucleus. This is particularly interesting given the evidence for an active galactic nucleus (AGN) from the emission line ratios, which we will discuss in Section 7.2. MRG-M0150 is instead well described by a single compact component with a nearly de Vaucouleurs profile ($n = 3.5$) and round isophotes. With imaging alone, we cannot tell whether MRG-M0150 is intrinsically rounder than the other two systems or is merely less inclined, but we will address this question using stellar kinematics in Paper II.

The disk-dominated structures of MRG-M0138 and MRG-M2129 are quite distinct from massive early-type galaxies in the local universe. In Section 8.2, we will compare our sample to local systems and will discuss the implications for their future evolution.

6. Unresolved Stellar Populations

We now use the photometric and spectroscopic data that were collected in Sections 3 and 4, together with the magnification factors derived in Section 5, to measure the redshifts, stellar masses, and ages of the lensed galaxies in our sample and to establish their quiescence. For the purposes of this paper, we will focus on representative values derived from the integrated spectra, deferring an analysis of their spatially resolved stellar populations.

6.1. Spectral Extraction

To define the extraction region for each target, we examined the flux distribution along the slit. The extraction box was chosen to cover the region where the flux profile was $\gtrsim 0.15$ times the peak intensity.[9] As discussed in Section 3.4, these approximately match the apertures in which the colors were measured. For our analysis, we consider the wavelength range from $\lambda_{rest} = 3600$ Å to $\lambda_{obs} \simeq 2.3\ \mu m$. The spectrum outside the range is generally too noisy to be useful. Only for the highest-redshift source, MRG-M0150, did we reduce the spectrum out to $\lambda_{obs} = 2.45\ \mu m$ to include Hα and [N II].

The spectra are shown in Figure 9, where they have been rebinned for display purposes, as indicated in the caption. The quality is remarkable considering the redshifts of these sources.

The median signal-to-noise ratio in the H band generally ranges from 21–32 per 300 km s^{-1} bin, which is approximately one velocity dispersion element. For the ultra-bright MRG-M0138, this figure reaches 137 and 77 for the MOSFIRE and FIRE spectra of Images 1 and 2, respectively.

6.2. Spectral Modeling

We jointly modeled the spectra and photometry using the `pyspecfit` code described by Newman et al. (2014). We used the Bruzual & Charlot (2003, BC03) population synthesis models and adopted an exponentially declining star formation history SFR $\propto \exp(-t/\tau)$, the Calzetti et al. (2000) dust attenuation curve, and the Chabrier (2003) initial mass function. For each galaxy we fit the redshift, velocity dispersion, age,[10] τ, metallicity Z, A_V, and the emission line parameters that are described below. As described in Newman et al. (2014), when comparing the data to a given model, we warped the spectrum continuum shape by the polynomial that minimizes the χ^2. This accounts for flux calibration errors in the spectrum and anchors its continuum shape to the photometry. We allowed a separate polynomial of order $\simeq \Delta\lambda_{rest}/200$ Å in each of the J, H, and K bands, where $\Delta\lambda_{rest}$ is the unmasked rest-frame wavelength interval.

Most of the galaxies in the sample exhibit weak emission lines whose measurement requires accurate modeling of the stellar continuum. Therefore, we fit both simultaneously. Emission lines from [O II], [O III], [N II], [S II] and the Balmer series were modeled as Gaussians with a velocity and velocity dispersion σ_{em} that are common to all emission lines but are distinct from those of the stellar component. For each galaxy, we modeled only the lines falling in the atmospheric transparency windows. The intensity ratios [O III] $\lambda 5008/\lambda 4960$ and [N II] $\lambda 6585/\lambda 6550$ were fixed to 2.98 and 3.05 (Storey & Zeippen 2000), while for the Balmer series we fixed the relative intensities assuming Case B recombination. Since Hβ is always in net absorption, we cannot separately constrain the emission line attenuation via the Balmer decrement. Therefore, once the emission line spectrum was constructed based on the aforementioned ratios, we attenuated it at each wavelength by the same factor as the starlight. We will consider the effects of possible differential extinction in our interpretation; these are expected to be small because the continuum attenuation is mild.

The posterior distributions were sampled using `MultiNest`. In addition to the model parameters, we also computed derived quantities, such as the specific star formation rate (sSFR) and emission line ratios. The priors were broad and uninformative, with two exceptions: we restricted the stellar metallicity to $Z = 0.01$–0.05 (the solar value is 0.02 in the BC03 models) and placed a Gaussian prior on $v_{gas} - v_{stars}$ with a mean of 0 and a dispersion of 200 km s^{-1}.[11]

The posterior constraints on the stellar populations and emission lines are listed in Table 5. As Figure 9 shows, these models generally fit the photometry and spectra well. For the photometry, the reduced χ^2_{phot} is in the range 0.4–1.5. Since deep NIR spectra often exhibit noise somewhat above the formal estimate, we rescaled the spectral uncertainties by a constant factor in the range 1.0–2.5 so that the reduced $\chi^2_{spec} \simeq 1$. We masked

---

[9] Relative to the peak position, the boundaries of these regions were $-5\rlap{.}''6$ to $2\rlap{.}''0$ for Image 1 of MRG-M0138, $\pm 1\rlap{.}''25$ for Image 2 of MRG-M0138, $-1\rlap{.}''5$ to $+1\rlap{.}''2$ for Image 1 of MRG-M0150, $\pm 1\rlap{.}''0$ for MRG-P0918, $\pm 0\rlap{.}''7$ for MRG-S1522, and $\pm 1\rlap{.}''6$ for MRG-M2129.

[10] The age refers to the time between the observation epoch and the beginning of the exponential star formation history.

[11] For MRG-M0138, the only emission line covered in the FIRE spectrum is [O II], which is not detected. In this case, we tied the gas velocity and dispersion to those of the stars to derive flux limits.
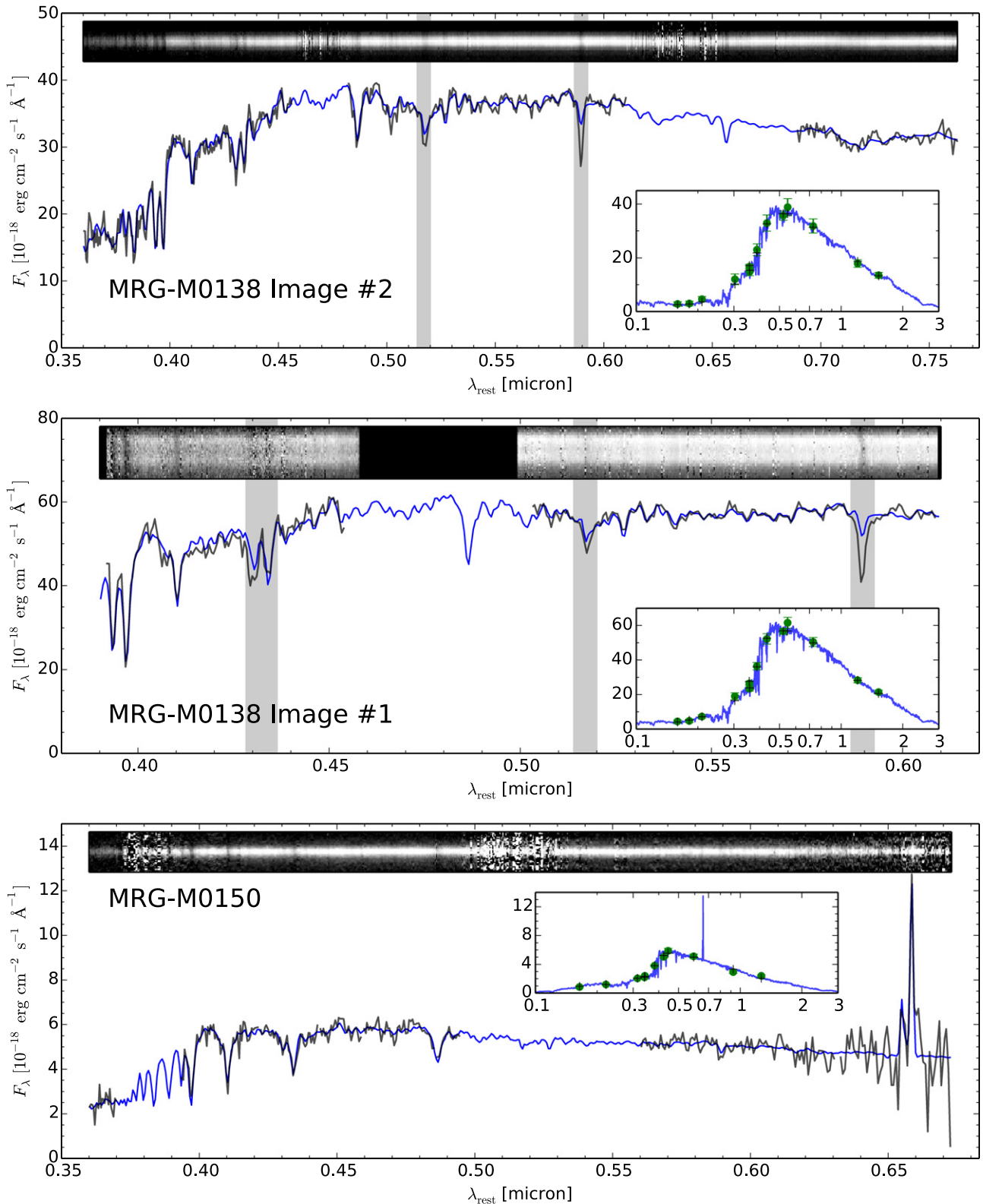
Figure 9. Integrated spectra of the lensed quiescent galaxy sample. Gray curves show the high-resolution spectra after taking the inverse variance weighted mean in 30 pixel (375 km s^{-1}) wide spectral bins. The blue curve shows the best-fit stellar and emission line model described in the text with the the same rebinning applied. Gray bands indicate regions that were masked for the fit. The inset shows the broadband photometry (green circles) and the model magnitudes (crosses). The axes of the insets have the same units as the main panels. The top of each panel shows the two-dimensional spectrum.

