



An Extremely Massive Quiescent Galaxy at $z = 3.493$: Evidence of Insufficiently Rapid Quenching Mechanisms in Theoretical Models*

Ben Forrest¹ , Marianna Annunziatella², Gillian Wilson¹ , Danilo Marchesini² , Adam Muzzin³ , M. C. Cooper⁴ , Z. Cemile Marsan³ , Ian McConachie¹, Jeffrey C. C. Chan¹ , Percy Gomez⁵, Erin Kado-Fong⁶ , Francesco La Barbera⁷ , Ivo Labbé⁸ , Daniel Lange-Vagle², Julie Nantais⁹ , Mario Nonino¹⁰ , Theodore Peña² , Paolo Saracco¹¹ , Mauro Stefanon¹² , and Remco F. J. van der Burg¹³

¹ Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA; benjamif@ucr.edu

² Physics and Astronomy Department, Tufts University, Medford, MA, USA

³ Department of Physics and Astronomy, York University, Toronto, Ontario, Canada

⁴ Center for Cosmology, Department of Physics and Astronomy, University of California, Irvine, Irvine, CA, USA

⁵ W. M. Keck Observatory, Kamuela, HI, USA

⁶ Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA

⁷ INAF–Osservatorio Astronomico di Capodimonte, Napoli, Italy

⁸ Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, VIC, Australia

⁹ Departamento de Ciencias Físicas, Universidad Andrés Bello, Santiago, Chile

¹⁰ INAF–Osservatorio Astronomico di Trieste, Trieste, Italy

¹¹ INAF–Osservatorio Astronomico di Brera, Milano, Italy

¹² Leiden Observatory, Leiden University, Leiden, The Netherlands

¹³ European Southern Observatory, Garching, Germany

Received 2019 October 22; revised 2019 November 17; accepted 2019 November 26; published 2020 February 5

Abstract

We present spectra of the most massive quiescent galaxy yet spectroscopically confirmed at $z > 3$, verified via the detection of Balmer absorption features in the H - and K -bands of Keck/MOSFIRE. The spectra confirm a galaxy with no significant ongoing star formation, consistent with the lack of rest-frame UV flux and overall photometric spectral energy distribution. With a stellar mass of $3.1^{+0.1}_{-0.2} \times 10^{11} M_{\odot}$ at $z = 3.493$, this galaxy is nearly three times more massive than the highest redshift spectroscopically confirmed absorption-line-identified galaxy known. The star formation history of this quiescent galaxy implies that it formed $>1000 M_{\odot} \text{ yr}^{-1}$ for almost 0.5 Gyr beginning at $z \sim 7.2$, strongly suggestive that it is the descendant of massive dusty star-forming galaxies at $5 < z < 7$ recently observed with ALMA. While galaxies with similarly extreme stellar masses are reproduced in some simulations at early times, such a lack of ongoing star formation is not seen there. This suggests the need for a quenching process that either starts earlier or is more rapid than that currently prescribed, challenging our current understanding of how ultra-massive galaxies form and evolve in the early universe.

Unified Astronomy Thesaurus concepts: [Extragalactic astronomy \(506\)](#); [Galaxy evolution \(594\)](#); [Galaxy formation \(595\)](#); [High-redshift galaxies \(734\)](#)

1. Introduction

Over the last decade, deeper and wider field near-infrared detected multiwavelength surveys have enabled the discovery and photometric investigation of rare ultra-massive galaxies (UMGs; $M_{*} > 10^{11} M_{\odot}$) at progressively higher redshifts (e.g., Rodighiero et al. 2007; Wiklind et al. 2008; Mancini et al. 2009; Marchesini et al. 2010; Stefanon et al. 2015; Marsan et al. 2017). Although most UMGs observed at $z > 2$ are still forming stars, often quite vigorously (Martis et al. 2016, 2019; Whitaker et al. 2017; Wang et al. 2019), the number of quiescent candidates has been increasing and exceeds the predictions of simulations by a factor of between 3 and 30, depending upon selection criteria (e.g., Straatman et al. 2014; Guarnieri et al. 2019; Alcalde Pampiega et al. 2019). A significant number of these massive quiescent systems have been spectroscopically confirmed at $1.5 < z < 2.5$, enabling a more precise characterization of their stellar populations, and improved modeling of their star formation histories due to the

detection of stellar continuum features (e.g., Belli et al. 2014, 2017, 2019; Newman et al. 2014; Kriek et al. 2016; Kado-Fong et al. 2017; Newman et al. 2018). Recently, three galaxies at slightly lower masses ($10.82 < \log(M/M_{\odot}) < 11.06$) at $3.75 < z < 4.01$ have also been confirmed via their Balmer absorption features (Tanaka et al. 2019; Valentino et al. 2019).

Due to the faintness of such objects at $z > 3$, the number of candidates spectroscopically confirmed at these higher redshifts has remained low (Marsan et al. 2015, 2017; Glazebrook et al. 2017; Schreiber et al. 2018b, 2018a, hereafter S18). While small, this higher redshift sample suggests that the selection techniques used for these candidates, typically involving rest-frame colors, yield relatively pure samples, though perhaps not complete (Marsan et al. 2015; Merlin et al. 2018, S18). The confirmation success rate in S18 also seems to confirm the aforementioned excess relative to simulations is indeed real.

The leading candidates for progenitors of these galaxies, which clearly must form stellar mass at extreme rates at early times, are high-redshift dusty star-forming galaxies (DSFGs). Recent ALMA observations of small numbers of these DSFGs at $5 < z < 7$ reveal large amounts of molecular gas and extreme star formation rates (e.g., Capak et al. 2011; Riechers

