The Astrophysical Journal, 908:135 (13pp), 2021 February 20 https://doi.org/10.3847/1538-4357/abce66

The Fundamental Plane of Massive Quiescent Galaxies at $z \sim 2$

Mikkel Stockmann[1,2,3], Inger Jørgensen[4], Sune Toft[1,2], Christopher J. Conselice[5,6], Andreas Faisst[7], Berta Margalef-Bentabol[8], Anna Gallazzi[9], Stefano Zibetti[9], Gabriel B. Brammer[1,2], Carlos Gómez-Guijarro[10], Michaela Hirschmann[3], Claudia D. Lagos[1,11,12], Francesco M. Valentino[1,2,3], and Johannes Zabl[13]
[1] Cosmic Dawn Center (DAWN), Denmark; mikkelstockmann@gmail.com
[2] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2100 Copenhagen Ø, Denmark
[3] DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2100 Copenhagen, Denmark
[4] Gemini Observatory, 670 N. A'ohoku Place, Hilo, HI 96720, USA
[5] Centre for Astronomy and Particle Theory, School of Physics & Astronomy, University of Nottingham, Nottingham NG7 2RD, UK
[6] Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester, UK
[7] IPAC, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA
[8] LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, France
[9] INAF-Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy
[10] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France
[11] International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
[12] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[13] Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France
Received 2019 December 29; revised 2020 November 25; accepted 2020 November 25; published 2021 February 18

Abstract

We examine the Fundamental Plane (FP) and mass-to-light ratio (M/L) scaling relations using the largest sample of massive quiescent galaxies at $1.5 < z < 2.5$ to date. The FP (r_e, σ_e, I_e) is established using 19 *UVJ* quiescent galaxies from COSMOS with Hubble Space Telescope H_{F160W} rest-frame optical sizes and X-shooter absorption line-measured stellar velocity dispersions. For a very massive, $\log(M_*/M_\odot) > 11.26$, subset of eight quiescent galaxies at $z > 2$, from Stockmann et al., we show that they cannot passively evolve to the local Coma cluster relation alone and must undergo significant structural evolution to mimic the sizes of local massive galaxies. The evolution of the FP and M/L scaling relations, from $z = 2$ to present day, for this subset are consistent with passive aging of the stellar population and minor merger structural evolution into the most massive galaxies in the Coma cluster and other massive elliptical galaxies from the MASSIVE Survey. Modeling the luminosity evolution from minor merger-added stellar populations favors a history of merging with "dry" quiescent galaxies.

Unified Astronomy Thesaurus concepts: Galaxy dynamics (591); Galaxy evolution (594); Galaxy physics (612); Elliptical galaxies (456); Giant elliptical galaxies (651); Quenched galaxies (2016); Scaling relations (2031); Galaxy mergers (608); Galaxy quenching (2040); Hubble Space Telescope (761); High resolution spectroscopy (2096)

1. Introduction

The most massive local elliptical galaxies, believed to be one of the most mature stages of galaxy evolution, have been shown to form the majority of their stars rapidly, in the densest environments at $z > 2$–5 (e.g., Blakeslee et al. 2003; Thomas et al. 2005; Greene et al. 2015). Understanding the formation and evolution of these systems is a complex task. One way to address this is to study their progenitors in the early universe and to see how their properties differentiate from their $z = 0$ counterparts, from which evolution can be inferred.

A population of massive, $\log(M_*/M_\odot) > 11$, quiescent galaxies which are the possible progenitor candidates of modern ellipticals have been located at $z > 2$ (Daddi et al. 2004; Kriek et al. 2009; Toft et al. 2012; Belli et al. 2017), which allows us to observe the evolution of the most massive systems in the universe. It is clear that these early massive galaxies have extremely compact sizes (Daddi et al. 2005; Trujillo et al. 2006, 2007; Buitrago et al. 2008; van Dokkum et al. 2008; Conselice et al. 2011; Szomoru et al. 2012), which are three to five times smaller than the present-day most massive elliptical galaxies at the same mass, and are also younger with more recent star formation (e.g., Stockmann et al. 2020).

From these observations rapid-size evolution has been inferred for field early-type galaxies across time (Newman et al. 2012; van der Wel et al. 2014; Faisst et al. 2017; Morishita et al. 2019; Mowla et al. 2019). Simulations have shown dry mergers to be an efficient process for making galaxies larger (e.g., Hopkins et al. 2009; Naab et al. 2009; Bezanson et al. 2009; Hilz et al. 2012, 2013; Remus et al. 2017; Lagos et al. 2018). Toft et al. (2014) proposed an evolutionary sequence of massive galaxies where the most massive elliptical galaxies, from the present-day universe, were formed in violent starbursts. These later quench, possibly via active galactic nuclei (AGNs), to become the compact quiescent galaxies at $z > 2$ suggested to undergo rapid size evolution and become the massive elliptical galaxies in the local universe (see also Cimatti 2008; Simpson et al. 2014; Gómez-Guijarro et al. 2018; Habouzit et al. 2019).

Scaling relations between different properties of galaxies, and how these evolve through time, can give us significant information about how galaxies are assembled over cosmic time. One way to do this is by studying the evolution of scaling relationships between various quantities. For example, massive local elliptical galaxies in the nearby universe are found to follow an empirical relation known as the Fundamental Plane (FP; Djorgovski & Davis 1987; Dressler et al. 1987) between surface brightness, internal velocity, and size.

Table 1
Sample Summary

Target ID	R.A.	Decl.	z_{spec}	log age yr^{-1}	$\log M_*/M_\odot$	$\log M_{dyn}/M_\odot$	$r_{e,circ}$	σ_e	$\log\langle I\rangle_{e,B}$
UV-105842	150.26265	2.0177791	2.0195	$9.19^{+0.26}_{-0.33}$	$11.68^{+0.16}_{-0.17}$	11.47 ± 0.19	2.91 ± 0.29	263 ± 57	4.40
UV-171687	149.88702	2.3506956	2.1020	$9.13^{+0.28}_{-0.32}$	$11.51^{+0.18}_{-0.19}$	11.31 ± 0.24	4.49 ± 0.45	182 ± 50	3.99
UV-90676	150.48750	2.2700379	2.4781	$9.09^{+0.29}_{-0.29}$	$11.78^{+0.17}_{-0.18}$	11.78 ± 0.21	4.08 ± 0.41	347 ± 82	4.46
UV-155853	149.55630	2.1672480	1.9816	$9.23^{+0.24}_{-0.33}$	$11.62^{+0.18}_{-0.17}$	11.57 ± 0.11	4.20 ± 0.42	247 ± 30	3.96
UV-230929	150.20842	2.7721019	2.1679	$9.10^{+0.28}_{-0.28}$	$11.48^{+0.16}_{-0.16}$	11.16 ± 0.07	1.48 ± 0.15	252 ± 21	5.11
CP-1243752	150.07394	2.2979755	2.0903	$9.23^{+0.24}_{-0.32}$	$11.79^{+0.17}_{-0.17}$	11.61 ± 0.13	2.54 ± 0.25	350 ± 53	4.59
CP-540713	150.32512	1.8185385	2.0409	$9.16^{+0.27}_{-0.32}$	$11.26^{+0.22}_{-0.23}$	11.53 ± 0.24	1.46 ± 0.15	353 ± 97	4.66
UDS-19627[a]	34.57125	−5.3607778	2.0389	$9.08^{+0.11}_{-0.10}$	$11.37^{+0.13}_{-0.10}$	11.33 ± 0.14	1.43 ± 0.14	318 ± 53	5.20

Notes. Column 1: target ID from S20; Column 2: R.A. in degrees (J2000); Column 3: decl. in degrees (J2000); Column 4: spectroscopic redshift; Column 5: mass-weighted age; Column 6: stellar mass; Column 7: circularized dynamical mass calculated using $\beta(n)$ from Cappellari et al. (2006); Column 8: circularized effective radius in kpc; Column 9: stellar velocity dispersion in km s^{-1}; Column 10: average surface brightness within $r_{e,circ}$ in $L_{B,\odot}pc^{-2}$ (see Appendix A).
[a] The spectroscopic redshift, age, stellar mass, and stellar velocity dispersion are from Toft et al. (2012) and the HST/WFC3 H_{F160W} sizes are from S20.

The zero-point of the edge-on FP has been observed to evolve with redshift, complementary to the mass-to-light ratio (M/L) (Faber et al. 1987), which has made this a preferred tool in studying the structural and luminosity evolution of early-type galaxies across time (e.g., Bender et al. 1992; Jørgensen et al. 1996, 2006; Jørgensen 1999; Treu et al. 2005; van der Wel et al. 2005; Cappellari et al. 2006; van der Marel & van Dokkum 2007; Saglia et al. 2010; Jørgensen & Chiboucas 2013). At $z < 1$, the FP zero-point offset has been interpreted as the result of purely passive (without structural) evolution of the stellar population (e.g., Jørgensen et al. 2006; Jørgensen & Chiboucas 2013). However, this is likely not the case at $z > 2$ where the red and quiescent galaxies are compact and must undergo significant size evolution to evolve into the sizes of the present-day galaxies.

Spectroscopic observations which are required to measure stellar velocity dispersions at $z > 2$ are, however, time-expensive and only the rarest, brightest, and most massive systems have been studied at this distance (van de Sande et al. 2013; Belli et al. 2017) using large cosmological fields like CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) and COSMOS (Scoville et al. 2007).

We present in this paper the FP study at $z > 2$ using a sample of massive field quiescent galaxies (MQGs) introduced in Stockmann et al. (2020) (hereafter S20). S20 find a shallow stellar velocity dispersion evolution and significant size growth between $z = 2$ and 0. In this paper, we explore whether this size growth, alongside the passive evolution of the stellar population, can account for the observed evolution of massive galaxies in the scaling relations from $z = 2$ to the present day.