Mg b since the galaxies may have non-solar abundance ratios. For MRG-M0138, we also masked Na D, which clearly has a non-solar abundance or is affected by interstellar absorption. For the

MOSFIRE spectrum of MRG-M0138, we also masked the region around the G band and Hγ due to imperfect correction of telluric absorption in this region.
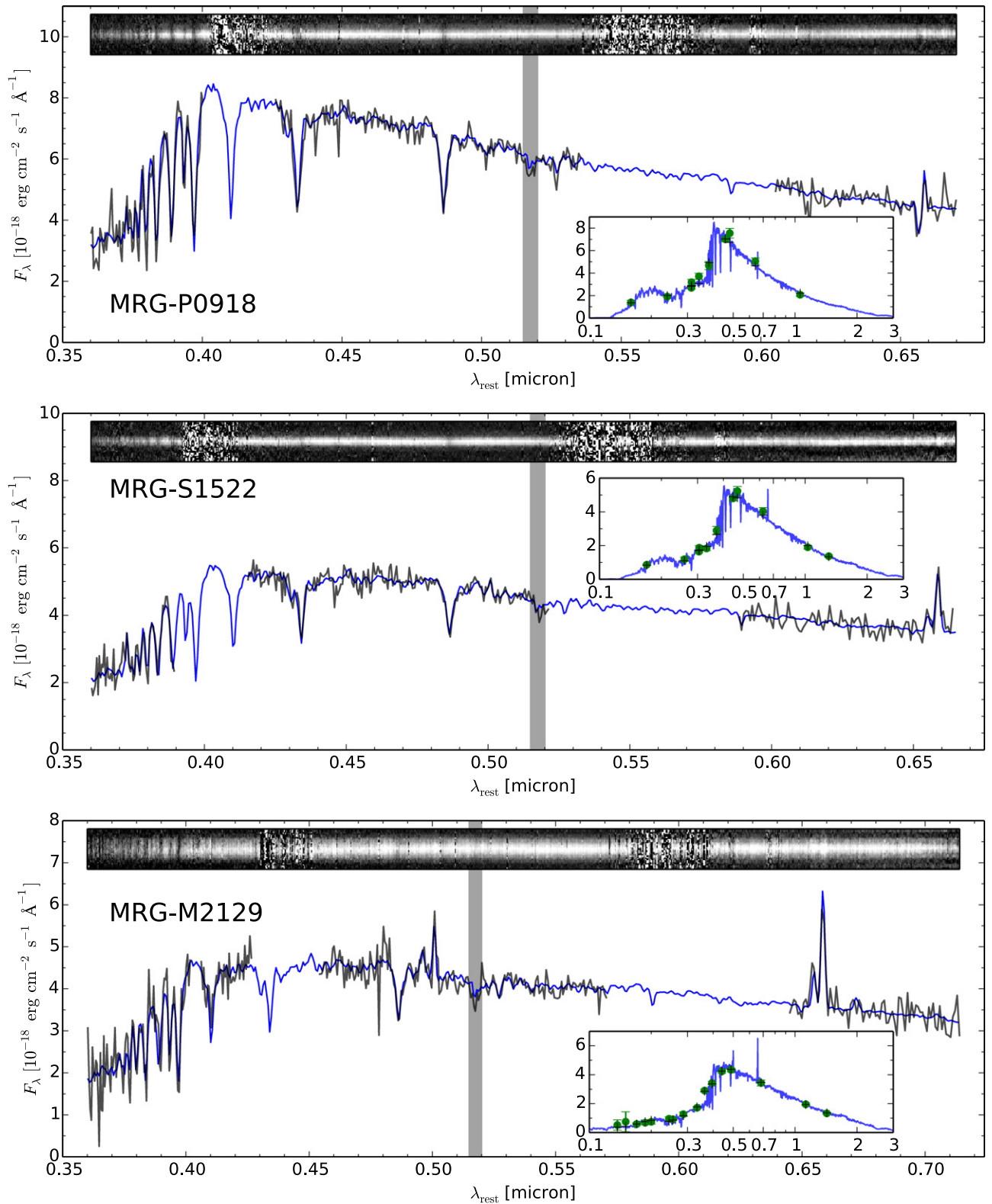
Figure 9. (Continued.)

6.3. Massive, Quiescent Stellar Populations

Consistent with the goal of our color–magnitude selection, our modeling of the stellar continuum indicates high stellar masses and low sSFRs for all five lensed galaxies in the sample. Their stellar masses, uncorrected for magnification, span the range $\mu M_* = 10^{11.6-12.8}\,M_\odot$. For the three systems with estimated magnification factors μ, we find $M_* = 10^{11.0-11.7}\,M_\odot$. The specific star formation rates (sSFRs), which are independent of magnification, are $10^{-10.7}\,\mathrm{yr}^{-1}$ or smaller. As we will discuss in Section 7, emission lines (when present) do not have ratios indicative of star formation.

Table 5
Spectroscopic Measurements

Quantity	Units	MRG-M0150	MRG-P0918	MRG-S1522	MRG-M2129	MRG-M0138 Image 1	MRG-M0138 Image 2
				Stellar Population Properties			
z	...	2.6355	2.3559	2.4503	2.1487	1.9486	1.9469
σ	km s^{-1}	261 ± 30	223 ± 16	241 ± 18	266 ± 21	298 ± 7	409 ± 11
Age	Gyr	0.76 ± 0.08	0.51 ± 0.02	0.61 ± 0.06	0.80 ± 0.10	1.35 ± 0.08	1.39 ± 0.16
τ	Myr	95 ± 35	<43	<71	103 ± 24	181 ± 17	178 ± 27
A_V	mag	0.61 ± 0.09	0.18 ± 0.06	0.34 ± 0.07	0.33 ± 0.09	0.35 ± 0.05	0.11 ± 0.05
[Z/H]	...	<0.33	0.02 ± 0.03	-0.03 ± 0.13	0.16 ± 0.13	0.01 ± 0.04	0.25 ± 0.09
log sSFR	yr^{-1}	<-10.72	<-12.80	<-11.77	-11.18 ± 0.54	-11.28 ± 0.11	-11.44 ± 0.22
log μM_*	M$_\odot$	12.06 ± 0.04	11.72 ± 0.02	11.74 ± 0.03	11.62 ± 0.05	12.77 ± 0.03	12.56 ± 0.04
μ SFR	M$_\odot$ yr^{-1}	<22	<1	<1	3 ± 2	31 ± 9	13 ± 3
log M_*	M$_\odot$	11.50 ± 0.17	...	...	10.96 ± 0.10	11.69 ± 0.19	11.68 ± 0.19
SFR	M$_\odot$ yr^{-1}	<6.1	...	...	0.6 ± 0.4	2.6 ± 1.4	1.8 ± 0.9
				Emission Line Properties			
$\sigma_{\rm em}$	km s^{-1}	213 ± 20	190 ± 49	345 ± 55	364 ± 22	...	...
$v_{\rm em} - v_{\rm stars}$	km s^{-1}	-20 ± 24	27 ± 37	40 ± 41	-12 ± 30	...	...
Hα EW	Å	4.3 ± 1.2	0.3 ± 0.5	4.2 ± 0.9	2.6 ± 0.9	...	...
[N II] $\lambda6585$ EW	Å	24.5 ± 2.3	3.7 ± 0.6	11.1 ± 1.8	17.1 ± 1.1	...	...
[O III] $\lambda5008$ EW	Å	...	0.1 ± 0.2	2.0 ± 0.3	5.3 ± 0.5	...	...
[O II] EW	Å	...	<3.5[a]	5.1 ± 0.9	2.5 ± 1.3	...	-0.1 ± 0.6
[S II] EW	Å	...	...	...	<3.6[a]	...	...
log [N II]/Hα	...	0.74 ± 0.12	>0.33	0.40 ± 0.11	0.79 ± 0.16	...	...
log [O III]/Hβ	...	...	...	$0.28^{+0.32}_{-0.11}$	$0.86^{+0.34}_{-0.15}$	...	...
			Line Fluxes Uncorrected for Magnification or Extinction:				
Hα flux	10^{-18} cgs	73 ± 20	5 ± 7	51 ± 11	30 ± 10	...	...
[N II] $\lambda6585$ flux	10^{-18} cgs	391 ± 38	54 ± 9	129 ± 21	183 ± 13	...	...
[O III] $\lambda5008$ flux	10^{-18} cgs	...	2 ± 4	31 ± 5	68 ± 6	...	...
[O II] flux	10^{-18} cgs	...	<44[a]	42 ± 7	17 ± 9	...	-3 ± 27
[S II] flux	10^{-18} cgs	...	...	...	<39[a]	...	...

Notes. BC03 models and a Chabrier (2003) IMF are assumed. Uncertainties are statistical and do not include model systematics (see text). All upper limits are 2σ. The flux magnification factor is μ; hence, log M_* is demagnified and log μM_* is not. EWs are reported in the rest frame. Line flux units are 10^{-18} erg cm^{-2} s^{-1} Å^{-1}. The [O II]$\lambda\lambda3727,3730$ and [S II]$\lambda\lambda6718,6733$ fluxes and EWs refer to the sum of the two doublet components. The spectra and derived extrinsic quantities, including line fluxes, are scaled to match the total photometric magnitudes.