* The spectroscopic data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

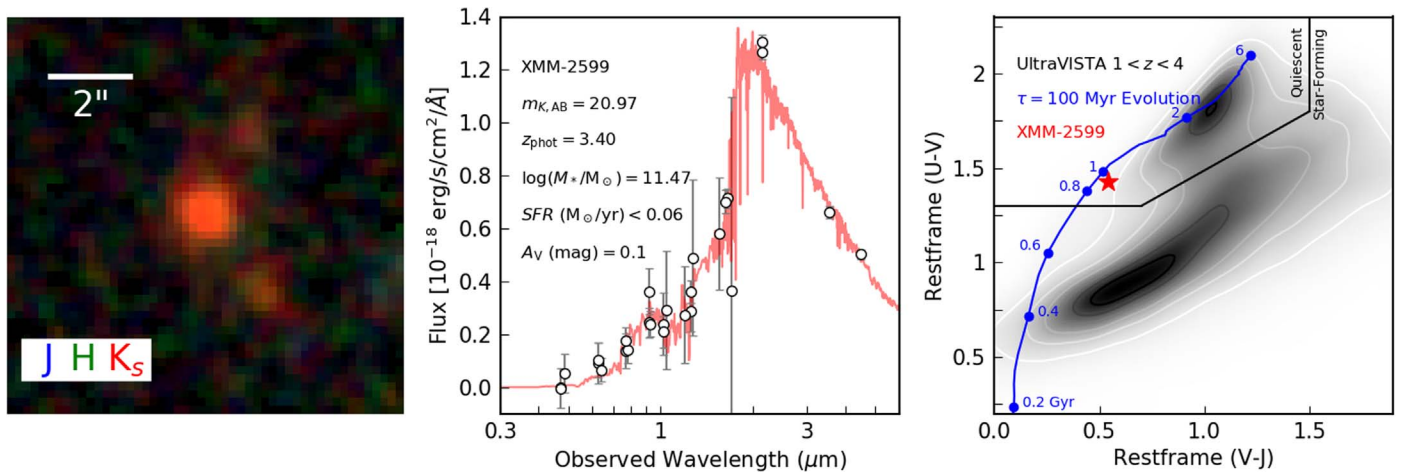


Figure 1. Photometric properties of XMM-2599. Left: near-infrared imaging of XMM-2599. Middle: photometric spectral energy distribution of XMM-2599. Data are shown in white with gray 1σ errorbars, while the best-fit template to the photometry alone is shown in red. Listed properties are also derived from the photometry alone. Right: XMM-2599 on the rest-frame UVJ diagram. A mass-complete sample of galaxies at $1 < z < 4$ from UltraVISTA is shown in gray for comparison. The evolution of a population with an exponentially declining star formation history parameterized by $\tau = 100$ Myr is shown in blue, with several ages labeled in gigayears. This track does not account for dust obscuration, and thus is not expected to agree with most of the star-forming population.

et al. 2013, 2017; Strandet et al. 2017; Marrone et al. 2018). The lack of deep stellar continuum spectra for these $z > 3$ UMGs, however (three UMGs with absorption features robustly detected; Glazebrook et al. 2017, S18), has prevented establishment of a firm link between these objects and the DSFGs, as photometric studies cannot robustly infer the past star formation history (SFH).

In this Letter, we present deep rest-frame optical spectra of XMM-2599, a quiescent UMG candidate at $z_{\text{phot}} \sim 3.4$. Our spectra confirm its quiescent nature and imply a period of intense star formation ($> 1000 M_{\odot} \text{yr}^{-1}$) in its $z \sim 5.5$ progenitor, consistent with most DSFGs observed at that epoch. The spectroscopic confirmation of XMM-2599, the most massive quiescent galaxy at $z > 3$, arguably represents the biggest challenge yet to the latest theoretical models of galaxy formation in the early universe, underlining the inadequate quenching mechanism(s) currently implemented in simulations.

Below we describe our target selection and spectral reduction in Section 2, our derivation of various galaxy characteristics in Section 3, and follow with a discussion (Section 4) and conclusions (Section 5). For this work we assume a Chabrier IMF, $H_0 = 70 \text{ km s}^{-1} \text{Mpc}^{-1}$, $\Omega_m = 0.3$, and $\Omega_{\Lambda} = 0.7$.

2. Data

2.1. Target Selection

Selected from deep, 28-band imaging catalogs of the VIDEO *XMM-Newton* field (spanning $0.3\text{--}4.5 \mu\text{m}$, M. A. Annunziatella et al., in preparation), the galaxy XMM-2599 (R.A. = $02^{\text{h}}27^{\text{m}}10^{\text{s}}.098$, decl. = $-04^{\circ}34'44''.988$) is luminous in the K_s -band ($m_{AB} = 20.97^{+0.02}_{-0.02}$), with a narrow singly peaked redshift probability distribution ($z_{\text{phot}} = 3.40^{+0.12}_{-0.10}$), and a spectral energy distribution (SED) consistent with a quenched galaxy (see Figure 1, which also lists the stellar population properties derived from SED modeling). Taken together, these three characteristics strongly suggest this galaxy is observed when the universe was only 1.5–2.0 billion years old, has a stellar mass $\log(M_*/M_{\odot}) \sim 11.5$, and is no longer forming stars at an appreciable rate. As shown in Figure 1, the galaxy also lies in the quiescent wedge of the rest-frame ($U-V$) versus

($V-J$) (UVJ) color-color diagram, consistent with the positions of poststarburst galaxies.

2.2. Spectroscopic Follow-up

We obtained deep spectra of XMM-2599 using the MOSFIRE spectrograph (McLean et al. 2010, 2012) on the Keck I telescope (PI Wilson; Figure 2). Observations were taken in November and December of 2018. A single mask was observed in K -band for $2^{\text{h}}45^{\text{m}}$, with an average seeing of $0''.6$, as determined from a slit star. Two masks in H -band were observed for on-source times of $2^{\text{h}}20^{\text{m}}$ and $2^{\text{h}}40^{\text{m}}$, with seeing of $0''.94$ and $1''.13$, respectively.

We began reduction by running the MOSFIRE Data Reduction Pipeline¹⁴ (DRP) to obtain 2D target and error spectra. The DRP constructs a pixel flat image, identifies slits, removes thermal contamination (K -band), performs wavelength calibration using sky lines, neon arc lamps, and argon arc lamps, removes sky background, and rectifies the spectrum. A custom Python code was written to perform 1D spectral extraction from the DRP outputs utilizing an optimal spectral extraction (Horne 1986).

Additional code was written to perform telluric corrections based on spectra of bright stars ($15 < m_{K_s} < 18$) included on the MOSFIRE slit masks. Similar to S18, this code uses the PHOENIX star models (Husser et al. 2013) to fit the near-infrared photometry of the stars and thus obtain intrinsic stellar spectra. The ratio of this model to the extracted 1D spectrum yields a telluric correction that is applied to other objects on the same mask.

The last piece of our reduction entailed identifying and masking out sky lines, which is of critical importance for such faint targets. To do this, we extracted 1D spectra of the sky from regions of slits that were uncontaminated by any object as determined from inspection of the 2D spectrum and K -band imaging. This resulted in ~ 10 spectra per mask, which were coadded to create a sky spectrum. The error spectra for these regions were added in quadrature, excluding wavelength regions of individual spectra that did not fall on the detector.