In Section 2, the $z > 2$ MQG sample from S20 together with a complementary quiescent galaxy sample at a similar redshift is presented alongside two local samples from the Coma cluster and the MASSIVE Survey. We present the M/L and FP scaling relations in Section 3.1 and 3.2, respectively. The predicted evolution of the size, stellar velocity dispersion, passive aging, and luminosity increase due to minor merger-driven growth are presented in Section 4. Finally, our results are interpreted and discussed in Section 5, following a summary of the main conclusions in Section 6.

Throughout the text, magnitudes are quoted in the AB system (Oke & Gunn 1983; Fukugita et al. 1996) and the following cosmological parameters, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, with $H_0 = 70$ km s^{-1} Mpc^{-1}, are used. All stellar masses are presented using the Chabrier (2003) initial mass function (IMF).

2. Data

2.1. A Sample of MQGs at z > 2

In S20, we presented a sample of MQGs at $z > 2$ studied with the X-shooter spectrograph (D'Odorico et al. 2006; Vernet et al. 2011) at the Very Large Telescope and Hubble Space Telescope (HST) H_{F160W} that are crucial to obtain both rest-frame optical stellar velocity dispersions and effective sizes. The sample is selected from the 2 square degree COSMOS field, using multi-wave band photometric fits (Muzzin et al. 2013). In summary, the sample is selected to be K-band bright and massive ($\log(M_*/M_\odot) > 11$) UVJ quiescent galaxies at $z > 2$. The adopted stellar population parameters such as mass-weighted age and stellar mass were based on the COSMOS15 photometry (Laigle et al. 2016), the continuum emission modeling of the X-shooter spectra, and the choice of star formation history (see details in S20, Section 4.3). Here we consider eight of the total 15 galaxies from S20 with measured stellar velocity dispersions, which is essential to study them in the scaling relations. We find no selection bias when comparing the size, age, stellar mass, and redshift of this sample to the parent sample in S20.

In Table 1, we list the mass-weighted age, stellar mass, stellar velocity dispersions, and sizes for these eight galaxies. The adopted stellar velocity dispersion will be referred to as the effective stellar velocity dispersion in this study due to minimal correction (<5%) when following the equations based on X-shooter observations in van de Sande et al. (2013). For further details, we refer to S20. Contrary to S20, we compute the dynamical masses in this paper using the circularized sizes ($r_{e,circ} = \sqrt{ba}$)[14] to make them consistent with the dynamical masses derived using the circularized radius from the study of local cluster scaling relations in previous work such as, e.g., Jørgensen & Chiboucas (2013). We have verified that the qualitative results from S20 remain when using circularized dynamical masses and sizes. Following local studies (e.g., Jørgensen 1999), we adopt the Bessel B-band luminosity, estimated in our case from the rest-frame fluxes obtained from the most recent COSMOS photometry (Laigle et al. 2016) using the photometric redshift code EAZY

[14] Here a and b are the semimajor and -minor axes.

(Brammer et al. 2008).[15] Normally, the magnitudes (luminosities) from the effective radius profile fitting are used when studying the scaling relations. We adopted another approach to, instead, extract the COSMOS magnitudes to treat the high-redshift samples (S20 and the galaxies from Section 2.2) consistently. We confirm that these magnitudes, when compared to the H_{F160W} magnitudes from the profile fits, can be considered representative with negligible differences. As a result, the method here is similar to the standard approach. The luminosity and average effective surface brightness are estimated using the method outlined in Appendix A. Hereafter, the dynamical mass-to-light ratio in the Bessel *B*-band is referred to as M/L. The eight galaxies in our sample have a mean age of ∼1.5 Gyr and span a similar stellar mass and size range to the full 15 galaxy parent sample from S20.

2.2. Complementary Sample of Quiescent Galaxies at $1.5 < z < 2.5$

In addition to the galaxies from S20, we adopt a sample of $1.5 < z < 2.5$ quiescent galaxies from Belli et al. (2017). We choose 11 out of 24 galaxies with stellar velocity dispersions, which have available COSMOS photometry in Laigle et al. (2016) to ensure consistent photometry extractions similar to the S20 sample. Out of the 11 galaxies, seven are at $z < 2$ and four at $z > 2$. The 11 galaxies from Belli et al. (2017) introduced here will be referred to as B17.

From Belli et al. (2017), we adopt the effective semimajor axis, absorption line-measured stellar velocity dispersions, and dynamical masses (see their Table 2). The effective semimajor axis are derived, similar to S20, using Sérsic profile fits to the rest-frame optical HST/WFC3 $H_{\rm F160W}$ images using *GALFIT* (Peng et al. 2002) and converted to circularized radius. We know from the S20 sample that the COSMOS magnitudes in the H-band are a good approximation of the magnitudes in the above profile fit method and thus indirectly related to the effective radii. The stellar velocity dispersions are derived from rest-frame optical Multi-Object Spectrometer For Infra-Red Exploration (McLean et al. 2012) spectra using Penalized PiXel-Fitting (Cappellari & Emsellem 2004) to fit Bruzual & Charlot (2003) (hereafter BC03) stellar population models which are qualitative similar to S20. The dynamical masses are converted from the semimajor axis to the circularized radius using the axis ratio. As in S20, the dynamical mass is estimated using the method from Cappellari et al. (2006).

The two $z \sim 2$ samples (Section 2.1 and 2.2) used here satisfy the *UVJ* quiescent galaxy selection from Muzzin et al. (2013). Together they allow for exploring a larger dynamical mass range, $10.5 < \log(M_{\rm dyn}/M_{\odot}) < 11.9$, and redshift, $1.5 < z < 2.5$, corresponding to a cosmological time span of ∼1.7 Gyr, considered short compared to their consequent 8-10 Gyr of evolution to the present day.

2.3. The MASSIVE Survey

The original volume-limited MASSIVE Survey sample is selected as the most massive and *K*-band brightest early-type galaxies within 108 Mpc of the northern hemisphere (Ma et al. 2014). Here, we use the 25 most massive, $\log(M_*/M_{\odot}) > 11.7$, MASSIVE galaxies (hereafter MASSIVE(n) sample), selected at fixed cumulative number density (CND) matching our massive $z \sim 2$ sample as described in S20. The CND of the $z \sim 2$ sample is estimated from the massive, $\log(M_*/M_{\odot}) > 11.2$, *UVJ* quiescent galaxies at $1.9 < z < 2.5$ in the Muzzin et al. (2013) catalog. The results in S20 are shown to be robust against the choice of CND method (fixed and probabilistic; Wellons & Torrey 2017), as well as the mass–rank scatter. Utilizing this CND approach is an attempt to minimize progenitor bias by predicting the local progenitors of the high-z galaxies. A thorough discussion of the assumptions and uncertainties is covered in S20 Section 5.1.

The magnitudes for 17/25 galaxies are obtained from the SDSS DR14 catalog photometry (Blanton et al. 2017) by cross-matching the MASSIVE(n) sample using the SDSS SkyServer.[16] The *de Vaucouleur* photometry ("deVMag") in the *u, g, r, i, z* bands is extracted and converted to rest-frame Bessel *B*-band magnitude using the EAZY code, in the same way as in the high-redshift samples in this study. The "cmodelMag" estimate, also used for the Coma sample (Section 2.4), is compared with the *de Vaucouleur* fit magnitudes to establish that the latter is a good representation, which is also confirmed by the "fracDeV" parameter. The luminosity and effective surface brightness are calculated from the apparent magnitude using the methods covered in Appendix A. Instead of estimating the luminosity distance from the redshift, these galaxies are close enough that peculiar velocities have a significant impact on their distance measurement. We therefore use the distance measurement from Ma et al. (2014), who correct for this effect. The 17 galaxy subsample used here is referred to as MASSIVE(n_{17}).

The *de Vaucouleur* effective radii ("deVRad") corresponding to extracted photometry from the SDSS SkyServer are used. After confirming that the radii are consistent among the different bands, we choose to use the *g*-band effective radii as the wavelength coverage is comparable to the Bessel *B*-band. The radii are circularized using the axis ratio ("deVAB"). The average luminosity weighted dispersion within the effective radius is adopted (Veale et al. 2018). The dynamical masses are estimated using the method in S20 (with the prescription from Cappellari et al. 2006) using $n = 4$ and the circularized effective sizes.

The stellar mass, size, and stellar velocity dispersion between the MASSIVE(n) and MASSIVE(n_{17}) samples are compared in Appendix B. Here we find that the MASSIVE(n_{17}) sample is uniformly sampled from the initial distribution of stellar mass, size, and stellar velocity dispersion and is 68% complete. As a result, our MASSIVE(n_{17}) selection with available photometry is representative of the parent CND matched sample and can be considered a suitable minimal progenitor-biased reference sample.

2.4. Coma Cluster

As a local reference cluster we use Coma/A1656 complete to $g'_{rf} \leqslant 16.1$ magnitudes, as used in Jørgensen et al. (2019). The scaling relations of the Coma cluster are well studied and we include them as a reference to other high-redshift comparisons. We adopt the 123 galaxies with average effective surface brightness, stellar velocity dispersion, and circularized sizes. The stellar velocity dispersion measurements (from Jørgensen et al. 2018) are derived from high signal-to-noise ratio (∼60 $\mathring{A}^{-1}$) spectra using the recipe presented in Jørgensen et al. (2017).