[a] These lines were formally detected at marginal significance, but since inspection of the spectra showed these may be spurious, we have chosen to quote only upper limits.

Inspection of the spectra immediately shows a diversity of ages from 0.5–1.4 Gyr. Coupled with the low current sSFRs, this implies a substantial decline in star formation rate over the past Gyr with e-folding times of $\tau \lesssim$ 100–200 Myr. A mild amount of reddening, corresponding to $A_V = 0.1$–0.6, is inferred for the full sample. We find stellar abundances near solar. For the most massive and oldest galaxy, MRG-M0138, the FIRE spectrum indicates $Z \simeq 2Z_\odot$. We will analyze the detailed chemical abundance pattern of this galaxy in a forthcoming paper.

Although the stellar continuum suggests only mild reddening, mid- and far-infrared observations are needed to address the possibility of highly extinguished star-forming regions. Stacking analyses have shown that quiescent galaxies at $z < 2.5$ identified by the *UVJ* diagram or similar criteria usually do *not* harbor much obscured star formation (Fumagalli et al. 2014; Man et al. 2016). For the galaxies in our sample, the only existing observation with adequate depth to address this question is a *Spitzer* MIPS 24 μm image that covers MRG-M2129 (PI: M. Yun, Program ID 50610). MRG-M2129 is not detected with a 2σ upper limit of 86 μJy, corresponding to $\mu L_{\rm IR} < 8 \times 10^{11}$ $L_\odot$ assuming the Wuyts

et al. (2008) template. In the limit where the dust is heated only by star formation, this limit implies SFR < 18 $M_\odot$ yr^{-1} (corrected for magnification).[12] This would still place the galaxy below the main sequence of star-forming galaxies, given its high mass; furthermore, this limit is conservative because the emission lines in MRG-M2129 show evidence for an active galactic nucleus (AGN; see Section 7.2) that could contribute to $L_{\rm IR}$.

6.4. Systematic Uncertainties and Robustness Tests

The formal uncertainties in Table 5 are usually very small, which is expected given the high quality of the data and the simplicity of the star formation history and other aspects of the model. We emphasize that the listed uncertainties are purely statistical. Systematic errors in the models certainly dominate (e.g., Muzzin et al. 2009). As a rough estimate of these systematic errors, we also analyzed the data using the FSPS

[12] We used a circular aperture with a radius of 3″5, a sky annulus extending from 7″–10″, and an aperture correction from the MIPS Instrument Handbook. Toft et al. (2017) find a more stringent 3σ limit of SFR < 5 $M_\odot$ yr^{-1} apparently due to differences in the estimated noise or aperture correction.
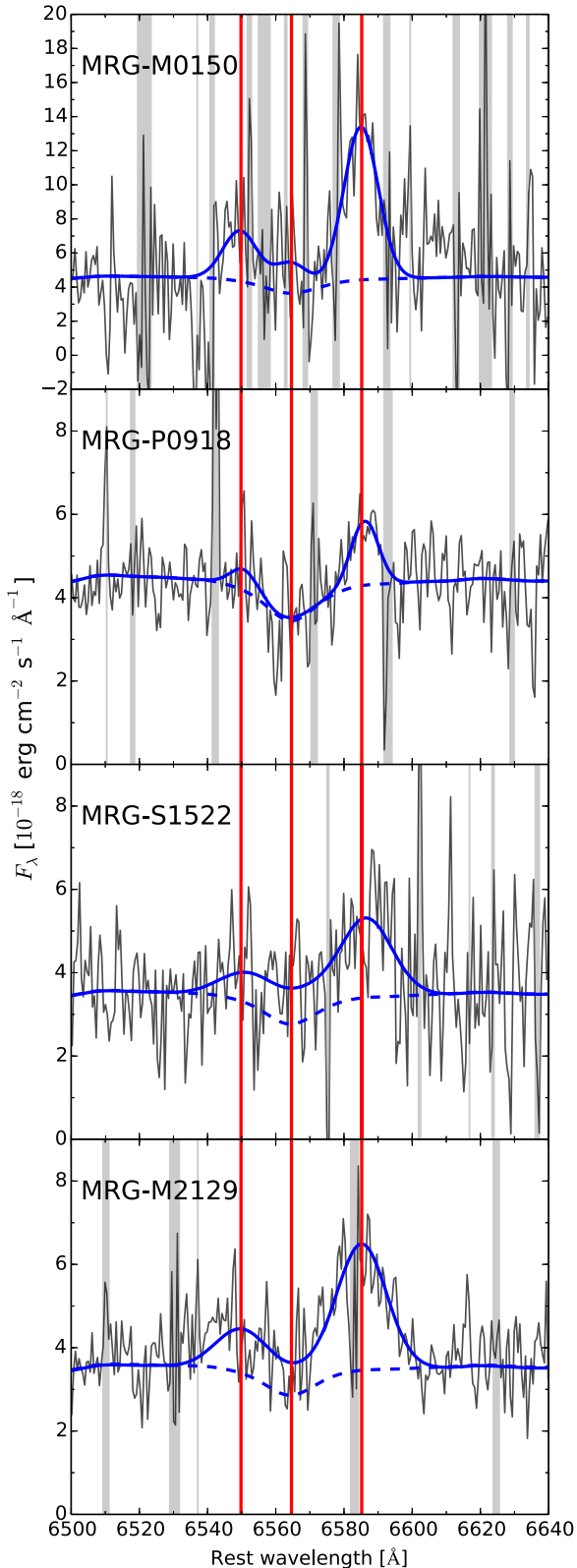
Figure 10. Integrated spectra in the [N II]+Hα region binned to 25 km s^{-1} pixel^{-1}. Solid blue lines show the best-fit models, including the stellar continuum and line emission. Dashed blue lines include only the stellar continuum. Vertical red lines show the positions of the [N II]λλ6550,6585 and Hα lines at the redshift of the stars. Gray bars indicate regions of bright sky emission.

v3.0 (Conroy et al. 2009; Conroy & Gunn 2010) stellar population synthesis models using the MIST isochrones (Choi et al. 2016). We find systematic differences of $t_{FSPS} - t_{BC03} =$

$+0.2$ Gyr in age, $+50$ Myr in τ, $+0.1$ mag in A_V, $+0.2$ dex in sSFR, and $+0.14$ dex in M_*, and $+0.3$ Å in emission line EWs. In [Z/H] the differences are not systematic, but there is a scatter of 0.2 dex, indicating that the metallicity constraints are the least robust.

For MRG-M0138, we have analyzed spectra of both Image 1 and 2, which provides a consistency test. The same colors were used in both fits, but we rescaled the flux level to match Image 2 when fitting the FIRE spectrum. Because MRG-M0138 is rotating (Paper II) and the slit cuts through the source at very different angles with respect to the major axis for Images 1 and 2, we do not expect to measure the same velocity dispersion σ in the two spectra. However, the stellar population parameters are reasonably consistent.

7. Measurement and Interpretation of Emission Lines

We detected emission lines in all of the lensed galaxies except MRG-M0138 (in that case, Hα, [N II], and [O III] are all contaminated by telluric absorption). In all cases the strongest line is [N II]. Hα is much weaker, with rest-frame equivalent widths (EWs) of $\lesssim 4$ Å. As seen in Figure 10, we detect this weak Hα emission from the infilling of the stellar absorption in all cases except MRG-P0918, for which we are only able to place an upper limit. The ratio [N II] $\lambda6585$/Hα ranges from 2–6. Figure 11 shows our sample in the BPT diagram (Baldwin et al. 1981). Comparing to the Kewley et al. 2001 extreme starburst line (magenta), we find that the [N II]/Hα ratios are all well in excess of those producible from photoionization by massive stars. Therefore, the nearly ubiquitous line emission in this sample is not indicative of low-level star formation.[13] In this section, we consider the robustness of the emission line measurements and explore the possible origins of the emission.

7.1. Robustness of Emission Line Measurements

It is clear that the stellar absorption correction significantly affects the inferred Hα emission. Are the high inferred [N II]/Hα ratios reliable? To test this, we fit only the [N II]+Hα region of the spectrum with the continuum fixed to a simple stellar population with an age of 400 Myr. Since the Hα absorption is maximum around this age, the test supplies an upper limit to the Hα flux and thus a lower limit to [N II]/Hα. Due to the weak sensitivity of the Hα absorption to age in the relevant range of ages, [N II]/Hα never shifted by more than half of its random uncertainty in this test. Systematic shifts in the line ratios that arise when using the FSPS versus BC03 stellar models are even smaller ($\simeq 0.02$ dex), with the exception that the lower limit on log [N II]/Hα is weakened from >0.54 to >0.33 for MRG-P0918 when using the FSPS models. We conservatively adopted this weaker constraint in Table 5. In all other cases, the [N II]/Hα ratio is robust to the stellar continuum modeling details.