¹⁴ <https://github.com/Mosfire-DataReductionPipeline/MosfireDRP>

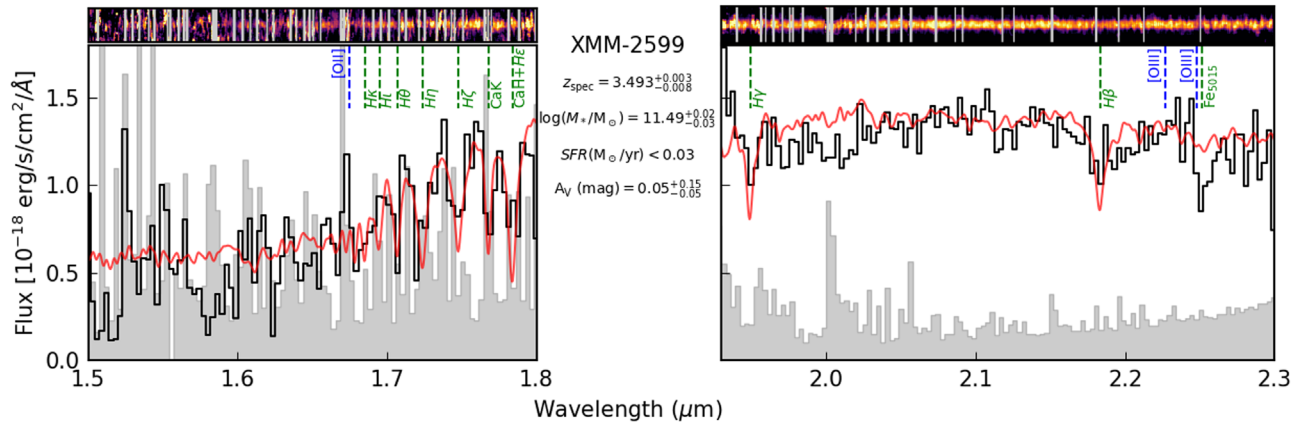


Figure 2. Near-infrared H - and K -band spectra for XMM-2599 and best-fit model. Top: the telluric corrected 2D spectra, smoothed for visual clarity. Strong sky lines are masked with gray lines. Bottom: the 1D extracted spectra, shown in bins $30 \text{ \AA}$ wide, are black, while the 1σ noise (including telluric correction) is gray. The best-fit template to the combined photometry and spectroscopy is plotted in red. The locations of absorption features are indicated in green, and the wavelengths corresponding to nebular emission from oxygen are blue.

We then fit and normalized by a “continuum” to the error curve to isolate noise spikes associated with sky lines. Any pixels on this curve above the 87.5th percentile of the sky spectra were considered to be strong sky lines, as was any adjacent pixel. This process reliably identifies sky lines when compared to a visual inspection of a 2D spectrum. Data from wavelengths affected by sky lines were then masked out for fitting purposes. For visualizations, sky line pixels were averaged with nearby nonaffected pixels to reduce the effects of the sky lines on the spectra, and data were binned.

2.3. Spectral Features

The final spectra of XMM-2599 show Balmer series absorption lines redshifted to $z = 3.493^{+0.003}_{-0.008}$ (Figure 2). These Balmer lines constrain the age of a galaxy, as they are associated with stars of mass $1.5\text{--}2 M_{\odot}$, which have main-sequence lifetimes of hundreds of millions of years, thereby breaking degeneracies in SED fitting associated with dust and stellar age. $H\gamma$, $H\epsilon$ + Ca H , $H\delta$, $H\eta$, and $H\theta$ are detected, while Ca K lies in a region of significant sky noise. $H\beta$ is seen in absorption, with the possibility of a small emission spike overlaid. We do not observe nebular emission from oxygen ($[\text{O III}]\lambda\lambda 4959, 5007$ and $[\text{O II}]\lambda\lambda 3726, 3729$), though the redshifted $[\text{O II}]$ doublet falls in a region of strong sky emission.

3. Analysis

3.1. Galaxy Fitting

For a consistent comparison with the sample from S18, we utilize the FAST++ code¹⁵ (Schreiber et al. 2018b, S18) to model the SEDs of our galaxies. FAST++ is a rewrite of FAST (Kriek et al. 2009) for C++ which allows for flexible SFH parameterizations as well as spectroscopic data of different wavelength resolutions. Furthermore, spectra are flux scaled to match the observed photometry for individual galaxies, and thus only spectral features/shapes contribute to the fit.

The spectrum from each bandpass was fit independently with the photometry to ensure that relative spectral flux calibrations between bandpasses did not affect the outcome. Both best-fit

templates were nearly identical, and the spectra were each scaled to match the resultant best fits. Said scaling differences here were $\sim 10\%$. Finally, each spectrum was allowed to vary relative to the other by up to 2 pixels to account for possible wavelength calibration errors. We then refit the photometry with the scaled spectra from both bandpasses—again yielding a best-fit template nearly identical to those produced with each band individually.

The grid of potential models tested with FAST++ included those with $3 < z < 4$, $8.0 < \log(\text{age}/\text{yr}) < \text{age of the universe at } z_{\text{model}}$, and $0 < A_V < 5$. Metallicities of $Z = 0.004$, 0.008 , 0.02 , and 0.05 were tested; however, the differences in χ^2 between the models of different metallicities are too small to differentiate given the signal-to-noise of our data. Throughout this work we have quoted results from the $Z = 0.02$ (Solar) metallicity run.

3.2. SFH

Given the ability of FAST++ to fit various functional forms of SFH, we begin with the form presented in S18, which can roughly reproduce the more complex shapes found in best-fit SFHs for massive quiescent galaxies at $z \sim 2$ (Belli et al. 2019):

$$\text{SFR}_{\text{base}}(t) \propto \begin{cases} e^{(t_{\text{burst}} - t)/\tau_{\text{rise}}}, & \text{for } t > t_{\text{burst}} \\ e^{(t - t_{\text{burst}})/\tau_{\text{decl}}}, & \text{for } t \leq t_{\text{burst}} \end{cases} \quad (1)$$

$$\text{SFR}(t) = \text{SFR}_{\text{base}}(t) \times \begin{cases} 1, & \text{for } t > t_{\text{free}} \\ R_{\text{SFR}}, & \text{for } t \leq t_{\text{free}} \end{cases}. \quad (2)$$

This SFH parameterization allows for a period of rising star formation, as well as decoupling the rising and falling exponential phases from the star formation at the time of observation (Papovich et al. 2011; Glazebrook et al. 2017; Schreiber et al. 2018b, S18). The grid of SFH parameters ranged from $7.0 < \log(t_{\text{burst}}/\text{yr}) < 9.2$, $7.0 < \log(\tau_{\text{rise}}/\text{yr}) < 9.5$, $7 < \log(\tau_{\text{decl}}/\text{yr}) < 9.5$, $7 < \log(t_{\text{free}}/\text{yr}) < 8.5$, and $-2.0 < \log(R_{\text{SFR}}) < 5.0$.