[15] https://github.com/gbrammer/eazy-photoz

[16] http://skyserver.sdss.org/dr14/en/tools/crossid/crossid.aspx

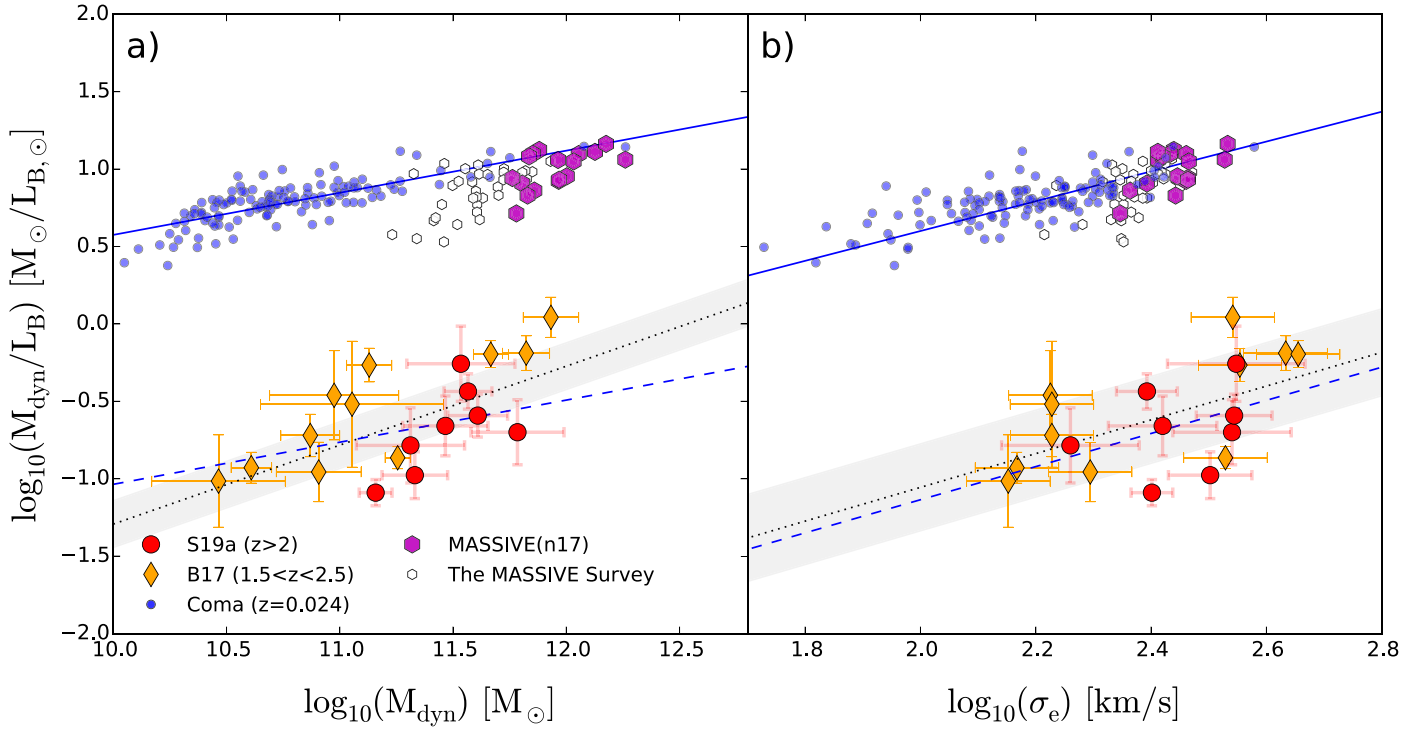

Figure 1. M/L as a function of dynamical mass (a) and stellar velocity dispersion (b) for the S20 massive quiescent galaxies at $z > 2$ (red symbols) and 11 COSMOS quiescent galaxies at $1.5 < z < 2.5$ from B17 (orange symbols). The Coma galaxies ($z = 0.0231$, blue symbols) are shown together with the best-fit relation (blue line) when assuming Jørgensen & Chiboucas (2013) Fundamental Plane parameters. The MASSIVE Survey galaxies with available rest-frame *B*-band photometry (all white and purple hexagons) and the cumulative number density-matched MASSIVE(n_{17}) sample (purple hexagons only) are shown. The best-fit (black dotted line) and rms (gray shading) of the combined high redshift samples (S20 and B17) are shown together with the Coma best-fit relation offset to the median M/L at $z \sim 2$ (dashed blue).

The surface brightness is calculated using rest-frame Bessel B magnitudes based on the SDSS *cmodelmag* magnitude (see Appendix A in Jørgensen et al. 2019). The circularized effective radii are constructed from SDSS parameters to be a pseudo-Sérsic effective radii with the purpose of matching the "cmodelMag" magnitudes (see the method in Jørgensen et al. 2019). Both data sets are calibrated to the *Legacy data* (Jørgensen 1999; Jørgensen & Chiboucas 2013) to provide a trustworthy low-redshift reference cluster. For further details on the physical properties of the Coma cluster, we refer to Section 2 in Jørgensen et al. (2018).

3. Results

In this section, we present the $M_{\rm dyn}/L_B$ and FP scaling relations as a tool to study the evolution of MQGs from $z = 2$ to 0. We do so by fitting these relations at $z \sim 2$ and comparing them to two local samples presented in Section 2.

3.1. Dynamical Mass-to-light Ratio, M_{dyn}/L_B

$M_{\rm dyn}/L_B$ reveals information about how the dynamical structure of galaxies compares to the luminosity of their stellar populations. This ratio has been found to increase in massive quiescent galaxies, believed to be driven by the non-star-forming and passive evolution of their stellar population. Here, we study the $M_{\rm dyn}/L_B$ relations for MQGs and their evolution in the past 10 billion years. We show these relations in Figure 1, both as a function of dynamical mass (1(a)) and stellar velocity dispersion (1(b)), for a sample of $z \sim 2$ MQGs (S20 + B17) alongside two local reference samples (Coma, MASSIVE(n_{17})).

Compared to the local Coma and MASSIVE(n_{17}) galaxies, the high-redshift samples (S20 and B17) have lower *M/L* as expected for brighter younger systems. The majority of the galaxies from S20 have, already at $z = 2$, dynamical masses similar to the ∼10% most massive galaxies in the Coma cluster (see Figure 1(a)). The MASSIVE(n_{17}) galaxies have dynamical masses similar to the 2% most dynamical massive Coma galaxies, which shows that they are among the most massive galaxies in the local universe. The distribution of the stellar velocity dispersions of S20 and B17 is consistent with the high-end measurements in Coma (see Figure 1(b)). When comparing the stellar velocity dispersions between the S20 and the MASSIVE(n_{17}) sample we find that a shallow evolution is expected between $z = 2$ and the present day. The MASSIVE(n_{17}) sample have high stellar velocity dispersions similar to the high end of the Coma cluster measurements.

The combined samples of S20 and B17, in Figure 1(a) and (b), are fit by minimizing the least-squares in the *y*-direction and the uncertainty on the slopes is estimated using a bootstrap procedure (see also Jørgensen et al. 1996; Jørgensen & Chiboucas 2013). A relation for the combined high-redshift (S20 and B17) sample is established, while this was not possible using the narrow dynamical mass range of the S20 sample alone. The fits are shown in Figure 1, together with the associated rms from the regression, and listed in Table 2. M/L with dynamical mass relation appears slightly steeper, although by less than 2σ. For both M/L versus $M_{\rm dyn}$ and M/L versus σ_e, we find best-fit slopes for the combined high-redshift sample to be consistent with the local Coma relation (see relations 4 and 6 in Table 2).

Table 2
Fundamental Plane and M/L Scaling Relations

No.	Sample	Relation	$N_{\rm gal}$	rms	σ_γ
(1)	Coma[a]	$\log r_e = (1.30 \pm 0.08)\log\sigma - (0.82 \pm 0.03)\log\langle I\rangle_{e,\rm B} - 0.443$	123	0.08	0.01
(2)	Combined sample ($1.5 < z < 2.5$)[b]	$\log r_e = (0.46 \pm 0.18)\log\sigma - (0.46 \pm 0.07)\log\langle I\rangle_{e,\rm B} + 1.275$	19	0.15	0.03
(3)	Coma[a]	$\log M/L = (0.27 \pm 0.03)\log M_{\rm dyn} - 2.145$	123	0.09	0.01
(4)	Combined sample ($1.5 < z < 2.5$)[b]	$\log M/L = (0.51 \pm 0.15)\log M_{\rm dyn} - 6.393$	19	0.26	0.06
(5)	Coma[a]	$\log M/L = (0.96 \pm 0.10)\log\sigma - 1.326$	123	0.10	0.01
(6)	Combined sample ($1.5 < z < 2.5$)[b]	$\log M/L = (1.09 \pm 0.41)\log\sigma - 3.236$	19	0.28	0.06

Notes. Column 1: sample; Column 2: fitting method; Column 3: scaling relations; Column 4: number of galaxies included in fit; Column 5: rms scatter along the y direction of the scaling relation; Column 6: uncertainty on the zero-point (γ) estimated as $\rm rms/\sqrt{N_{gal}}$ (assuming fixed coefficients).
[a] Fits are from Jørgensen et al. (2019) Table 4.
[b] Samples from S20 and B17 (see Section 2)

3.2. The FP

The FP is spanned by the effective size, r_e, stellar velocity dispersion, σ_e, and average effective surface brightness, $\langle I\rangle_{e,\rm B}$. Its edge-on and face-on orientations are defined as

$$\log_{10} r_e = \alpha \log_{10}\sigma_e + \beta \log_{10}\langle I\rangle_{e,\rm B} + \gamma \quad (1)$$

and

$$(2.22\log_{10} r_e + \beta\log_{10}\langle I\rangle_{e,\rm B} + \alpha\log_{10}\sigma_e)/2.7 = (\alpha\log_{10}\langle I\rangle_{e,\rm B} - \beta\log_{10}\sigma_e)/1.54, \quad (2)$$

respectively. The best-fit Coma relation slopes ($\alpha = 1.30 \pm 0.08$, $\beta = -0.82 \pm 0.03$) and zero-point ($\gamma = -0.443$) in the rest-frame B-band are adopted as our local reference orientation of the plane (Jørgensen et al. 2006; see also Table 2).