The [O III]/Hβ ratio is inferred less directly. Since the Hβ and higher-order emission is much weaker than the stellar absorption, the intensity of the Balmer emission lines in the fit is driven by Hα infilling, and we cannot constrain the Balmer decrement Hβ/Hα. Our spectrum model assumes that the nebular lines and the stars are equally attenuated. If the line emission is more

---

[13] Even if the weak Hα emission were interpreted as arising from star formation, assuming that its attenuation is similar to the starlight, the inferred SFRs would generally be consistent with limits from the stellar continuum modeling: μSFR $= 28 \pm 8$, 3 ± 2, 14 ± 3, and $5 \pm 2 M_\odot$ yr^{-1} for MRG-M0150, MRG-P0918, MRG-S1522, and MRG-M2129, respectively.
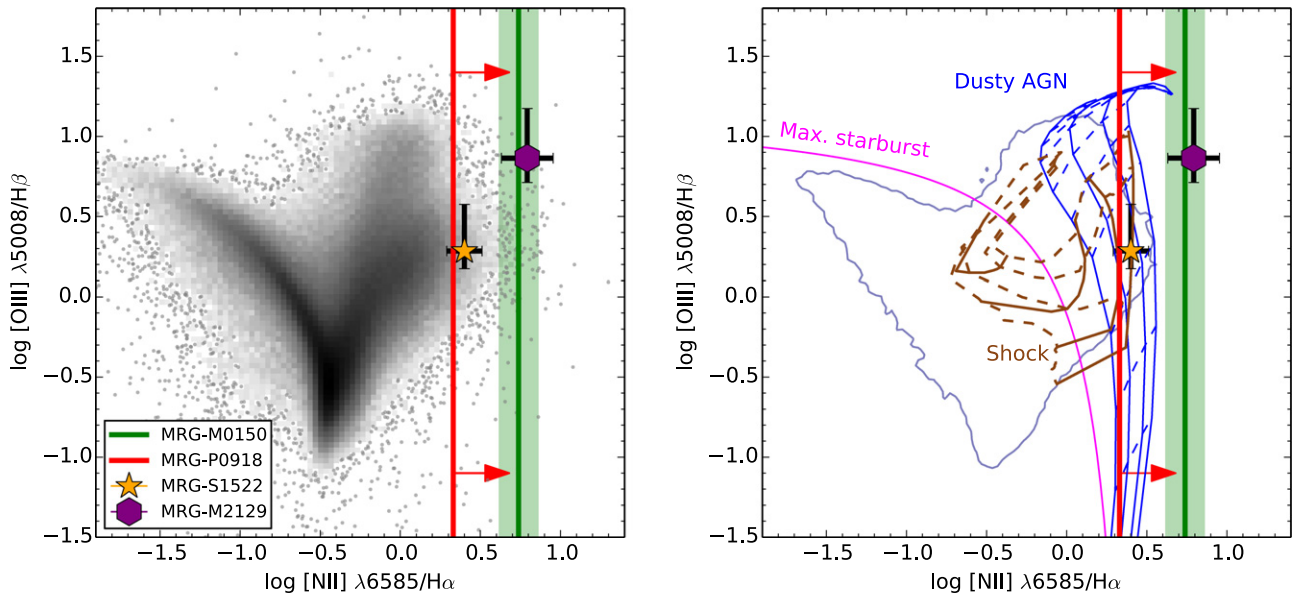
Figure 11. Left: the BPT diagram showing the lensed quiescent galaxies with detected emission lines. Note that there is no constraint on [O III]/Hβ for MRG-P0918 and MRG-M0150, and only a lower limit on [N II]/Hα for the former. The gray shading shows the location of local galaxies in the SDSS selected from the MPA-JHU catalog (Brinchmann et al. 2004) with $>2\sigma$ detections in all lines. Individual gray points are shown where the density is low. Right: the lensed quiescent galaxies are compared to several theoretical models. The magenta line is the Kewley et al. (2001) theoretical maximum starburst curve. The blue lines show a grid of dusty AGN models (Groves et al. 2004a, 2004b) with $4Z_\odot$ metallicity, $n = 10^4$ cm^{-3}, $\alpha = -2$ to -1.2, and $\log U = -4$ to 0. The brown lines show a grid of shock models (Allen et al. 2008) without precursor emission, $2Z_\odot$ metallicity, $n = 1$ cm^{-3}, $B = 10^{-4}$ to 10 μG, and velocities 100–1000 km s^{-1}. The grids were generated using `itera` (Groves & Allen 2010). The gray contour outlines the SDSS locus from the left panel for reference.

extinguished, then this would push the inferred [O III]/Hβ upward. Reddy et al. (2015) found that the attenuation of the Hα photons can exceed that of the stellar continuum by up to 1.5 mag in high-z star-forming galaxies. We think that such a large difference is unlikely to hold in quiescent galaxies where attenuation is mild, but assuming 1.5 mag of differential attenuation as a limiting case, we would infer [O III]/Hβ typically 0.3 dex higher. The asymmetric error bars in Figure 11 and Table 5 include this uncertainty added in quadrature.

7.2. Origins of Line Emission

Determining the excitation mechanisms in the Seyfert/LIER[14] region of the BPT diagram is difficult, even in local galaxies with much more detailed information. Although we cannot expect to definitively identify the excitation mechanisms in every case, we can consider which scenarios are consistent with the available evidence. The ratio [N II]/Hα $\simeq$ 2–6 is high throughout our sample, but the strength of the oxygen lines and the spatial distribution and kinematics of the gas relative to the stars provide additional constraints and show more diversity.

The spatial variation of emission line EWs is shown in Figure 12. These measurements were made from spectra extracted in bins along each image. (These bins are the same that are used to extract stellar kinematics in Paper II, where further details can be found.) We then modeled each spectrum as described in Section 6.2, except that we did not include photometric constraints, which are not necessary to measure the emission line EWs.

Photoionization by an AGN is one way to produce a high [N II]/Hα ratio. In this scenario, we expect high [O III]/Hβ

and [O III]/[O II] ratios and a negative gradient in the emission line EWs (i.e., centrally concentrated emission). MRG-M2129 is the only galaxy that fits these criteria. Its high values of $\log$ [O III]/Hβ $= 0.86 \pm$ 0.15 and $\log$ [O III]/[O II] $= 0.5 \pm$ 0.2 are characteristic of Seyfert rather than LIER-type emission (Veilleux & Osterbrock 1987; Kewley et al. 2006).[15] The emitting gas extends out to radii of at least 2 kpc, with a declining [O III] EW that is qualitatively consistent with photoionization by a central source. No other emission line in any of the galaxies in our sample shows $>2\sigma$ evidence for radial variation in Figure 12. A compact central source is detected in the *HST* images of MRG-M2129 (Section 5.3), which could be a Seyfert nucleus. All of these lines of evidence point toward photoionization by an AGN as one component of the excitation. Toft et al. (2017) came to similar conclusions on MRG-M2129 with some quantitative differences, which we discuss in the Appendix.

Although present in MRG-M2129, AGN photoionization is likely to not be the dominant source of excitation in MRG-S1522 or MRG-P0918. The lower [O III]/Hβ ratio of MRG-S1522 is more consistent with a LIER than a Seyfert classification in the BPT diagram, although the measurement uncertainties make this ambiguous. Confirmation comes from its stronger [O II] emission than [O III], as seen in LIER-type spectra, with $\log$ [O III]/[O II] $= -0.2 \pm 0.1$. In MRG-P0918, we do not detect [O III] to sensitive limits. Although we cannot constrain the ratio [O III]/Hβ and place this galaxy on the BPT diagram, the low EW of [O III] and its weakness compared to [N II] (with a ratio $\lesssim 0.1$) are not typical of Seyfert-type spectra. For MRG-M0150, the only strong emission lines that are observable from the ground are [N II] and Hα; so, for more

---

[14] LIERs are low-excitation emission line regions. Following Belfiore et al. (2016) and others, we drop the "nuclear" designation of the more traditional LINER term since the emission is not confined to the central regions.

---

[15] Our measurements of [O III]/[O II] assume that the emission line and stellar light are equally attenuated. We caution that the [O II] emission in MRG-M2129 is very weak and lies in a spectral region with significant residuals. Nonetheless, at a minimum we can exclude that [O II] is stronger than [O III] as in LIERs.
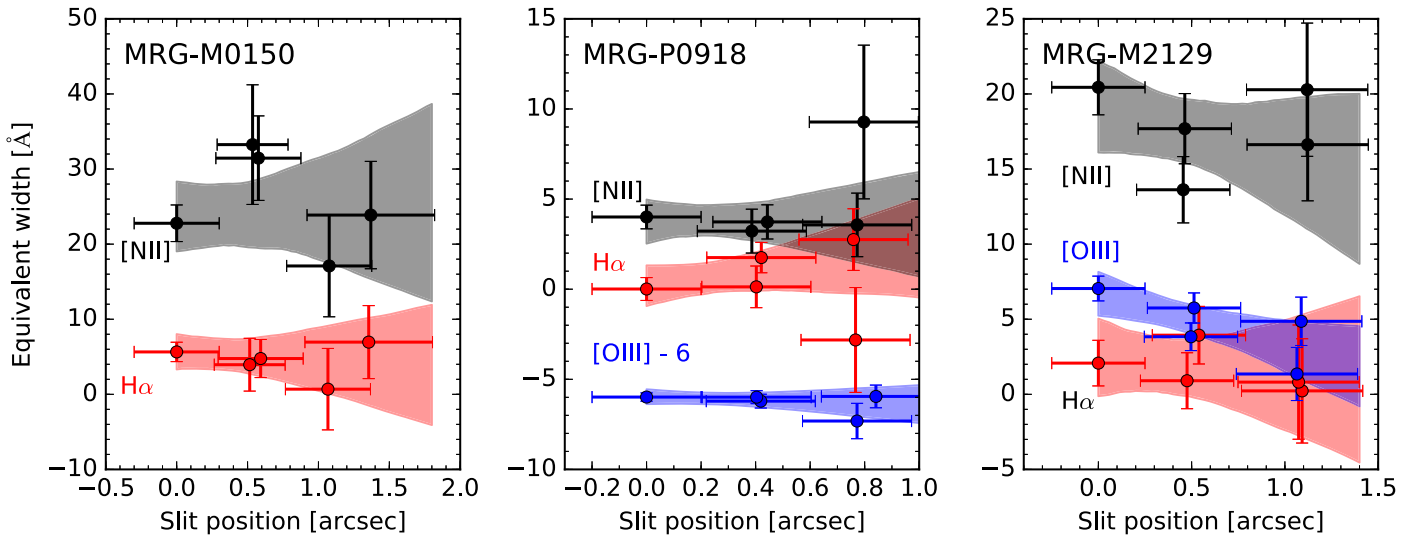
Figure 12. Spatially resolved emission line rest-frame EWs as a function of distance along the slit from the center of the image. Measurements on both sides of center are plotted. The horizontal error bars indicate the bin sizes. Points have been shifted horizontally by $\lesssim 0\rlap{.}''1$ for clarity, and the [O III] EW in the middle panel is shifted downward 6 Å. Colored bands show the 95% credible range for a linear fit. The only gradient with $>2\sigma$ significance is the decline in [O III] with radius in MRG-M2129.

information on this galaxy, we will have to await improved spectral coverage with *JWST*.