The best-fit SFH of the form given above implies that this galaxy formed $> 1000 M_{\odot} \text{ yr}^{-1}$ for almost 0.5 Gyr beginning at $z \sim 7.2$ (Figure 3). Our analysis makes use of this SFH, including in the derivation of the mass formation history in Figure 4. However, we also fit the data using a variety of other

¹⁵ <https://github.com/cschreib/fastpp>

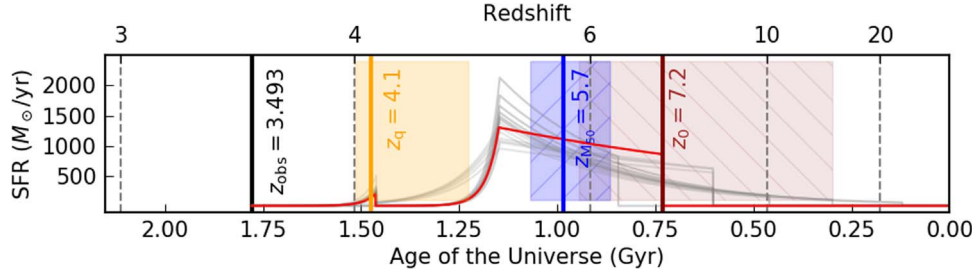


Figure 3. Best-fit star formation history for XMM-2599. The red curve indicates the SFR over cosmic time, with the maximum SFR and a characteristic average SFR shown in solar masses per year on the y-axis. The black line indicates the spectroscopic redshift and the maroon line is the time that the galaxy began forming stars (z_0). The orange line is the time at which SFR drops below 10% of the previous average SFR (z_q) while the blue line denotes the time at which half of the final stellar mass has been formed (z_{M50}). The width of shaded regions corresponds to 1σ confidence intervals on z_q , z_{M50} , and z_0 . Other $Z = Z_\odot$ models with $\chi^2_{\text{model}} - \chi^2_{\text{best}} < 1$ are shown in gray.

common functional forms of SFH, including exponentially declining, delayed exponentially declining, truncated, and top-hat forms. Aside from the delayed exponentially declining SFH, which builds stellar mass to unreasonable levels in the early universe, all functional forms yield similar results, with significant star formation completed by $z \sim 5$, and highly suppressed star formation possibly continuing until $z = 4\text{--}4.5$.

3.3. Star Formation Rates

Star formation rates are calculated in several ways for XMM-2599, as shown in Figure 5. Values for other UMGs are obtained from S18, and are calculated in the same way. The SED-derived SFR for XMM-2599 was calculated from FAST++, using the same parameter grid as above. The ultraviolet SFR is calculated from the best-fit SED template by integrating flux density over a 350 Å top-hat filter centered on 2800 Å rest frame and converting to a star formation rate (Kennicutt 1998; Muzzin et al. 2013a):

$$\text{SFR}_{\text{UV}}[M_\odot \text{ yr}^{-1}] = 3.23 \times 10^{-10} L_{2800}[L_\odot] \quad (3)$$

Similar calculations are done to determine SFR based on integrated line fluxes from the MOSFIRE spectra for [O II] (Kennicutt 1998, S18) and H β (Kewley et al. 2004, S18):

$$\text{SFR}_{[\text{O II}]}[M_\odot \text{ yr}^{-1}] = 1.59 \times 10^{-8} L_{[\text{O II}]}[L_\odot] \quad (4)$$

$$\text{SFR}_{\text{H}\beta}[M_\odot \text{ yr}^{-1}] = 5.46 \times 10^{-8} L_{\text{H}\beta}[L_\odot]. \quad (5)$$

In the case of XMM-2599, we note that strong emission is not obvious in either case. H β may have a small amount of emission overlaid on the stronger absorption feature, while there is strong sky emission on the wavelengths corresponding to [O II], yielding a signal-to-noise ratio of $\text{SNR}_{[\text{O II}]} \sim 0.2$. Using the above equations we calculate an upper limit of $\text{SFR}_{\text{H}\beta} < 4 M_\odot \text{ yr}^{-1}$ for XMM-2599, and find that a line flux of $f_{[\text{O II}]} = 5.5 \times 10^{-18} \text{ erg/s/cm}^2$ is necessary to reproduce this value. Assuming an emission feature width of 10 Å and a continuum level of $\sim 6 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$ from the best-fit SED, this corresponds to a peak line flux density of $f_{\lambda, [\text{O II}]} \sim 1 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$. Although this is broadly consistent with the spectra, we do not plot this value or a limit on Figure 5 due to the very low signal-to-noise.

4. Discussion

4.1. Progenitors of Quiescent UMGs

In order to build up such a large stellar mass at early times, the progenitors of systems like XMM-2599 must have been explosively star-forming at $z \sim 5\text{--}6$. DSFGs at $z > 5$ have been confirmed using longer wavelength data, such as that provided by ALMA, but those with large published gas and/or stellar masses remain few (Capak et al. 2011; Riechers et al. 2013; Cooray et al. 2014; Spilker et al. 2016; Strandet et al. 2017). While the low number densities of these DSFGs suggest that they cannot account for all of the quiescent galaxies photometrically identified at $3 < z < 4$ (Straatman et al. 2014), it seems possible that they could be progenitors of the most massive end of the quiescent UMG population, such as XMM-2599.

In Figure 4, we explore this possibility for a sample of high-redshift DSFGs with published stellar masses, molecular gas masses, and star formation rates (Capak et al. 2011; Riechers et al. 2013, 2017; Cooray et al. 2014; Ma et al. 2015; Strandet et al. 2017; Marrone et al. 2018; Williams et al. 2019; Jin et al. 2019). These systems have masses consistent with the mass evolution of XMM-2599 derived from our best-fit SFH. Additionally, the available gas allows for nearly all of them to reach a stellar mass of $\log(M/M_\odot) > 11$ by $z \sim 3.5$ with a plausible star formation efficiency through cosmic time. While such massive high-redshift DSFGs are rare, their existence implies that other galaxies as massive as XMM-2599 at $z \sim 3.5$ exist. Moreover, though many of these DSFGs have clear optical counterparts, the recent discovery of a significant number of DSFGs at $3 < z < 8$ with no such counterpart indicates that such galaxies may exist in sufficient numbers to be progenitors of the $z > 3$ quiescent UMG population down to even lower masses, and have simply avoided detection thus far (Wang et al. 2019; Williams et al. 2019).