The best-fit Coma data, from Jørgensen et al. (2019), were fit using the same method as described below while adopting the FP parameters from Jørgensen & Chiboucas (2013) ($\alpha = 1.30 \pm 0.08$, $\beta = -0.82 \pm 0.03$) to get the zero-point ($\gamma = -0.443$). We use this as our local reference cluster fit and list the parameters along with their uncertainties in Table 2.

In Figure 2, the FP edge-on and face-on projections, as described in Equations (1) and (2), are shown. We use this to examine how $z \sim 2$ MQGs populate and evolve to $z = 0$ in this plane. In the edge-on FP, the dominating errors from the stellar velocity dispersion are shown on the y-axis. For the face-on plane, the errors are calculated similarly to the approximation used in Jørgensen et al. (2006).

The COSMOS quiescent galaxies from S20 and B17 are found to be below (in the edge-on plane) and above (in the face-on plane) the local Coma FP relation. These galaxies have compact sizes and younger stellar populations (due to their high redshift and more recent quenching), effectively increasing their mean effective surface brightness.

An edge-on FP cannot be clearly established using the S20 sample alone. However, when fitting the S20 and B17 samples together, an FP is in place at $1.5 < z < 2.5$ (hereafter referred to as $\rm FP_{z\sim2}$). The FP is fitted using the least-squares method, minimizing the least-squares in the y-direction, with uncertainties from a bootstrapping method (see relation 2 in Table 2). In van de Sande et al. (2014), an FP was indicated for a similar epoch ($1.5 < z < 2.5$). However, in this study, the sample of $z > 2$ galaxies is three times more numerous, robustly confirming the existence of a plane at $1.5 < z < 2.5$.

An $\rm FP_{z\sim2}$ is established using 19 massive quiescent galaxies that as a result must have been a relatively homogeneous population already at this epoch, signaling even earlier formation and significant evolution from $z = 2$ to the present day.

4. Evolution of the Scaling Relations

In Section 3 we established an FP for MQGs at $1.5 < z < 2.5$. Here, we explore how the $z = 2$ MQGs from S20 evolve through these scaling relations to the CND-matched minimal progenitor-biased (Section 2.3) local MASSIVE($\rm n_{17}$) sample. The $z = 2$ MQGs were, in S20, shown to undergo structural ($\Delta\log r_{e,circ} \sim 0.6$) and stellar mass ($\Delta\log M_* \sim 0.3$) evolution with a stellar-to-dynamical mass ratio, $\Delta\log_{10} M_*/\Delta\log_{10} M_{\rm dyn} \sim 0.5$, from $z = 2$ to 0. These effects were in S20 suggested to arise from minor merger-driven size growth. Additionally, the luminosity is expected to change due to passive evolution of the stellar population and with the addition of new stellar mass from the minor merger-driven size evolution. We adopt the structural and dynamical mass evolution from S20 and in Sections 4.1 and 4.4 model the aforementioned luminosity evolution components. Combining these effects we explore whether they show a consistent picture of evolution when analyzed in the scaling relations.

Finally, to highlight how the different physical mechanisms affect the evolution of the galaxies, throughout the scaling relations, Equation (8) is derived. The evolution of the FP parameters can be formalized (Saglia et al. 2010, 2016) under the assumption of homology, where α, β are constant over time (Beifiori et al. 2017):

$$\Delta\log_{10} L = \frac{1+2\beta}{\beta}\Delta\log_{10} r_e - \frac{\alpha}{\beta}\Delta\log_{10}\sigma_e - \frac{\Delta\gamma}{\beta}. \quad (3)$$

The logarithmic difference is defined by $\Delta\log_{10} X = \log_{10}(X_{z=2}/X_{z=0})$ where X, in this case, is either the luminosity, size, or dispersion. The zero-point evolution is described by $\Delta\gamma = \gamma_z - \gamma_{z=0}$. This leads to the relation between the zero-point evolution and the average change in M/L, $\Delta\log_{10} M/L = \Delta\gamma/\beta$. We express the change in M/L by effective size, stellar velocity dispersion, and luminosity evolution:

$$\Delta\log_{10} M/L = \frac{1+2\beta}{\beta}\Delta\log_{10} r_e - \frac{\alpha}{\beta}\Delta\log_{10}\sigma_e - \Delta\log_{10} L_{\rm total}. \quad (4)$$

The change in effective size and stellar velocity dispersion are adopted from S20 as mentioned above. The total luminosity contribution from $z = 2$ to 0 can be described by both the

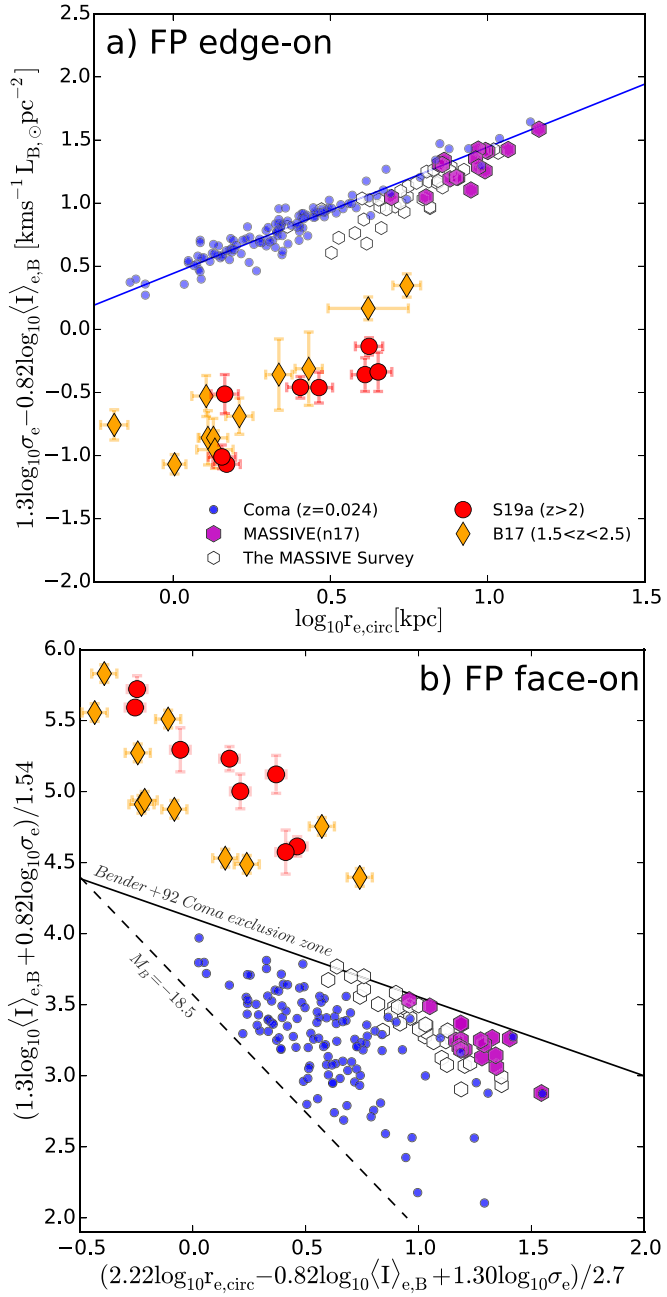

Figure 2. The Fundamental Plane (FP) projected edge-on (a) and face-on (b) with the symbols as in Figure 1. The exclusion zone for local spheroidal galaxies from Bender et al. (1992) and the Coma luminosity limit from Jørgensen et al. (2006) are shown in the face-on plane. An $FP_{z\sim2}$ is established when combining the COSMOS quiescent galaxies from S20 and B17.

luminosity change due to passive evolution, $\Delta \log_{10} L_{\rm passive}$, and that associated with the newly incorporated stellar mass within the effective radius, $\Delta \log_{10} L_{M_*}$,

$$\Delta \log_{10} L_{\rm total} = \Delta \log_{10} L_{\rm passive} + \Delta \log_{10} L_{M_*}. \quad (5)$$

$\Delta \log_{10} L_{M_*}$can be expressed in terms of the stellar mass and $M_* - L$ evolution:

$$\Delta \log_{10} L_{M_*} = \Delta \log_{10} M_* - \Delta \log_{10} \frac{M_*}{L} \quad (6)$$

$$= x \Delta \log_{10} M_{\rm dyn} - \Delta \log_{10} \left(\frac{M_*}{L} \right)_{M_*}. \quad (7)$$

Here $x = \Delta \log_{10} M_* / \Delta \log_{10} M_{\rm dyn}$ and $\Delta \log_{10} M_{\rm dyn}$ are adopted from S20 (see the first paragraph). The total change in M/L can thus be expressed in terms of the passive (Section 4.1), structural (Section 4.2), stellar-to-dynamical mass ratio (Section 4.3), and luminosity contribution from the stellar mass increase due to minor merger-driven size growth (Section 4.4),

$$\begin{aligned} \Delta \log_{10} M/L = & \frac{1+2\beta}{\beta} \Delta \log_{10} r_e - \frac{\alpha}{\beta} \Delta \log_{10} \sigma_e \\ & - \Delta \log_{10} L_{\rm passive} \\ & - x \Delta \log_{10} M_{\rm dyn} + \Delta \log_{10} \left(\frac{M_*}{L} \right)_{M_*}. \end{aligned} \quad (8)$$

We show the effect of the mean size and dispersion evolution from S20 and the modeled effect on the luminosity in Figure 3 and in Appendix C. The effects of passive (red) and structural (blue) evolution with the effect of the stellar-to-dynamical mass ratio (black) inclusion, and the luminosity increase from the stellar mass (green), are shown as arrows from the median of the S20 sample of $z = 2$ MQGs. The effects of passive and structural evolution together with the change in dynamical-to-stellar mass ratio present a realistic scenario of how these S20 MQGs could evolve into local galaxies.