Photoionization of diffuse interstellar gas by hot evolved stars, such as post-asymptotic giant branch (post-AGB) stars, is often invoked to explain LIER emission in local early-type galaxies (e.g., Yan & Blanton 2012; Singh et al. 2013; Belfiore et al. 2016). At the ages relevant to quiescent galaxies at $z = 2$, i.e., $\simeq 0.5$–3 Gyr, stellar population synthesis models predict Hα EWs of $\simeq 0.1$–0.8 Å (Cid Fernandes et al. 2011, Figure 2). For MRG-M0150, MRG-S1522, and MRG-2129, we observe Hα EWs in the range 2.6–4.3 Å, which means there is insufficient ionizing flux from post-AGB stars to explain most of the emission. In MRG-P0918, however, the lower Hα EW of 0.3 ± 0.5 Å is consistent with expectations for post-AGB stars. Furthermore, in this scenario the gas should share the distribution and kinematics of the stars, which we observe. The gaseous and stellar velocity dispersions are consistent (Table 5) and the emission line EWs are essentially constant (Figure 12). The main difficulty with this interpretation is that photoionization models of 3–13 Gyr old populations produce log [N II]/Hα $\lesssim 0.1$ (Binette et al. 1994; Byler et al. 2017), which is smaller than observed in MRG-P0918. Since models at the relevant age of 0.5 Gyr have not been explored and probably have uncertainties in the shape of the ionizing spectrum, we consider that post-AGB stars may still be a viable explanation for most of the line emission in MRG-P0918 but not for the other galaxies in our sample.

Shocked gas is another possible source of high-[N II]/Hα emission with low ionization. Shocks are consistent with the lower values of [O III]/Hβ and [O III]/[O II] seen in MRG-S1522. Supporting evidence comes from the high line width of the ionized gas ($\sigma = 345$ km s^{-1}) relative to the stars (241 km s^{-1}), which indicates outflowing or turbulent ionized gas. We think that shocks are likely to contribute to the line emission in MRG-M2129 as well, since the line width of the gas is also elevated well above that of the stars (364 versus 266 km s^{-1}). Such a situation could in general be explained by a differing spatial distribution of the gas and stars. However, in the case of MRG-M2129, the EW of the [N II] emission does

not change much across the image (Figure 12), and the elevated $\sigma_{\rm gas}$ is not confined to the nucleus but extends to $R \simeq 2$ kpc. This suggests that both AGN photoionization and shocks likely contribute to the line emission in MRG-M2129.

We have explored whether the Allen et al. (2008) shock models can reproduce the line ratios in Figure 11. For MRG-S1522, we find that a model with twice solar metallicity, $n = 1$ cm^{-3}, and $B = 2 \mu$G can reproduce the [N II]/Hα and [O III]/Hβ ratios for shock velocities $v \gtrsim 400$ km, comparable to $\sigma_{\rm gas}$.[16] However, the agreement only holds if we consider the shock emission and not the radiative precursor, which may be an unphysical scenario. The emitting area required to produce the observed Hα luminosity is $\lesssim 100 \, (\mu/5)^{-1}$ kpc^{-2}, which corresponds to a spherical radius $\lesssim 3 \, (\mu/5)^{-1/2}$ kpc. Since this is comparable to the galaxy size, it requires a large-scale shock, but the energetics are feasible. On the other hand, the high [N II]/Hα ratios of MRG-M2129 and MRG-0150 cannot be reproduced by any of the Allen et al. (2008) shock models or the Groves et al. (2004a, 2004b) AGN models. This is demonstrated by the grids shown in Figure 11, where we have chosen parameters that maximize the [N II]/Hα ratio. Extracting more detailed physical information about these intriguing systems may require further development in these models.

In summary, four of the five lensed galaxies in our sample show line emission with very high [N II]/Hα ratios that are inconsistent with a star formation origin. (For the fifth, MRG-M0138, all of the strong nebular lines except [O II] are inaccessible from the ground.) MRG-M2129 shows clear evidence of AGN photoionization. Based on emission line ratios and the distribution and kinematics of the ionized gas, we argue that shocked gas is present in MRG-S1522 and MRG-M2129. Shocks may well power the line emission in MRG-M0150, but we lack the information needed to distinguish AGN photoionization. They could also contribute in MRG-P0918, but post-AGB stars might instead be the main ionizing source in that system. The high incidence of line emission in galaxies seen soon (<800 Myr) after quenching,

---

[16] We note that $\sigma_{\rm gas}$ includes unresolved rotation and, therefore, does not measure purely turbulent motion.
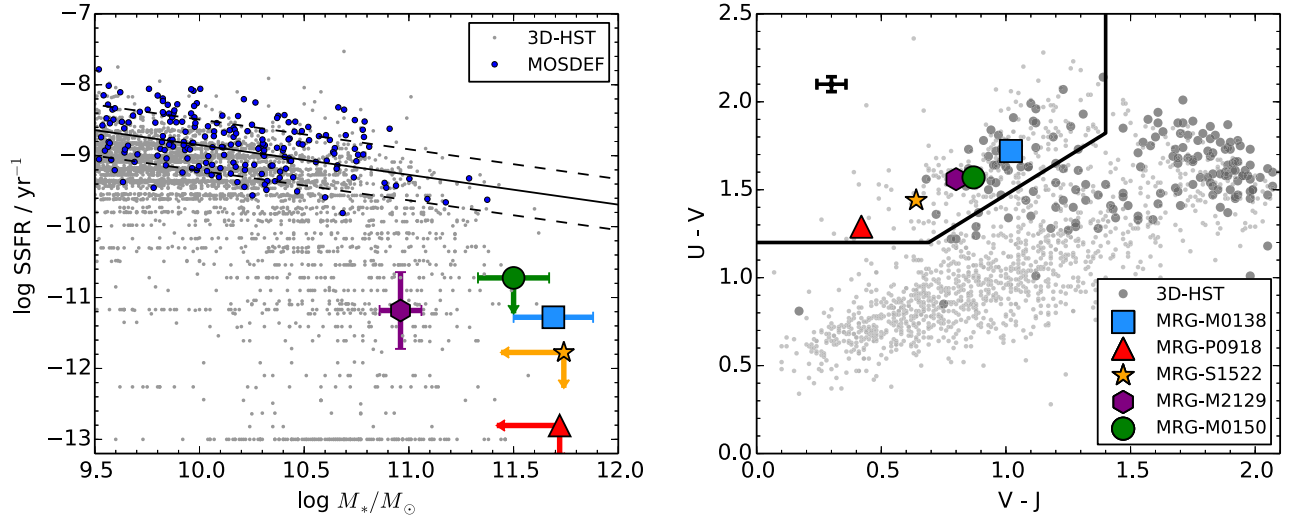
Figure 13. Left: the stellar mass–sSFR relation (the "main sequence") defined by (a) a stellar mass-selected sample at $z = 2$–2.6 from the 3D-*HST* catalogs (Skelton et al. 2014), where the SFRs are derived from SED fitting, and (b) Hα-based SFRs, corrected for extinction, for galaxies in the MOSDEF survey (Shivaei et al. 2015) at $z = 2.09$–2.61. The linear fit and 1σ scatter for the MOSDEF sample is shown. The lensed galaxies (see legend in right panel) fall $\gtrsim 1.5$ dex below the "main sequence" of star-forming galaxies. The stellar masses of lensed galaxies with unknown magnifications are plotted as upper limits. Right: the rest-frame *UVJ* colors of galaxies in the 3D-*HST* fields at $z = 2$–2.6 are compared to those of the lensed sample. Smaller and larger gray symbols denote 3D-*HST* galaxies having $10^{10} M_\odot < M_* < 10^{11} M_\odot$ and $M_* > 10^{11} M_\odot$, respectively. Colors of the lensed galaxies are determined by integrating the fits in Figure 9; representative uncertainties are indicated by the black error bars. The black line separates the star-forming and quiescent regions defined by Whitaker et al. (2011). The lensed sample spans the quiescent sequence, except for the reddest cases.
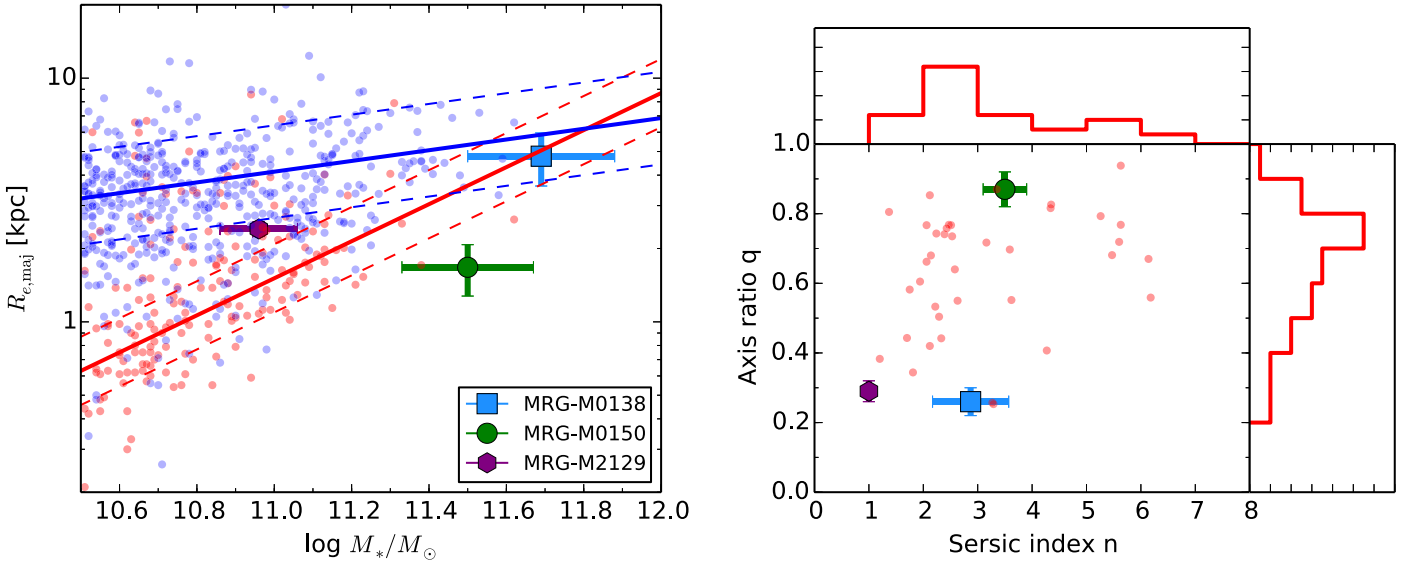
Figure 14. Left: the stellar masses and sizes of the lensed quiescent galaxies with lens models are compared to the relations defined by star-forming (blue) and quiescent (red) galaxies at $z = 2$–2.6 in the 3D-*HST* fields, as classified by the *UVJ* criterion (Whitaker et al. 2011) and measured by van der Wel et al. (2014). Linear fits to the relation at $z = 2.25$ are overlaid, with dashed lines indicating the 1σ scatter. Right: Sérsic indices n and projected axis ratios $q = b/a$ for the lensed quiescent sample (see legend in left panel) are compared to those of similarly massive ($M_* > 10^{11} M_\odot$) quiescent galaxies at $z = 2$–2.6 in the 3D-*HST* fields. Histograms show the marginal distributions of the 3D-*HST* sample. Measurements from our single-Sérsic models are shown for the lensed galaxies. The lensed sample is reasonably representative of coeval quiescent galaxies in their sizes and Sérsic indices, but high ellipticities (low q) may be over-represented.