4.2. Comparison to Simulations

Quenched galaxies such as XMM-2599 are extremely rare as the stellar mass function for the quiescent population declines steeply at the high-mass end. Data from the 1.62 deg² UltraVISTA survey (McCracken et al. 2012) implies that quiescent UMGs at $3 < z < 4$ with $\log(M/M_\odot) > 11$ have a density of $n \sim 10^{-5.83} \text{ Mpc}^{-3}$, while those with $\log(M/M_\odot) > 11.5$ are estimated to be more than a factor of 10 rarer, at $n \sim 10^{-6.97} \text{ Mpc}^{-3}$ (Muzzin et al. 2013b). However, they are observed in numbers significantly higher than those predicted by simulations (see, e.g., Figure 14 of Alcalde

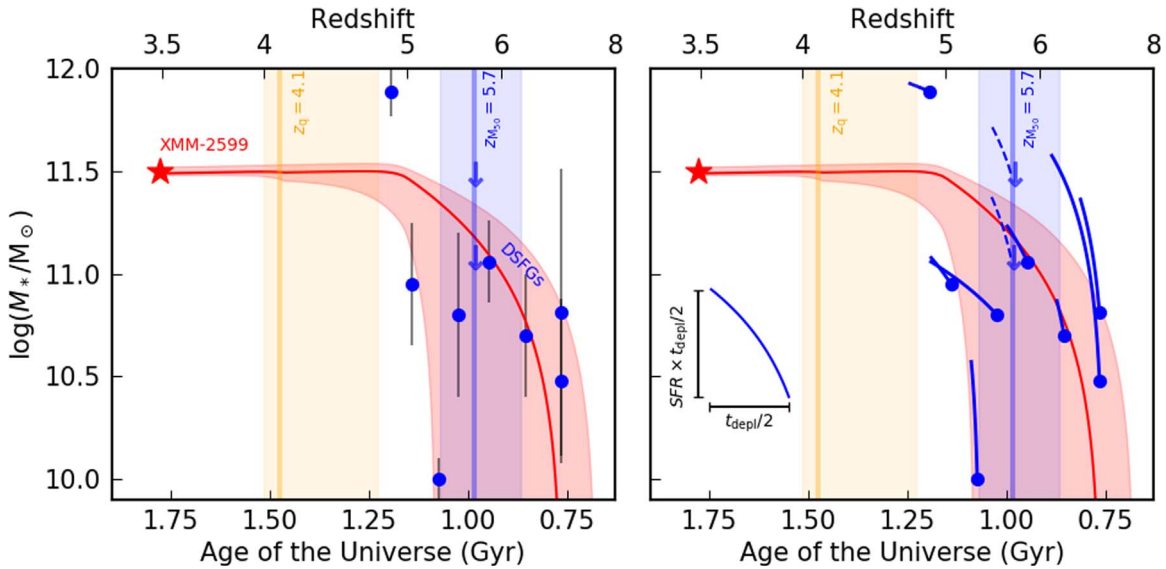


Figure 4. High-redshift DSFGs as potential progenitors of XMM-2599. We show the stellar mass evolution for XMM-2599 in red as calculated from our best-fit SFH, with a shaded 68% confidence interval. Left: several high-redshift DSFGs are shown in blue with errors on masses (Capak et al. 2011; Riechers et al. 2013, 2017; Cooray et al. 2014; Ma et al. 2015; Strandet et al. 2017; Marrone et al. 2018; Jin et al. 2019; Williams et al. 2019). Reported upper limits are plotted as arrows. Right: blue segments show the evolution of the DSFGs assuming the published star formation rate held constant over half the gas depletion timescale (i.e., half of the available gas is turned into stars). When no gas depletion timescale or gas mass is reported, we set $t_{\text{depl}} = 0.1$ Gyr, a value typical of the population. The overlap of these tracks with the mass evolution of XMM-2599 suggests that they are potential high-redshift progenitor systems.

Pampliega et al. 2019). Tens of $z > 3$ UMGs have been spectroscopically confirmed via detection of faint emission lines implying ongoing star formation or AGN activity (Kubo et al. 2015; Marsan et al. 2015, 2017, S18). However, only three such systems have robust redshifts from the detection of absorption lines alone: ZF-COS-20115 at $z = 3.715$ with $M_* = 1.15^{+0.16}_{-0.09} \times 10^{11} M_\odot$ (Glazebrook et al. 2017, S18), 3D-EGS-18996 at $z = 3.239$ with $M_* = 9.8^{+0.04}_{-0.06} \times 10^{10} M_\odot$ (S18), and 3D-EGS-40032 at $z = 3.219$ with $M_* = 2.03^{+0.16}_{-0.14} \times 10^{11} M_\odot$ (S18).

Given the low observed number densities, large volume simulations are required for comparison. In Figure 5 we compare observed absorption-line UMGs to simulated galaxies in snapshots from Illustris TNG300 (302.6 Mpc on a side; Naiman et al. 2018; Nelson et al. 2018; Marinacci et al. 2018; Pillepich et al. 2018; Springel et al. 2018). TNG300 is able to suppress star formation in massive galaxies at high redshift and easily reproduces 3D-EGS-18996 and 3D-EGS-40032, and reproduces ZF-COS-20115 within the observational errors. Still, at $z = 3.49$ TNG300 has low number densities for high-mass galaxies with $SFR < 5 M_\odot \text{ yr}^{-1}$; $n \sim 10^{-6.24} \text{ Mpc}^{-3}$ for $\log(M/M_\odot) > 11$ and $n \sim 10^{-7.44} \text{ Mpc}^{-3}$ for $\log(M/M_\odot) > 11.5$. The photometric sample from which XMM-2599 was selected has several other massive, quiescent candidates at $z_{\text{phot}} \sim 3.5$, corresponding to a number density of $\log(n) \sim -6.64^{+0.15}_{-0.23} \text{ Mpc}^{-3}$. This number density predicts ~ 6 similar objects in the simulated box, while having zero would correspond to a 3σ outlier using a simple Poisson calculation. Additionally, XMM-2599 has an SFR roughly 2–10 $\times$ below that of any simulated galaxy of similar mass based on the various SFR limit determinations.