4.1. Passive Evolution

The expected passive evolution of the stellar population, from the redshift of formation to the present day, is estimated based on the evolution in M_*/L of a BC03 simple stellar population model with Chabrier IMF and solar metallicity. The formation redshift is estimated from the mass-weighted age and the redshift of observation. The passive evolution takes into account mass loss during stellar evolution, and the M_* is thus the mass locked into stars at a given age. The mass-weighted mean stellar age is indicative of the epoch by which the bulk of the stellar mass was formed. The M_*/L uncertainties are dominated by the 1σ standard deviation of the age ($\sim$0.2 dex) when assuming solar metallicity. The definition of passive evolution assumes that no new stellar mass is added and thus M_*/L directly relates to the $\Delta \log_{10} L_{\rm passive}$ from Equation (8).

The average estimate of the passive evolution is shown by the red arrow in Figure 3. This is what is expected as Equation (8) reduces to $\Delta \log_{10} M/L = \Delta\gamma/\beta = -\Delta \log_{10} L_{\rm passive}$ for no structural ($\Delta \log_{10} r_e = \Delta \log_{10} \sigma_e = 0$) and no luminosity change (from a stellar mass increase). The passive evolution cannot fully explain the direct evolution, of MQGs from $z = 2$ to 0, to Coma galaxies at fixed dynamical mass (or size). Furthermore, it is inadequate in alone explaining the expected evolution to the MASSIVE(n_{17}) sample.

4.2. Structural Evolution

We learned from S20 that the size growth with shallow stellar velocity dispersion change was driven by minor merger structural evolution. Using this information we can

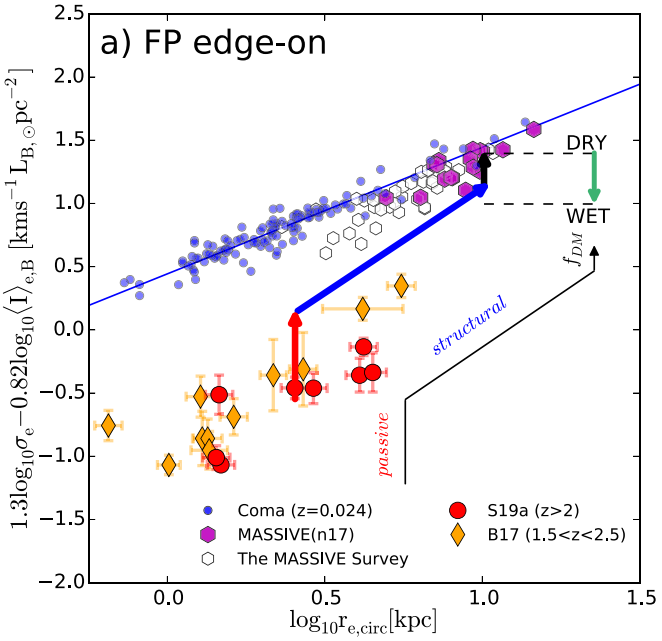

Figure 3. FP projected edge-on (as portrait in Figure 2) shown with the predicted evolution (see Section 4) of the S20 sample from $z = 2$ to 0. The average decrease in luminosity from passive evolution of the stellar population (Section 4.1) is shown by the red arrow. The size increase from structural evolution due to minor mergers is shown as the blue arrow (Section 4.2). The black arrow shows the effect of a changing $M_*/M_{\rm dyn}$ from 1 to 0.5 (Section 4.3). The green arrow represents the surface brightness change for dry or wet minor merger-driven structural evolution (see Section 4.4). The massive quiescent galaxies from S20 are consistent with evolving, via both passive and primarily dry minor merger structural evolution when taking the change in stellar-to-dynamical mass ratio into account, into the local most massive Coma galaxies and the cumulative number density-matched minimal progenitor biased MASSIVE(n$_{17}$) sample. These evolutionary trends are shown in Figure C.1, incorporating Figures 1 and 2, in Appendix C.

simplify Equation (8):

$$\Delta \log_{10} M/L = \frac{\Delta\gamma}{\beta} = \frac{1+2\beta}{\beta}\,\Delta \log_{10} r_e. \tag{9}$$

The effect of the structural evolution adopted from S20 is based on the size increase from the stellar mass–size plane, their Figure 8 ($\Delta \log r_{e,circ} \sim 0.6$), alongside insignificant stellar velocity dispersion evolution ($\Delta \log \sigma_e \sim 0$). This effect is shown in the edge-on FP in Figure 3 as the sloped blue arrow. The combined effect of the mean passive evolution of the stellar population (red arrow) and mean the structural evolution from minor merger-driven size increase (blue arrow) brings the S20 MQGs in the vicinity of the MASSIVE(n$_{17}$) sample.

4.3. Stellar-to-dynamical Mass Ratio

S20 found that the median stellar-to-dynamical mass ratio $x = \Delta \log_{10} M_* / \Delta \log_{10} M_{\rm dyn} \sim 0.5$. Specifically the stellar mass doubles ($\Delta \log M_* \sim 0.3$ dex) from $z = 2$ to 0. Following Equation (8), we show how this ratio affects the evolution of the scaling relations when not unity:

$$\Delta \log_{10} M/L = \frac{\Delta\gamma}{\beta} = x\,\Delta \log_{10} M_{\rm dyn}. \tag{10}$$

Here the effect of a changing stellar-to-dynamical mass ratio can be seen to directly affect the change in the scaling relations. The effect for $x = 0.5$ is shown in Figure 3. When taking into account the stellar-to-dynamical mass ratio, which affects the conversion between the stellar and dynamical M/L, we obtain a predicted evolution closer to the high-mass end of the Coma relation and the median position of the MASSIVE(n$_{17}$) sample.

4.4. Luminosity Increase from Wet Minor Merger Stellar Populations

From S20 we learned that $z = 2$ MQGs grow their size and stellar mass through minor mergers in their evolution to $z = 0$. In addition to the decrease in the B-band luminosity, due to passive evolution, it is expected that such merger events could add to the B-band luminosity if these are star-forming galaxies at the time of merging. From now on, this type of merger is referred to as "wet," contrary to the "dry" minor mergers that are passive before merging. In Equation (8) this reduces to the effect of the stellar M/L change from new stars accreted from minor mergers:

$$\Delta \log_{10} M/L = \Delta \log_{10}\left(\frac{M_*}{L}\right)_{M_*}. \tag{11}$$

The B-band luminosity increase from merging galaxies between $z = 2-0$ is modeled with composite stellar population models from the BC03 library with solar metallicity to obtain the stellar M/L (similar to Section 4.1). The star formation history follows the evolution of the main sequence (Speagle et al. 2014). It is assumed that, after merging, the galaxies stop forming stars and follow a passive evolution.

The median stellar mass increase ($\Delta \log M_* = 0.3$ dex) from minor mergers, predicted in S20, is used assuming a 1:20 merger ratio. Note that in our simplistic model, the correct mass ratio does not play a significant role. We also investigate a more realistic scenario with mergers distributed across redshift (z = 1.8–0.1) together with two extreme cases of all the mass added at $z = 1.8$ or 0.1. A B-band luminosity increase of 0.4–0.45 dex is found in all cases (for more details see Appendix D).

In Figure 3 the effect of increasing luminosity, due to wet minor merger stellar populations, is shown by the green arrow (0.4 dex). This effect appears to be in disagreement with the location of the Coma relation and the MASSIVE(n$_{17}$) sample. This suggests that dry minor merger galaxies, with no additional luminosity increase, are a preferred evolution scenario for MQGs.

5. Discussion

5.1. Passive Evolution of MQGs from z = 2 to 0

Studies of passive galaxies in $0.8 < z < 1.8$ clusters (among others Jørgensen et al. 2006, 2019; van der Marel & van Dokkum 2007; Jørgensen & Chiboucas 2013; Beifiori et al. 2017) find that the change in M_*/L can be explained by passive evolution to $z = 0$. Below we explore if a similar analysis can account for the evolution of the scaling relations at $z \sim 2$.

Jørgensen & Chiboucas (2013) derived the $M_{dyn}/L_{\rm B}$ evolution (based on models from Maraston 2005), as a function of stellar age and metallicity, to be $\log M_{dyn}/L_{\rm B} = 0.935 \log \text{age} + 0.337[\text{M/H}] - 0.053$. Assuming passive evolution from $z = 2$ to the best-fit Coma relation, at $\log(M_{\rm dyn}/M_\odot) = 11.5$, we find a formation redshift of $z_{\rm form,Coma} = 2.01^{+0.1}_{-0.04}$ (for details see Appendix E). The formation redshift is similar to the redshift of observation,

which leaves too short a time to form the S20 MQGs at this epoch. The formation redshift, derived from our stellar population mass-weighted ages (assuming the median age), is $z_{\rm form} = 3.41^{+4.92}_{-0.91}$. The uncertainties are estimated using the 1σ age uncertainties. Based on this, we conclude that the S20 MQGs at $z > 2$ cannot evolve to the Coma relation by passive evolution alone.

5.2. Minor Merger-driven Structural Evolution of MQGs

The fixed CND-matched MASSIVE(n_{17}) sample allows us to study the evolution of massive galaxy scaling relations from $z = 2$ to 0 with minimal progenitor bias. Evidence against purely passive evolution to $z = 0$ is present in both the scaling relations (Figures 3 and C1) as explicitly shown in the previous section. The S20 MQGs at $z > 2$ are consistent with evolving into the most massive Coma galaxies, and the MASSIVE(n_{17}) sample, through passive, structural, and stellar-to-dynamical mass ratio evolution.