the indications of shocked gas in many (and potentially all) cases, and the lack of ongoing star formation are all consistent with the idea that AGN-driven outflows or turbulence could play an important role in maintaining the quiescence of these galaxies. We will discuss these implications further in Section 8.3.

8. Discussion

We have discovered a sample of five quiescent galaxies at $z = 2$–2.6 that are significantly magnified by galaxy

clusters. These galaxies are extraordinarily bright in the near-infrared—$H_{AB} < 20$ in all cases—due to their high stellar masses ($M_* \gtrsim 10^{11} M_\odot$ for the three galaxies with lens models) combined with lensing magnification by factors of $\mu \approx 4$–13. Observations with the Magellan/FIRE and Keck/MOSFIRE spectrographs confirmed the redshifts and evolved stellar populations (ages 0.5–1.4 Gyr) for the full sample. The integrated spectra are among the most detailed yet obtained for quiescent galaxies at these redshifts, particularly for the remarkable case of MRG-M0138, which is the NIR-brightest lensed distant galaxy yet discovered. In addition to the

flux amplification, lensing affords the possibility of spatially resolving the stellar continuum with ground-based NIR spectrographs and measuring their internal kinematics and stellar population gradients. The lensed galaxies in the present sample are the only quiescent systems at $z > 2$ for which this is currently practical, making them unique and valuable resources.

8.1. Assessing the Representative Nature of the Lensed Galaxy Sample

Further papers will present spatially resolved measurements of the stellar kinematics and populations. In order to interpret these measurements, we must be able to place this sample in the context of the full galaxy population. Here we compare our lensed galaxies to coeval samples selected from deep field surveys.

Figure 13 (left panel) compares the sample to the star formation "main sequence" derived in the 3D-*HST* and MOSDEF surveys. The lensed sample falls $\gtrsim 1.5$ dex below the main sequence. The right panel of Figure 13 shows that all galaxies in our sample fall in the quiescent region of the *UVJ* diagram. Consistent with the wide range of ages spanning 0.5–1.4 Gyr that we measured from the spectra and photometry (Section 6), our sample spans nearly the full range of the *UVJ* quiescent sequence. Only the reddest and presumably oldest galaxies are not represented.

In Figure 14, the structural properties of the three galaxies with a lens model are compared to coeval quiescent galaxies in the 3D-*HST* survey. The left panel shows that the lensed galaxies scatter around the mass–size relation and so can be considered typically "compact." The right panel compares the axis ratios and Sérsic indices of our sample to those of similarly massive, coeval quiescent galaxies. The lensed sample displays Sérsic indices of $n \simeq 1$–4 and is consistent with being drawn from the 3D-*HST* distribution. On the other hand, the axis ratios of the lensed galaxies are at the round and flat extremes of the 3D-*HST* distribution.

Two aspects of the sample that are perhaps surprising are the high ellipticities of two systems (MRG-M0138 and MRG-M2129) and the high stellar masses ($\log M_* = 11.69 \pm 0.19$ for MRG-M0138 and $\log M_* = 11.50 \pm 0.17$ for MRG-M0150) of two others. Considering first the ellipticities, we could suppose that errors in the lens model could lead to spuriously high values. However, we found that the ellipticity is robustly recovered across three lens models for MRG-M2129. In Paper II, we will show that their stellar kinematics imply that both MRG-M0138 and MRG-M2129 are intrinsically flat. These comparisons suggest that the high ellipticities are genuine. One might suppose that a selection effect could enhance their probability of inclusion in our sample. This might be the case if images were selected for follow-up based on their length-to-width ratio, for example, but our follow-up observations were based on a complete color–magnitude–selected sample in the central regions of the galaxy clusters in our survey. Given the small number of objects, it may be that high ellipticities are simply randomly over-represented in the three lensed galaxies with lens models. Regardless of its origin, this over-representation is important to bear in mind for our kinematic analysis in Paper II.

Turning to the remarkably high stellar masses inferred for MRG-M0150 and MRG-M0138, we note that since we are studying galaxies on the bright tail of the luminosity function, we expect the lensed galaxies to pile up near the flux limit of our selection box. This is generally the case for the *H* band flux (Figure 1), with the exception of the ultra-bright MRG-M0138. In Paper II, we will show that its dynamical mass agrees with the stellar mass inferred in this paper. Since the stellar and dynamical masses have different dependences on the lens mapping, this consistency provides some reassurance that the magnification is not very far in error, i.e., well beyond the ~ 0.2 dex uncertainty we estimated in Section 5.1.2. For MRG-M0150, we will show in Paper II that the dynamical mass is $\simeq 0.2$ dex smaller than the stellar mass. This could indicate that the magnification is underestimated and the stellar mass is overestimated, but the difference is within the estimated uncertainties and, in any case, still implies a very massive galaxy.

We conclude that the sample is broadly representative of the colors and sizes of massive quiescent galaxies at $z = 2$–2.6. The stellar masses and ellipticities are likely to be accurate within our estimated uncertainties, but the subset of three galaxies with lens models has a high proportion of high-ellipticity galaxies compared to unlensed samples and so is not fully representative in that property. Future lens models with additional constraints from any newly identified multiple image systems would be useful to help validate our magnification estimates.

8.2. Disky Quiescent Galaxies and Evolution to z = 0

The structures of the galaxies in our sample are very different from typical early-type galaxies in the local universe with a similar mass. Their sizes are comparable to other $z = 2$ quiescent galaxies (Figure 14) and are smaller than $z = 0$ galaxies of equal mass. Even more striking, however, is the presence of a dominant disk component in two of the three galaxies for which we have reconstructed the source, MRG-M0138 and MRG-M2129. Both galaxies have ellipticities $e > 0.7$. To place these shapes in the context of low-redshift galaxies, we selected galaxies with $\log M_* > 11$ and $z = 0.05$–0.25 from the GAMA survey (Driver et al. 2009; Taylor et al. 2011; Kelvin et al. 2012). Only 4% of the local galaxies have ellipticities $e > 0.7$. Even more rare is the apparent lack of a bulge in MRG-M2129, which is well-described by a pure exponential disk ($n = 1$), as also remarked by Toft et al. (2017). In our local comparison sample, only 0.9% of galaxies have $n \leqslant 1$.

Therefore, these galaxies need to evolve in size, shape, and concentration to resemble their $z \sim 0$ descendants. The extreme differences in structure imply that relatively little of the total transformation from star-forming galaxies into local early-type systems occurred when these galaxies were quenched. Instead, these changes must have occurred later, likely through a series of major and minor mergers. The structures of the galaxies in our sample also have implications for their earlier evolution. In particular, whatever processes quenched star formation in these galaxies did not destroy the stellar disk or, in the case of MRG-M2129, even produce a significant bulge. In Paper II, we will discuss the past and future evolution of these galaxies in greater detail by bringing additional kinematic evidence to bear.