Three possible analogs for ZF-COS-20115 were found in the MERAXES semianalytic model (Mutch et al. 2016; Qin et al. 2017) (box size = $125 h^{-1} \text{ Mpc}$), though none of these approach the mass of XMM-2599. Other large simulations such as Millenium ($500 h^{-1} \text{ Mpc}$ on a side; Springel et al. 2005;

Henriques et al. 2015) do not come close to reproducing any of these quiescent UMGs. In order to do so simulations require either a more rapid buildup of stellar mass in situ during the epoch of reionization or a faster quenching mechanism than is currently prescribed.

We also compare the evolution of XMM-2599 based on our best-fit SFH, as shown in Figure 5. This shows that at $5 < z < 6$, the characteristics of XMM-2599, i.e., large stellar masses and extreme SFRs, are well reproduced by TNG300. This is also clear from the ability of TNG300 to reproduce the observed properties of the DSFGs. However, TNG300 is unable to match the rapidity with which XMM-2599 is quenched at $3.5 < z < 4$. Various parametric forms of SFH were tested, as well as different metallicities, and none of these eliminate this issue. For a given SFH form, a lower metallicity generally results in a best-fit model with larger SFR for shorter durations occurring at earlier epochs before rapidly quenching, thus exacerbating the disagreement with simulations. While the highest metallicity tested ($Z = 0.05$) moves the peak of star formation to later times, the SFH looks nearly identical after $z = 4.1$, and thus the disagreement remains. Several studies have shown that high-redshift massive quiescent galaxies have metallicities only slightly below solar (e.g., Kriek et al. 2016; Belli et al. 2019; Estrada-Carpenter et al. 2019).

4.3. Possible Alternatives

Upon follow-up with high-resolution *Hubble Space Telescope* (HST) imaging, a number of red, massive, high-redshift galaxies detected with near-infrared ground-based imaging have been revealed to be close pairs (Marsan et al. 2019; Mowla et al. 2019). We lack high-resolution HST imaging for XMM-2599, and thus the case of two compact galaxies in extreme proximity cannot be ruled out. However, we also note that examples of this, as shown in Figures 3 and 4 of Marsan et al. (2019) and Figure 2 of Mowla et al. (2019), exhibit clear

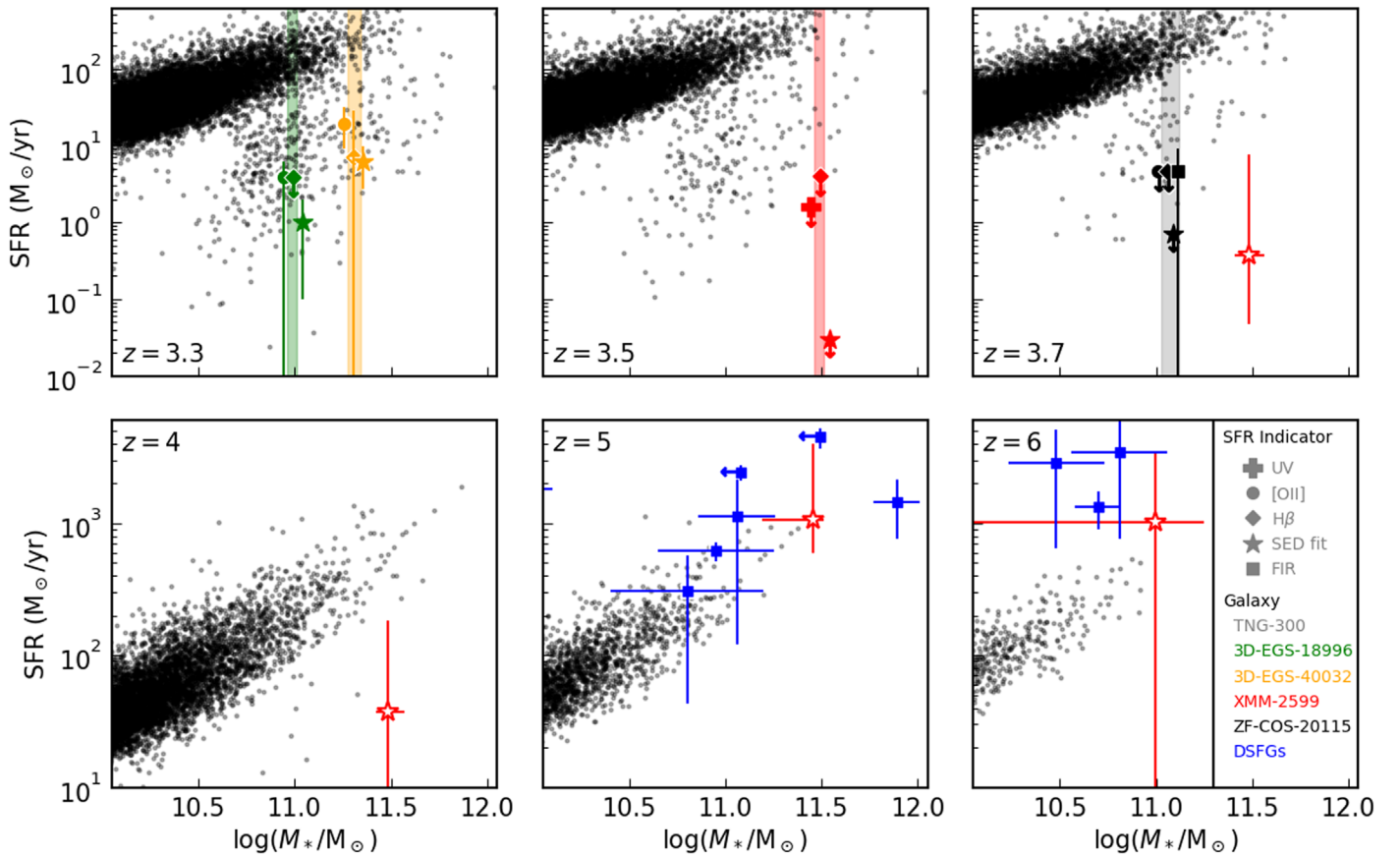


Figure 5. Comparison to the Illustris TNG-300 simulation on the SFR- M_* plane. We show the spectroscopically confirmed absorption-line identified UMGs at $z_{\text{spec}} > 3$ (green, orange, red, and black), simulated galaxies from six snapshots in Illustris TNG-300 (gray), and the DSFGs from Figure 4 (blue). Several probes of star formation are shown differentiated by marker style, many as 1σ upper limits. These are offset along the abscissa for visual clarity, while the best-fit stellar mass is shown as a column, with the width indicating the 68% statistical error. Using the best-fit SFH from Figure 3, we plot the position of XMM-2599 at previous epochs as well (open red stars). Note that the range of the ordinate axis differs in the two rows.

deviations from a compact, circular object in the near-infrared imaging, which XMM-2599 does not (Figure 1).