The size increase of MQGs in cosmological simulations could be explained by adiabatic expansion due to AGNs, decreasing the central mass density and puffing up the galaxies (Dubois et al. 2013; Choi et al. 2018). Major mergers, as the dominant mechanism for size growth, have become less popular as they make the galaxies too massive to be consistent with massive nearby galaxies. In S20, the structural evolution is interpreted to be from minor mergers in line with the scenario presented in the idealized simulations from Hopkins et al. (2009), Naab et al. (2009), and Hilz et al. (2012, 2013). Here, the effective half-light radius grows by adding stars to the outskirts of the galaxy from tidally stripped minor mergers. This scenario is shown to cause inside-out growth, starting from a compact elliptical galaxy (core) that causes, through minor mergers, a build-up of the surface density profile wings, a present-day analog of a giant elliptical galaxy (core–envelope). A consequence of the inside-out minor merger growth scenario from Hilz et al. (2012) is an increasing dark matter fraction which has been suggested to cause a tilt in the FP over time (Boylan-Kolchin et al. 2005; D'Onofrio et al. 2013). Essentially this is a consequence of using the effective radius as a foundation for analyzing the evolution, as this results in the stellar velocity dispersion mainly tracing the stars at $z = 2$, but at $z = 0$ a higher fraction of dark matter to stars. This effectively introduces a systematic difference between comparing stellar velocity dispersions across epochs that only trace stars with a similar sample where the stellar-to-dynamical mass ratio evolves. This systematic difference, included in Section 4.3, must be taking into account and can be seen as the black arrow in Figures 3 and C1.

5.3. Dry Minor Merger Evolution

In Section 4.4, the predicted luminosity increase from wet minor mergers is modeled under the assumption that they are the primary drivers of the size growth. For the realistic scenario of adding wet minor mergers continuously from $z = 2$ to 0, we find that the luminosity increases by roughly 0.4 dex.

In Figures 3 and C1, the predicted position of the S20 MQGs (following passive, structural, and stellar-to-dynamical mass evolution) is indicated alongside the effect of the luminosity from wet minor mergers. The green arrow indicates how the predictions would move compared to the local best-fit relation of Coma and around the locus of the MASSIVE(n_{17}) sample, strongly favoring dry minor mergers.

In the inside-out growth scenario, the rest-frame B-band luminosity increase takes place in the outer parts of the galaxy. The luminosity from the MASSIVE(n_{17}) sample is measured using SDSS deVMag which represents the luminosity of the galaxy out to $8r_e$. Thus, an underestimation of the luminosity, by only sampling the central part of the galaxy and missing the outskirts, is unlikely.

The wet minor merger luminosity increase offsets the predictions from the local relation and the MASSIVE(n_{17}) sample and thus appears to not be a favored way to grow MQGs at $z \sim 2$. Another possibility is that the minor mergers already have quenched stellar populations (before their merger) with low rest-frame B-band luminosity (Oogi & Habe 2013; Naab et al. 2014; Tapia et al. 2014). The evolution from $z = 2$ to 0 of the FP and M/L scaling relations is consistent with such a scenario, caused primarily by dry minor mergers, and passive and stellar-to-dynamical mass ratio evolution for MQGs.

5.4. Caveats

Data from S20 and B17 are combined to establish the $FP_{z\sim2}$ at $1.5 < z < 2.5$ with more than half of the sources at $z > 2$. A large fraction of the quiescent galaxies from B17 is found to be disk-like (based on Sérsic index, $n < 2.5$; Belli et al. 2017), which could mean that an unknown contribution from rotation is included in the measured stellar velocity dispersion. For spherical dispersion-dominated systems, the circularized radius and semimajor axis are comparable methods of size measurement. However, for more disk-like systems the difference grows between the two size measuring methods, further causing a bias between dispersion- and rotation-dominated galaxies. We estimate, based on the axis ratios, that the circularized sizes differ by 7%–30% compared to the semimajor axes. This is well within the quoted uncertainties of the predicted position of the S20 MQGs at $z = 0$. This bias could potentially affect the zero point and coefficients of the best-fit in Figures 1 and 2. This issue could be solved by spatially resolved spectroscopy disentangling the contribution from rotation and dispersion.

The SDSS modelMag luminosities of the most luminous galaxies have been underestimated (Bernardi et al. 2017). In our study, the deVMag and cmodelMag methods have been used to estimate the luminosity of the MASSIVE and Coma sample, respectively. The median offset for the brightest galaxies ($M_r \sim -24$), in the r-band, are $\sim$15% (see Figure 7 in Bernardi et al. 2017). This translates to a difference in $\Delta \log_{10} L \propto \log_{10}(10^{0.4\Delta M_r}) = 0.4 \cdot 0.15 = 0.06$, assuming that the r-band magnitudes are representative for the Bessel B-band used in our study. In the case that this assumption is valid, the M/L and average effective surface brightness would change by 0.06 dex, thus moving the local galaxies in the positive y-direction, by the same amount, in Figures 1, 2, and 3. Such an effect is minimal and would not affect the general trends, results, and conclusions made in this paper.

The dominating uncertainty of the mass–rank scatter from the CND-matching of the local MASSIVE(n_{17}) sample does not affect the conclusions of this study. Furthermore, if using a probabilistic CND-matching approach (see Wellons & Torrey 2017), this would increase the number of galaxies in the MASSIVE(n_{17}) sample from 17 to 30. In Figures 1 and 2 this would correspond to a greater number of white hexagons

becoming purple, which causes no noticeable effects on the trends in the figures. On the other hand, if not all massive galaxies at $z > 2$ have similar merger histories, descendants that become quiescent systems at late times would have been missed (see, e.g., Naab et al. 2014). A study of the average stellar population age within the effective radius of the most massive ($\sigma_* > 220\,\mathrm{km\,s^{-1}}$) MASSIVE Survey galaxies (Greene et al. 2015) find their ages to be >10 Gyr. This suggest that our sample of MASSIVE(n_{17}) galaxies already were quenched at $z > 2$, and likely have similar merger histories from $z = 2$ to 0. Studies also show that the most massive end of the stellar mass function ($\log(M_*/M_\odot) > 11.5$) evolves very little, if at all, between $0 < z < 2$ (McLeod et al. 2020, their Figure 5).

6. Summary and Conclusion

In this work, we present the highest-redshift study of quiescent galaxy scaling relations with a sample size two times larger than previous studies at this redshift. The M/L of MQGs at $z > 2$ is observed to be ~30−40 times smaller than the local Coma relation (at fixed dynamical mass) and it requires significant passive luminosity evolution to match the $z = 0$ relation. In S20, the same galaxies are shown to undergo considerable structural evolution by quadrupling their sizes from $z = 2$ to 0, while their effective dispersion remains nearly unchanged. In this paper, the FP and M/L established scaling relations at $z \sim 2$, and the expected structural and passive evolution, are explored for the S20 MQGs from $z = 2$ to 0. The main conclusions of this study are listed below.

1. The FP and M/L relations are established at $z \sim 2$. Compared to the local Coma cluster and the CND-matched MASSIVE(n_{17}) sample, the quiescent galaxies at high redshift are found to be both compact and rest-frame B-band brighter, the latter due to more recently quenched stellar populations. The position of the MASSIVE(n_{17}) sample broadly agrees with the best-fit Coma relation for the most massive and largest galaxies (Figures 1 and 2).
2. Interpreting the M/L offset as purely passive evolution of the stellar population leads to a formation redshift of $z \sim 2$, lower than the formation redshift inferred from the stellar population analysis of S20 MQGs, $z_{\mathrm{form}} = 3.41^{+4.92}_{-0.91}$ (Section 5.1). As a result, the S20 MQGs are not consistent with their evolution into the local Coma FP and M/L scaling relations by passive evolution alone.
3. The S20 MQGs at $z \sim 2$ are consistent with passive, structural, and stellar-to-dynamical mass ratio evolution into the most massive Coma galaxies and the minimal progenitor-biased MASSIVE(n_{17}) sample (See Section 4 and Figures 3 and C1).
4. In the case that the observed size evolution can be attributed entirely to minor mergers, the FP and M/L evolutions are consistent with the accretion of dry minor merger stellar populations. A scenario of wet minor mergers increases the rest-frame B-band luminosity by 0.4 dex, inconsistent with the evolution of MQG at $z \sim 2$ into the local most massive Coma and the CND-matched MASSIVE(n_{17}) galaxies.

M.S., S.T., C.G., and G.B. acknowledge support from the European Research Council (ERC) Consolidator Grant funding scheme (project ConTExt, grant No. 648179). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. Based on observations made with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. Support for this work was provided by NASA through grant No. HST-GO-14721.002 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS 5-26555. M.S. thanks Nina Voit for her encouragement and unparalleled love, help, and support. This research made use of Astropy (version 1.1.1),[17] a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). I.J. is supported by the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., on behalf of the international Gemini partnership of Argentina, Brazil, Canada, Chile, the Republic of Korea, and the United States of America. M.H. acknowledges financial support from the Carlsberg Foundation via a Semper Ardens grant (CF15-0384). F.V. acknowledges support from the Carlsberg Foundation research grant CF18-0388 "Galaxies: Rise And Death."