8.3. The Nature of Emission Lines in $z \sim 2$ Quiescent Galaxies

The nearly ubiquitous presence of line emission with high [N II]/Hα ratios in our sample of massive, quiescent galaxies is striking. While massive ($M_* \gtrsim 10^{10.9}\, M_\odot$) star-forming galaxies at $z = 1$–3 commonly host nuclear outflows that are thought to be driven by AGN (Genzel et al. 2014), less is known about quiescent galaxies in this redshift range. Belli et al. (2017a) studied the incidence of line emission in $z = 0.7$–2.7 galaxies in the KMOS^{3D} survey. They detected unambiguous line emission in 17% of the quiescent galaxies (as classified by UVJ colors) and in about half of these attributed the emission to shocks. It is interesting that in our sample the incidence is much higher—emission lines are detected in all four galaxies for which the strong lines are observable from the ground—and that our sample reaches [N II]/H$\alpha \simeq 6$, whereas values above ~ 1 are absent from the Belli et al. sample.

These differences might be explained in part by differences in the sensitivities of the observations and in the stellar masses characterizing each sample. Belli et al. found a trend for galaxies with lower Hα EWs and higher stellar masses to have higher [N II]/Hα ratios. Since the lensed galaxies in our sample have, on average, lower Hα EWs and higher masses than the KMOS^{3D} quiescent galaxies, it is possible that future observations will show that the KMOS^{3D} and lensed galaxies sample different parts of a common sequence. We emphasize that we can characterize the weak emission lines in our sample of quiescent galaxies only because of the galaxies' brightness (due to their high masses combined with lensing magnification) and our accurate modeling of the stellar continuum. With shallower data, we could not measure the infilling of the stellar Hα absorption. If a stellar redshift could not be measured, then one could easily mistake the stronger [N II] $\lambda 6585$ line for Hα, leading to a completely different interpretation.

Shocks are most clearly present in two cases (MRG-S1522 and MRG-M2129), as judged by the line ratios, gas kinematics, and in the case of MRG-M2129, the extended distribution of the [N II] emission. Although shocks could also be present in MRG-P0918 and MRG-M0150, other explanations remain viable in these cases because of the fewer constraints on MRG-M0150 and the low Hα EW seen in MRG-P0918 (see Section 7.2). Given that the current rates of star formation are very low in MRG-S1522 and MRG-M2129 ($\lesssim 1\, M_\odot$ yr^{-1}; Table 5), sources of energy are needed other than massive stars and core-collapse supernovae. Type Ia supernovae occurring at the expected rate (Maoz et al. 2012), which is much higher than in local massive ellipticals, can supply kinetic power comparable only to the Hα luminosities seen in our sample, which is expected to be $\lesssim 2\%$ of the shock luminosity (Allen et al. 2008). Therefore, shocks powered by type Ia supernovae ejecta cannot generate the observed line emission. Mergers could produce shocks, but we do not find evidence for ongoing mergers in our sample except for a possible faint companion of MRG-M2129, and this explanation seems inconsistent with the spatial uniformity of the [N II] emission. Therefore, AGNs provide the most obvious energy source.

Observations of some low-redshift quiescent galaxies have revealed significant reservoirs of turbulent molecular gas with $\sim$kpc (or larger) sizes and suppressed star formation relative to the Kennicutt–Schmidt relation, often combined with warm shocked gas (e.g., Alatalo et al. 2015; Guillard et al. 2015; Lanz et al. 20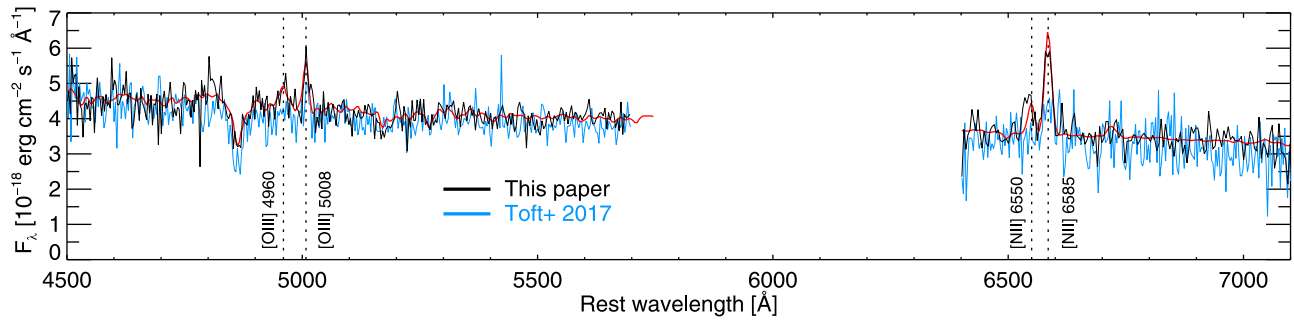16). A significant fraction of intermediate-mass quiescent galaxies in the local universe appear to have ionized gas outflows (Cheung et al. 2016). The injection of turbulence into the interstellar medium via an AGN jet has been suggested to suppress star formation in these galaxies. Our observations are at least consistent with a similar mechanism acting in MRG-S1522 and MRG-M2129, and potentially in *all* of the galaxies in our sample that have ages $\lesssim 1$ Gyr.

The high fraction of massive systems in which an AGN seems to be affecting the gas—both in high-mass star-forming galaxies at $z \sim 2$ and in recently quenched galaxies—suggests than AGN activity might play an important role in quenching star formation and maintaining low star formation rates in these galaxies, although proving causality is very difficult. Our sample affords detailed observations but is small. More deep spectra of $z \gtrsim 2$ quiescent galaxies are needed to firmly establish the prevalence of emission lines and the variation of their properties with the time since quenching. Observations of the distribution and kinematics of any molecular gas in such galaxies may also be revealing.

8.4. Future Work

In four of the lensed quiescent galaxies discussed in this paper (i.e., all but MRG-S1522), we are able to spatially resolve the stellar continuum in our NIR spectra. In the companion Paper II, we present the resolved stellar kinematics of these galaxies. In a future paper, we plan to study their resolved stellar populations in order to dissect the star formation histories of early quiescent galaxies. We will also use observations of the extraordinarily bright MRG-M0138 to measure its multi-element stellar abundance pattern. Many of these measurements are currently only possible for quiescent galaxies that are gravitationally lensed. Observations with *JWST* will only marginally resolve a typical compact quiescent galaxy at $z \sim 2$. While such data will be very valuable, lensed quiescent galaxies will continue to offer the highest-resolution views of these galaxies and will provide unique insights into their formation, even in the *JWST* era.

Figure 15. Integrated H- and K-band spectra of MRG-M2129 from this paper (black, with the model from Figure 9 shown as the red curve) and from Toft et al. (2017, blue) binned to $\simeq 200$ km s^{-1} pixel^{-1}. Significant differences in the EWs of [O III] and [N II] are clearly visible.

the NASA/ESA *Hubble Space Telescope*, and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA) and the Canadian Astronomy Data Centre (CADC/NRC/CSA).

Appendix

Here we compare the stellar population, structure, and emission line properties that we derived for MRG-M2129 with those recently published by Toft et al. (2017). Despite using quite different formulations of the star formation history, we find stellar population parameters that are consistent with Toft et al. Comparing their Extended Data Table 1 to our Table 5, we find that the stellar mass and age are consistent and, although the centers of our posteriors favor a higher metallicity and lower dust attenuation than Toft et al., these too are consistent within the quoted uncertainties.

Our fiducial magnification from Monna et al. (2017) is $\mu = 4.5$, which is within 3% of the Toft et al. model. The magnification in the more extreme of the two Zitrin models (Section 5.3) is only 18% higher. As described by Toft et al., the magnification is relatively well-constrained in this system because it is located far from the cluster center where the mass distribution is dominated by the smooth dark matter halo.

Toft et al. fit a single Sérsic model to the source, so we compare their parameters to our single Sérsic+Gaussian model. Like Toft et al., we infer a nearly exponential ($n = 1$) surface brightness profile with a compact size. The effective radius $R_{e,\mathrm{maj}}$ and PA are consistent within $\simeq 6\%$ and 6°, respectively. The main difference is that our reconstruction gives a much flatter source with $b/a = 0.29 \pm 0.03$ compared to their $0.59^{+0.03}_{-0.09}$. In correspondence with Toft and co-workers, it was found that a likely explanation is the manner in which the PSF was treated in their analysis. Once a discrepancy in their treatment was corrected, they found a value $b/a \simeq 0.4$ that is much closer to our measurement (S. Toft et al. 2018, private communication).

Finally, Toft et al. find a relatively high [N II]/Hα ratio coupled with an emission line velocity dispersion in excess of the stars. Like us, they interpret this as possible evidence of AGN-driven turbulence or outflows. However, beyond this qualitative agreement, there are substantial quantitative differences. Toft et al. find log [N II]/H$\alpha = -0.06 \pm 0.10$ versus our 0.79 ± 0.16. Figure 15 shows that this difference arises from manifestly different emission line EWs at the level of the raw spectra. Although our spectrum is deeper, the differences

do not seem to be consistent with noise fluctuations, especially for [N II]. This suggests some systematic effect in the observations or data reduction.

The presence of the [N II] and [O III] doublets with the correct ratios rules out the stronger lines in our spectra as arising from a simple data reduction artifact. The extraction apertures that are used to produce the spectra in Figure 15 are nearly matched ($\pm 1\rlap{.}''6$ for this paper and $\pm 1\rlap{.}''4$ for Toft et al.). This fact, combined with the weak spatial variation in [N II] EW that we see along the arc (Figure 12), makes it hard to understand these differences in the spectra as arising either from target acquisition errors affecting one of the observations or from unintended self-subtraction of the galaxy wings during the data reduction. In addition, unlike Toft et al., who find that the emission lines are redshifted by 238 km s^{-1} relative to the stars, we find no such velocity offset in the integrated spectrum ($v_{\mathrm{em}} - v_{\mathrm{stars}} = -12 \pm 30$ km s^{-1}) and we also do not detect He II $\lambda 5413$. Future integral field observations are desirable to investigate the possible origins of the differences.