Such close pairs are evidence of future mergers, and therefore XMM-2599 may be the result of a recent dry merger, which lacked sufficient cold gas to trigger substantial star formation. Future high-resolution imaging could pick-up more structural features and shed light on whether this object is the result of a recent dry merger, or indeed a pair of galaxies. We note that, assuming a 1:1 mass ratio, these galaxies/progenitors would still have stellar masses $\log(M/M_\odot) \sim 11.2$, making them both UMGs.

Nearby neighbors can also contaminate bands with lower spatial resolution, in particular the IRAC bandpasses. ZF-COS-20115 provides a case study of this, as an optically invisible neighbor led to an initial overestimate of the stellar mass by $\sim 40\%$ (Glazebrook et al. 2017; Schreiber et al. 2018b). While XMM-2599 has two neighbors in the near-infrared ($\sim 1''$ – $2''$ away), they are sufficiently distant as to not contaminate the photometry, and the light profile of XMM-2599 is consistent with a roughly circular, singly peaked distribution perturbed by noise. Refitting XMM-2599 assuming extreme contamination from these neighbors in IRAC in line with Schreiber et al. (2018b), i.e., 15% in $3.6\mu\text{m}$ imaging and 28% in $4.5\mu\text{m}$ imaging, still results in a stellar mass of $\log(M/M_\odot) \sim 11.4$, more massive than any other $z > 3$ quiescent UMG.

While massive quiescent populations remain rare at high redshift, star-forming systems in this mass regime, nearly all of

which are dust obscured, are more common (Marchesini et al. 2014; Martis et al. 2016; Whitaker et al. 2017; Martis et al. 2019). Since heavily dust-obscured galaxies and quiescent galaxies can have similar UV-NIR photometry, it is important to rule out the possibility that XMM-2599 is a dusty galaxy. Although large amounts of dust can severely dampen emission line signatures in spectra, reproduction of absorption lines by dust is difficult and requires an old stellar population. Long wavelength data is a certain way to rule out ongoing dust-obscured star formation but the only far-infrared imaging in the region, with Herschel-PACS in HerMES (Oliver et al. 2012), shows no detection near XMM-2599. However, the imaging would only detect objects with $\text{SFR}_{\text{IR}} > 1000 M_\odot$ and is thus insufficiently deep to constrain the nature of XMM-2599.

ALMA follow-up of massive galaxies at $z > 3$ has shown that UVJ color selection also does a good job of identifying truly quiescent galaxies (Schreiber et al. 2018b, S18). XMM-2599 has rest-frame colors $(U-V) = 1.43^{+0.03}_{-0.02}$ and $(V-J) = 0.54^{+0.06}_{-0.02}$, thus placing it within the quiescent wedge of the UVJ diagram (Figure 1). More specifically, XMM-2599 lies in the blue corner of the quiescent wedge, typically associated with younger, poststarburst galaxies, as opposed to redder galaxies that quenched in the distant past. This limits the amount of dust obscuration possible, as substantial dust would move the galaxy toward the red side of the wedge. Additionally, we calculate the rest-frame equivalent width

value of $EW_{H\beta,0} = 13.8^{+5.8}_{-2.9}$ Å. Of the 77 low-redshift LIRGS and ULIRGS with $EW_{H\beta}$ values in Table 2 of Poggianti & Wu (2000), only 4 have absorption, and XMM-2599 has about twice the equivalent width of these, implying that dust-obscured SFGs have trouble reproducing such deep Balmer lines (see also Section 4.1 of Schreiber et al. 2018b). Given the lack of emission lines, the superior fit of quiescent galaxy templates to the data due to depth of the Balmer lines, and the rest-frame colors of XMM-2599, all the evidence suggests this galaxy is quiescent. Nonetheless, deep far-infrared data is important to completely exclude the possibility of extreme dust-obscuration of ongoing star formation.

5. Conclusions

In this work we presented spectra confirming the existence of a quiescent galaxy at $z = 3.493$ with a stellar mass of $3.1 \times 10^{11} M_{\odot}$. The rest-frame colors combined with the lack of emission lines from nebular oxygen reduce the likelihood of ongoing, dust-obscured star formation. This galaxy’s SFH suggests a period of intense star formation, $> 1000 M_{\odot} \text{ yr}^{-1}$ for several hundred Myr at $z \sim 6$, consistent with the most gas-rich DSFGs observed at that epoch.

Simulations have improved substantially in the last few years, and are able to reproduce the massive, star-forming DSFGs observed at high redshift that are considered possible progenitors for massive quenched galaxies such as XMM-2599. However, they are still unable to reproduce massive, quiescent galaxies at $z \sim 4$. The specific mechanisms that enable the rapid transformation of these galaxies is unclear, and may in fact be the result of several concurrent events. While gas-rich major mergers are important in building up the stellar mass at early times, a reduction in the number of these events would limit the amount of gas available for star formation. Virial shocks and increased feedback from active galactic nuclei could provide the energy necessary to keep any remaining gas heated, thus prevent the cooling and collapse necessary for forming stars (e.g., Steinborn et al. 2015; Faisst et al. 2017; Man & Belli 2018; Spilker et al. 2018). Improved ability to replicate these events in the early universe is required to reproduce this extreme galaxy in simulations.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. This work is supported by the National Science Foundation through grants AST-1517863, AST-1518257, and AST-1815475, by *HST* program number GO-15294, and by grant numbers 80NSSC17K0019 and NNX16AN49G issued through the NASA Astrophysics Data Analysis Program (ADAP). Support for program number GO-15294 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. Further support was provided by the Faculty Research Fund (FRF) of Tufts University and by Universidad Andrés Bello grant number DI-12-19/R. This research made use of Astropy,¹⁶ a community-developed core Python package for Astronomy (Robitaille et al. 2013; Price-Whelan et al.

2018). B.F. thanks L. Alcorn for discussions regarding the MOSFIRE DRP and E. Conant for advice. We also thank the anonymous referee for a constructive report that improved the manuscript.