Appendix A
Derivation of Luminosity and Effective Surface Brightness, $\langle I \rangle_{e,\mathrm{B}}$

The luminosity and average effective surface brightness are estimated by converting the EAZY (Brammer et al. 2008) calculated rest-frame B-band fluxes to apparent AB magnitudes (assuming no extinction correction) and calculating the absolute Vega magnitudes and luminosity by

$$M_{Vega,\mathrm{B}} = m_{Vega,\mathrm{B}} - 5 \cdot (\log(D_L/\mathrm{pc}) - 1)$$
$$\frac{L_{\mathrm{B,gal}}}{L_{\mathrm{B},\odot}} = 10^{-0.4(M_{Vega,\mathrm{B}} - M_{\odot,\mathrm{B}})}. \quad \text{(A1)}$$

Here the luminosity distance (D_L) and $M_{\odot,\mathrm{B}} = 5.45$[18] are used. The effective surface brightness in Bessel B-band is calculated as

$$\langle I \rangle_{e,\mathrm{B}} = \frac{L_{\mathrm{B,gal}}/L_{\mathrm{B},\odot}}{2\pi r_e^2}. \quad \text{(A2)}$$

Note that cosmological redshift dimming is included when converting the radius from arcseconds to parsecs.

Appendix B
How Representative is the MASSIVE(n_{17}) Sample?

To ensure that we do not introduce a bias we compare the MASSIVE(n_{17}) and the parent MASSIVE(n) samples' stellar mass, size, or stellar velocity dispersion. This is presented in Figure B1 where we show that the MASSIVE(n_{17}) sample is representative of the structural and kinematical parameters.

[17] http://www.astropy.org
[18] http://mips.as.arizona.edu/~cnaw/Sun.html

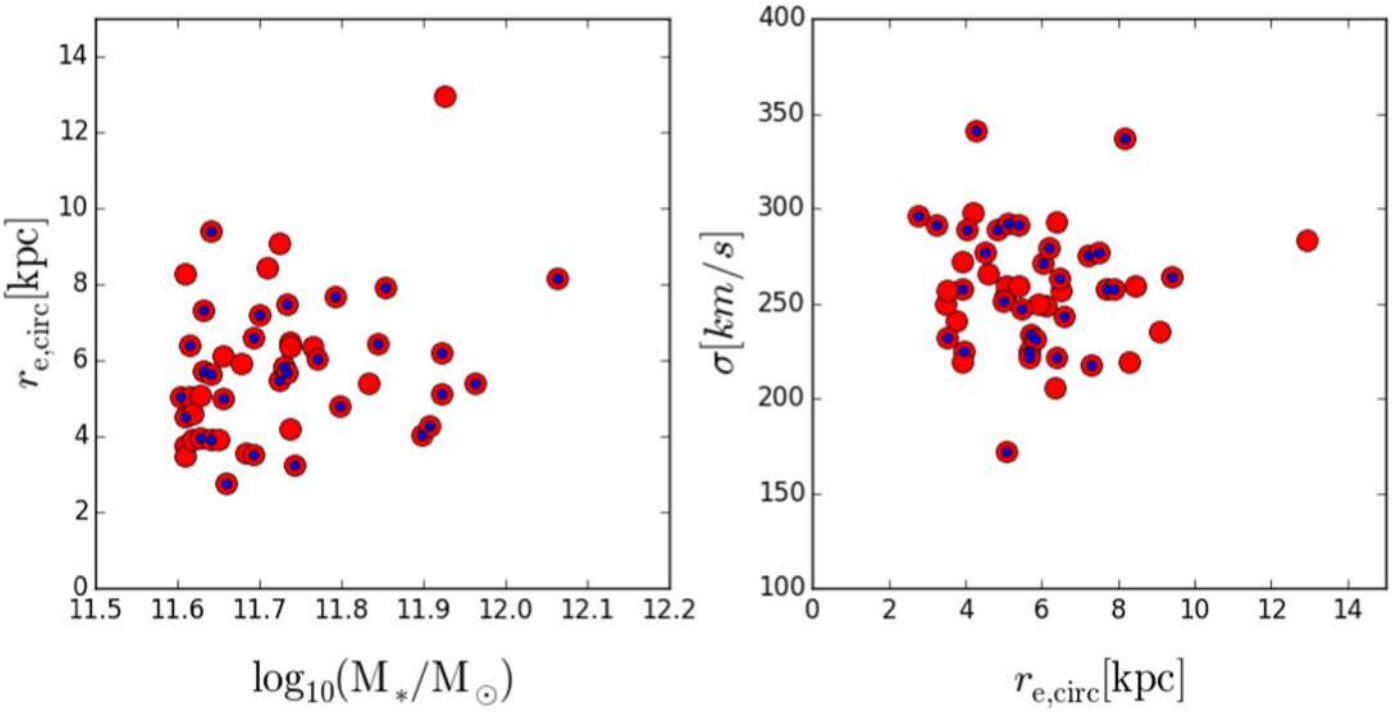

Figure B1. Distribution of MASSIVE(n) (red) and MASSIVE(n_{17}) (blue) for the stellar mass–size and size–dispersion plane. The MASSIVE(n_{17}) sample, with available SDSS photometry, is selected uniformly from the parent MASSIVE(n) sample and can be considered representative for the cumulative number density-matched parent sample.

Appendix C
Evolution of the FP and *M/L* Relations

We present Figure C1, showing Figures 1 and 2 from Section 3, here with the mean evolutionary trends covered in Section 4.

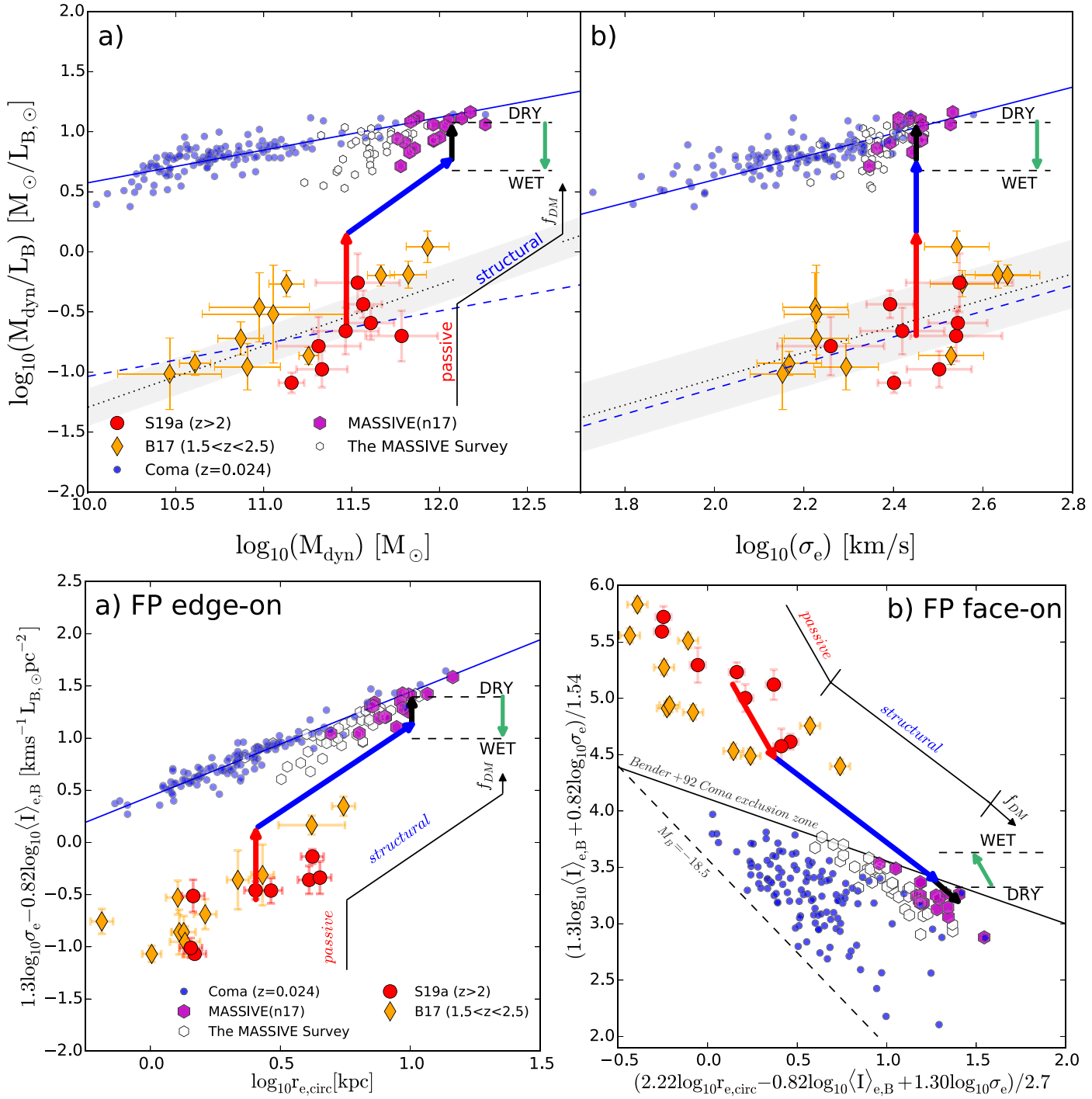

Figure C1. FP and M/L scaling relations from Section 3 shown with the predicted passive, structural, dark matter fraction evolution (red, blue, and black arrows) and luminosity change due to minor mergers (green arrow).

Appendix D
Details on the Modeling of the *B*-band Luminosity Increase Due to Minor Merger-added Stellar Populations

The amount of *B*-band luminosity increase due to the minor wet and dry newly added merging stellar populations is constrained between redshifts $z = 2$ to 0, based on simple assumptions. Figure D1 shows the *B*-band luminosity increase due to minor mergers (on top of the luminosity decrease due to passive evolution) as a function of redshift for three scenarios. In scenario A, it is assumed that all the merging happens about 300 Myr after the galaxies are observed at $z = 1.8$. In scenario B, the galaxies merge at z = 0.1. Note that, since the merging galaxies follow the global star-forming main sequence, and hence have lower star formation rates at lower redshifts on average, the increase in luminosity is less at $z = 0.1$ than at $z = 1.8$. Finally, scenario C shows a more realistic merger history for which 10%, 20%, 30%, and 40% of the stellar mass increase happens via mergers at $z = 0.1$, 0.5, 1.5, and 1.8, respectively. These follow roughly the measured trends of merger fraction in the literature (e.g., Man et al. 2012, 2016; Newman et al. 2012). Although the merger history in the different scenarios is very different, the final increase in rest-frame *B*-band luminosity is very similar between 0.4 and 0.45 dex.