ORCID iDs

Andrew B. Newman ⓘ https://orcid.org/0000-0001-7769-8660
Sirio Belli ⓘ https://orcid.org/0000-0002-5615-6018
Richard S. Ellis ⓘ https://orcid.org/0000-0001-7782-7071

References

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Alatalo, K., Lacy, M., Lanz, L., et al. 2015, ApJ, 798, 31
Allam, S. S., Tucker, D. L., Lin, H., et al. 2007, ApJL, 662, L51
Allen, M. G., Groves, B. A., Dopita, M. A., Sutherland, R. S., & Kewley, L. J. 2008, ApJS, 178, 20
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2016, ApJL, 827, L32
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2017, ApJL, 851, L40
Barro, G., Trump, J. R., Koo, D. C., et al. 2014, ApJ, 795, 145
Bedregal, A. G., Scarlata, C., Henry, A. L., et al. 2013, ApJ, 778, 126
Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, MNRAS, 461, 3111
Belli, S., Genzel, R., Förster Schreiber, N. M., et al. 2017a, ApJL, 841, L6
Belli, S., Newman, A. B., & Ellis, R. S. 2015, ApJ, 799, 206
Belli, S., Newman, A. B., & Ellis, R. S. 2017b, ApJ, 834, 18
Binette, L., Magris, C. G., Stasińska, G., & Bruzual, A. G. 1994, A&A, 292, 13
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buitrago, F., Trujillo, I., Conselice, C. J., et al. 2008, ApJL, 687, L61
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Cabanac, R. A., Alard, C., Dantel-Fort, M., et al. 2007, A&A, 461, 813
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Chabrier, G. 2003, PASP, 115, 763
Chang, Y.-Y., van der Wel, A., Rix, H.-W., et al. 2013, ApJ, 762, 83
Cheung, E., Bundy, K., Cappellari, M., et al. 2016, Natur, 533, 504

Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Dahle, H., Aghanim, N., Guennou, L., et al. 2016, A&A, 590, L4
Damjanov, I., Abraham, R. G., Glazebrook, K., et al. 2011, ApJL, 739, L44
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, A&G, 50, 5.12
Ebeling, H., Barrett, E., Donovan, D., et al. 2007, ApJL, 661, L33
Ebeling, H., Edge, A. C., & Henry, J. P. 2001, ApJ, 553, 668
Ebeling, H., Stockmann, M., Richard, J., et al. 2018, ApJL, 852, L7
Elíasdóttir, Á., Limousin, M., Richard, J., et al. 2007, arXiv:0710.5636
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Franx, M., Labbé, I., Rudnick, G., et al. 2003, ApJL, 587, L79
Fumagalli, M., Franx, M., van Dokkum, P., et al. 2016, ApJ, 822, 1
Fumagalli, M., Labbé, I., Patel, S. G., et al. 2014, ApJ, 796, 35
Geier, S., Richard, J., Man, A. W. S., et al. 2013, ApJ, 777, 87
Genzel, R., Förster Schreiber, N. M., Rosario, D., et al. 2014, ApJ, 796, 7
Groves, B. A., & Allen, M. G. 2010, NewA, 15, 614
Groves, B. A., Dopita, M. A., & Sutherland, R. S. 2004a, ApJS, 153, 9
Groves, B. A., Dopita, M. A., & Sutherland, R. S. 2004b, ApJS, 153, 75
Guillard, P., Boulanger, F., Lehnert, M. D., et al. 2015, A&A, 574, A32
Hennawi, J. F., Gladders, M. D., Oguri, M., et al. 2008, AJ, 135, 664
Hill, A. R., Muzzin, A., Franx, M., & van de Sande, J. 2016, ApJ, 819, 74
Johnson, T. L., Sharon, K., Bayliss, M. B., et al. 2014, ApJ, 797, 48
Jones, T. A., Swinbank, A. M., Ellis, R. S., Richard, J., & Stark, D. P. 2010, MNRAS, 404, 1247
Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, NJPh, 9, 447
Kelson, D. D. 2003, PASP, 115, 688
Kelson, D. D., van Dokkum, P. G., Franx, M., Illingworth, G. D., & Fabricant, D. 1997, ApJL, 478, L13
Kelson, D. D., Williams, R. J., Dressler, A., et al. 2014, ApJ, 783, 110
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, MNRAS, 421, 1007
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961
Kneib, J. P., Mellier, Y., Fort, B., & Mathez, G. 1993, A&A, 273, 367
Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2003, in HST Calibration Workshop: Hubble After the Installation of the ACS and the NICMOS Cooling System, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore, MD: STScI), 337
Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, Natur, 540, 248
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Krist, J. E., Hook, R. N., & Stoehr, F. 2011, Proc. SPIE, 8127, 81270J
Krogager, J.-K., Zirm, A. W., Toft, S., Man, A., & Brammer, G. 2014, ApJ, 797, 17
Lanz, L., Ogle, P. M., Alatalo, K., & Appleton, P. N. 2016, ApJ, 826, 29
Lee-Brown, D. B., Rudnick, G. H., Momcheva, I. G., et al. 2017, ApJ, 844, 43
Leethochawalit, N., Jones, T. A., Ellis, R. S., et al. 2016a, ApJ, 820, 84
Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. 2016b, ApJ, 831, 152
Man, A. W. S., Greve, T. R., Toft, S., et al. 2016, ApJ, 820, 11
Maoz, D., Mannucci, F., & Brandt, T. D. 2012, MNRAS, 426, 3282
McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, Proc. SPIE, 8446, 84460J
Meneghetti, M., Natarajan, P., Coe, D., et al. 2017, MNRAS, 472, 3177
Monna, A., Seitz, S., Balestra, I., et al. 2017, MNRAS, 466, 4094
Muzzin, A., Labbé, I., Franx, M., et al. 2012, ApJ, 761, 142
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, ApJS, 206, 8

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, ApJ, 777, 18
Muzzin, A., Marchesini, D., van Dokkum, P. G., et al. 2009, ApJ, 701, 1839
Nelson, E., van Dokkum, P., Franx, M., et al. 2014, Natur, 513, 394
Newman, A. B., Belli, S., & Ellis, R. S. 2015a, ApJL, 813, L7
Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018, ApJ, 862, 126
Newman, A. B., Ellis, R. S., Andreon, S., et al. 2014, ApJ, 788, 51
Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162
Newman, A. B., Ellis, R. S., & Treu, T. 2015b, ApJ, 814, 26
Newman, A. B., Smith, R. J., Conroy, C., Villaume, A., & van Dokkum, P. 2017, ApJ, 845, 157
Newman, A. B., Treu, T., Ellis, R. S., et al. 2013, ApJ, 765, 24
Oldham, L., Auger, M. W., Fassnacht, C. D., et al. 2017, MNRAS, 465, 3185
Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, PASP, 125, 654
Pickles, A. J. 1998, PASP, 110, 863
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A29
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Priewe, J., Williams, L. L. R., Liesenborgs, J., Coe, D., & Rodney, S. A. 2017, MNRAS, 465, 1030
Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259
Repp, A., & Ebeling, H. 2018, MNRAS, 479, 844
Richard, J., Smith, G. P., Kneib, J.-P., et al. 2010, MNRAS, 404, 325
Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, ApJ, 785, 104
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Shivaei, I., Reddy, N. A., Shapley, A. E., et al. 2015, ApJ, 815, 98
Simcoe, R. A., Burgasser, A. J., Schechter, P. L., et al. 2013, PASP, 125, 270
Singh, R., van de Ven, G., Jahnke, K., et al. 2013, A&A, 558, A43
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
Stark, D. P., Auger, M., Belokurov, V., et al. 2013, MNRAS, 436, 1040
Stark, D. P., Swinbank, A. M., Ellis, R. S., et al. 2008, Natur, 455, 775
Storey, P. J., & Zeippen, C. J. 2000, MNRAS, 312, 813
Swinbank, A. M., Webb, T. M., Richard, J., et al. 2009, MNRAS, 400, 1121
Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, ApJ, 749, 121
Tadaki, K.-i., Kodama, T., Nelson, E. J., et al. 2017, ApJL, 841, L25
Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Toft, S., Franx, M., van Dokkum, P., et al. 2009, ApJ, 705, 255
Toft, S., Gallazzi, A., Zirm, A., et al. 2012, ApJ, 754, 3
Toft, S., Zabl, J., Richard, J., et al. 2017, Natur, 546, 510
Treu, T., Ellis, R. S., Liao, T. X., et al. 2005, ApJ, 633, 174
Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, ApJ, 650, 18
Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2005, ApJ, 631, 145
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van der Wel, A., Rix, H.-W., Wuyts, S., et al. 2011, ApJ, 730, 38
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
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295
Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, MNRAS, 449, 361
Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, ApJ, 735, 86
Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, ApJL, 770, L39
Wong, K. C., Zabludoff, A. I., Ammons, S. M., et al. 2013, ApJ, 769, 52
Wuyts, E., Barrientos, L. F., Gladders, M. D., et al. 2010, ApJ, 724, 1182
Wuyts, S., Labbé, I., Schreiber, N. M. F., et al. 2008, ApJ, 682, 985
Yan, R., & Blanton, M. R. 2012, ApJ, 747, 61
Zitrin, A., Broadhurst, T., Umetsu, K., et al. 2009, MNRAS, 396, 1985
Zitrin, A., Meneghetti, M., Umetsu, K., et al. 2013, ApJL, 762, L30
Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327