ORCID iDs

Ben Forrest  <https://orcid.org/0000-0001-6003-0541>
 Gillian Wilson  <https://orcid.org/0000-0002-6572-7089>
 Danilo Marchesini  <https://orcid.org/0000-0001-9002-3502>
 Adam Muzzin  <https://orcid.org/0000-0002-9330-9108>
 M. C. Cooper  <https://orcid.org/0000-0003-1371-6019>
 Z. Cemile Marsan  <https://orcid.org/0000-0002-7248-1566>
 Jeffrey C. C. Chan  <https://orcid.org/0000-0001-6251-3125>
 Erin Kado-Fong  <https://orcid.org/0000-0002-0332-177X>
 Francesco La Barbera  <https://orcid.org/0000-0003-1181-6841>
 Ivo Labbé  <https://orcid.org/0000-0002-2057-5376>
 Julie Nantais  <https://orcid.org/0000-0002-7356-0629>
 Mario Nonino  <https://orcid.org/0000-0001-6342-9662>
 Theodore Peña  <https://orcid.org/0000-0002-0033-5041>
 Paolo Saracco  <https://orcid.org/0000-0003-3959-2595>
 Mauro Stefanon  <https://orcid.org/0000-0001-7768-5309>
 Remco F. J. van der Burg  <https://orcid.org/0000-0003-1535-2327>

References

- Alcalde Pampliega, B., Pérez-González, P. G., Barro, G., et al. 2019, *ApJ*, **876**, 135
- Belli, S., Newman, A. B., & Ellis, R. S. 2017, *ApJ*, **834**, 18
- Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, **874**, 17
- Belli, S., Newman, A. B., Ellis, R. S., & Konidaris, N. P. 2014, *ApJL*, **788**, L29
- Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, *Natur*, **470**, 233
- Cooray, A., Calanog, J., Wardlow, J. L., et al. 2014, *ApJ*, **790**, 40
- Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2019, *ApJ*, **870**, 133
- Faisst, A. L., Carollo, C. M., Capak, P. L., et al. 2017, *ApJ*, **839**, 71
- Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, *Natur*, **544**, 71
- Guarnieri, P., Maraston, C., Thomas, D., et al. 2019, *MNRAS*, **483**, 3060
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, *MNRAS*, **451**, 2663
- Home, K. 1986, *PASP*, **98**, 609
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, **553**, A6
- Jin, S., Daddi, E., Magdis, G. E., et al. 2019, *ApJ*, **887**, 144
- Kado-Fong, E., Marchesini, D., Marsan, Z. C., et al. 2017, *ApJ*, **838**, 57
- Kennicutt, R. C. 1998, *ARA&A*, **36**, 189
- Kewley, L. J., Geller, M. J., & Jansen, R. A. 2004, *AJ*, **127**, 2002
- 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
- Kubo, M., Yamada, T., Ichikawa, T., et al. 2015, *ApJ*, **799**, 38
- Ma, J., Gonzalez, A. H., Spilker, J. S., et al. 2015, *ApJ*, **812**, 88
- Man, A., & Belli, S. 2018, *NatAs*, **2**, 695
- Mancini, C., Matute, I., Cimatti, A., et al. 2009, *A&A*, **500**, 705
- Marchesini, D., Muzzin, A., Stefanon, M., et al. 2014, *ApJ*, **794**, 65
- Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, **725**, 1277
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, **480**, 5113
- Marrone, D. P., Spilker, J. S., Hayward, C. C., et al. 2018, *Natur*, **553**, 51
- Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2015, *ApJ*, **801**, 133
- Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2017, *ApJ*, **842**, 21
- Marsan, Z. C., Marchesini, D., Muzzin, A., et al. 2019, *ApJ*, **871**, 201
- Martis, N. S., Marchesini, D., Brammer, G. B., et al. 2016, *ApJ*, **827**, L25
- Martis, N. S., Marchesini, D. M., Muzzin, A., et al. 2019, *ApJ*, **882**, 65
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, **544**, A156
- McLean, I. S., Steidel, C. C., Epps, H., et al. 2010, *Proc. SPIE*, **7735**, 77351E
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, *Proc. SPIE*, **8446**, 84460J
- Merlin, E., Fontana, A., Castellano, M., et al. 2018, *MNRAS*, **473**, 2098
- Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, *ApJ*, **880**, 57
- Mutch, S. J., Geil, P. M., Poole, G. B., et al. 2016, *MNRAS*, **462**, 250
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, *ApJS*, **206**, 8

¹⁶ <http://www.astropy.org>

- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, *ApJ*, **777**, 18
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **477**, 1206
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018, *ApJ*, **862**, 125
- Newman, A. B., Ellis, R. S., Andreon, S., et al. 2014, *ApJ*, **788**, 51
- Oliver, S. J., Bock, J., Altieri, B., et al. 2012, *MNRAS*, **424**, 1614
- Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, *MNRAS*, **412**, 1123
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *MNRAS*, **475**, 648
- Poggianti, B. M., & Wu, H. 2000, *ApJ*, **529**, 157
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, *AJ*, **156**, 123
- Qin, Y., Mutch, S. J., Duffy, A. R., et al. 2017, *MNRAS*, **471**, 4345
- Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, *Natur*, **496**, 329
- Riechers, D. A., Leung, T. K. D., Ivison, R. J., et al. 2017, *ApJ*, **850**, 1
- Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, *A&A*, **558**, A33
- Rodighiero, G., Cimatti, A., Franceschini, A., et al. 2007, *A&A*, **470**, 21
- Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018a, *A&A*, **618**, A85
- Schreiber, C., Labbé, I., Glazebrook, K., et al. 2018b, *A&A*, **611**, A22
- Spilker, J. S., Aravena, M., Béthermin, M., et al. 2018, *Sci*, **361**, 1016
- Spilker, J. S., Marrone, D. P., Aravena, M., et al. 2016, *ApJ*, **826**, 112
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Stefanon, M., Marchesini, D., Muzzin, A., et al. 2015, *ApJ*, **803**, 11
- Steinborn, L. K., Dolag, K., Hirschmann, M., Almudena Prieto, M., & Remus, R. S. 2015, *MNRAS*, **448**, 1504
- Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJ*, **783**, L14
- Strandet, M. L., Weiss, A., Breuck, C. D., et al. 2017, *ApJ*, **842**, L15
- Tanaka, M., Valentino, F., Toft, S., et al. 2019, *ApJL*, **885**, L34
- Valentino, F., Tanaka, M., Davidzon, I., et al. 2019, *ApJ*, in press (arXiv:1909.10540)
- Wang, T., Schreiber, C., Elbaz, D., et al. 2019, *Natur*, **572**, 211
- Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, *ApJ*, **850**, 208
- Wiklund, T., Dickinson, M., Ferguson, H. C., et al. 2008, *ApJ*, **676**, 781
- Williams, C. C., Labbe, I., Spilker, J., et al. 2019, *ApJ*, **884**, 154