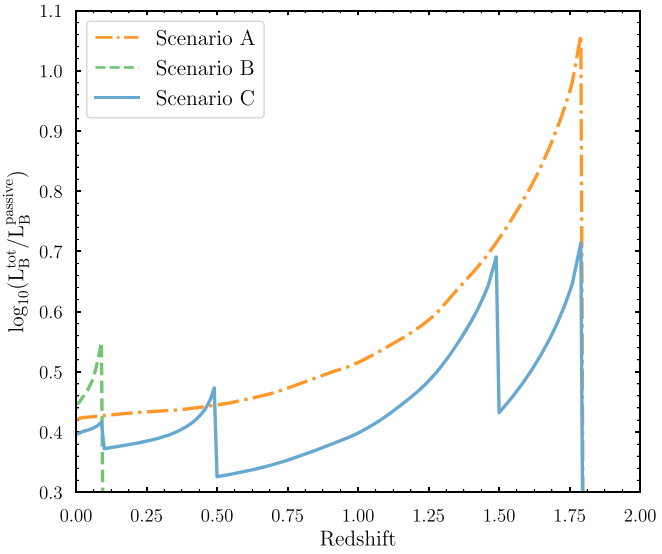

Figure D1. Change in *B*-band luminosity of different galaxies during passive evolution (after turn-off of star formation) after evolving along the star-forming main sequence. The luminosity is compared to a galaxy on the main sequence until $z = 0$. The minor mergers added are matched to the predicted stellar mass increase. Model C, the most realistic one, adds relative fractions 0.1/0.2/0.3/0.4 of minor mergers at a redshift of $z = 0.1/0.5/1.5/1.8$, whereas models A and B assume the extremes of adding all the stellar populations at either 1.8 and $z = 0.1$, respectively. The luminosity increase is scaling directly with the added stellar mass and is independent of the mass ratio. The final relative luminosity increase due to minor merger-driven structural evolution is not strongly affected by adding stellar mass at different epochs, and gives values of $\log_{10}(L_B^{tot}/L_B^{passive}) \sim 0.4$.

Appendix E
Formation Redshift from the *M/L* Relation

Jørgensen & Chiboucas (2013), using Maraston (2005) models, predict the M/L_B ratio evolution as a function of age and metallicity (Jørgensen & Chiboucas 2013, Table 9):

$$\log M/L_B = 0.935 \log \text{age} + 0.337[\text{M/H}] - 0.053. \quad \text{(E1)}$$

For passive evolution with constant metallicity, the difference in $\log_{10} M/L_B$ can be related to the age of the stellar population:

$$\Delta \log M/L_B = 0.935\ \Delta \log \text{age} \quad \text{(E2)}$$

If the MQGs at $z > 2$ are the progenitors of the local Coma relation, the change in M/L_B (at fixed dynamical mass) can be used to estimate a corresponding formation time. The age difference can be written in terms of look-back times and expressed as the formation time:

$$\begin{aligned}\Delta \log \text{age} &= \log \text{age}_{z=0} - \log \text{age}_{z\sim 2} \\ &= \log(t_{\text{form}} - t_{\text{obs},z=0}) - \log(t_{\text{form}} - t_{\text{obs},z\sim 2}) \\ &= \log\left(\frac{1 - t_{\text{form}}/t_{\text{obs},z=0}}{1 - t_{\text{form}}/t_{\text{obs},z\sim 2}}\right)\end{aligned} \quad \text{(E3)}$$

$$\begin{aligned}\Leftrightarrow t_{\text{form}} &= \frac{10^{\Delta \log \text{age}} \cdot t_{\text{obs},z\sim 2} - t_{\text{obs},z=0}}{10^{\Delta \log \text{age}} - 1} \\ &= \frac{10^{(\Delta \log M/L_B)/0.935} \cdot t_{\text{obs},z\sim 2} - t_{\text{obs},z=0}}{10^{(\Delta \log M/L_B)/0.935} - 1}\end{aligned} \quad \text{(E4)}$$

The uncertainty in the formation redshift is estimated by varying the M/L uncertainties (∼0.25 dex).

ORCID iDs

Mikkel Stockmann https://orcid.org/0000-0001-5983-6273
Inger Jørgensen https://orcid.org/0000-0003-3002-1446
Sune Toft https://orcid.org/0000-0003-3631-7176
Christopher J. Conselice https://orcid.org/0000-0003-1949-7638
Andreas Faisst https://orcid.org/0000-0002-9382-9832
Berta Margalef-Bentabol https://orcid.org/0000-0001-8702-7019
Anna Gallazzi https://orcid.org/0000-0002-9656-1800
Stefano Zibetti https://orcid.org/0000-0003-1734-8356
Gabriel B. Brammer https://orcid.org/0000-0003-2680-005X
Carlos Gómez-Guijarro https://orcid.org/0000-0002-4085-9165
Michaela Hirschmann https://orcid.org/0000-0002-3301-3321
Claudia D. Lagos https://orcid.org/0000-0003-3021-8564
Francesco M. Valentino https://orcid.org/0000-0001-6477-4011
Johannes Zabl https://orcid.org/0000-0002-9842-6354

References

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Beifiori, A., Mendel, J. T., Chan, J. C. C., et al. 2017, ApJ, 846, 120
Belli, S., Newman, A. B., & Ellis, R. S. 2017, ApJ, 834, 18
Bender, R., Burstein, D., & Faber, S. M. 1992, ApJ, 399, 462
Bernardi, M., Fischer, J. L., Sheth, R. K., et al. 2017, MNRAS, 468, 2569
Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290
Blakeslee, J. P., Franx, M., Postman, M., et al. 2003, ApJL, 596, L143
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2005, MNRAS, 362, 184
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buitrago, F., Trujillo, I., Conselice, C. J., et al. 2008, ApJL, 687, L61
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Chabrier, G. 2003, PASP, 115, 763
Choi, E., Somerville, R. S., Ostriker, J. P., Naab, T., & Hirschmann, M. 2018, ApJ, 866, 91
Cimatti, A. 2008, Natur, 451, 253
Conselice, C. J., Bluck, A. F. L., Buitrago, F., et al. 2011, MNRAS, 413, 80
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJL, 600, L127
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
D'Odorico, S., Dekker, H., Mazzoleni, R., et al. 2006, Proc. SPIE, 6269, 626933
D'Onofrio, M., Fasano, G., Moretti, A., et al. 2013, MNRAS, 435, 45
Dressler, A., Lynden-Bell, D., Burstein, D., et al. 1987, ApJ, 313, 42
Dubois, Y., Gavazzi, R., Peirani, S., & Silk, J. 2013, MNRAS, 433, 3297
Faber, S. M., Dressler, A., Davies, R. L., et al. 1987, in Nearly Normal Galaxies. From the Planck Time to the Present, ed. S. M. Faber (New York: Springer), 175
Faisst, A. L., Carollo, C. M., Capak, P. L., et al. 2017, ApJ, 839, 71
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
Gómez-Guijarro, C., Toft, S., Karim, A., et al. 2018, ApJ, 856, 121
Greene, J. E., Janish, R., Ma, C.-P., et al. 2015, ApJ, 807, 11
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Habouzit, M., Genel, S., Somerville, R. S., et al. 2019, MNRAS, 484, 4413
Hilz, M., Naab, T., Ostriker, J. P., et al. 2012, MNRAS, 425, 3119
Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924
Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, MNRAS, 398, 898
Jørgensen, I. 1999, MNRAS, 306, 607
Jørgensen, I., & Chiboucas, K. 2013, AJ, 145, 77
Jørgensen, I., Chiboucas, K., Berkson, E., et al. 2017, AJ, 154, 251
Jørgensen, I., Chiboucas, K., Flint, K., et al. 2006, ApJL, 639, L9
Jørgensen, I., Chiboucas, K., Webb, K., & Woodrum, C. 2018, AJ, 156, 224
Jørgensen, I., Franx, M., & Kjærgaard, P. 1996, MNRAS, 280, 167
Jørgensen, I., Hunter, L. C., O'Neill, C. R., et al. 2019, ApJ, 881, 42
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Lagos, C. d. P., Stevens, A. R. H., Bower, R. G., et al. 2018, MNRAS, 473, 4956
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Ma, C.-P., Greene, J. E., McConnell, N., et al. 2014, ApJ, 795, 158
Man, A. W. S., Toft, S., Zirm, A. W., Wuyts, S., & van der Wel, A. 2012, ApJ, 744, 85
Man, A. W. S., Zirm, A. W., & Toft, S. 2016, ApJ, 830, 89
Maraston, C. 2005, MNRAS, 362, 799
McLean, I. S., Ramsay, S. K., & Takami, H. (ed.) 2012, Ground-based and Airborne Instrumentation for Astronomy IV, Vol. 8446 (Bellingham, WA: SPIE)
McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2020, arXiv:2009.03176
Morishita, T., Abramson, L. E., Treu, T., et al. 2019, ApJ, 877, 141
Mowla, L., van Dokkum, P., Brammer, G., et al. 2019, ApJ, 880, 57
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178
Naab, T., Oser, L., Emsellem, E., et al. 2014, MNRAS, 444, 3357
Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Oogi, T., & Habe, A. 2013, MNRAS, 428, 641
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Remus, R.-S., Dolag, K., Naab, T., et al. 2017, MNRAS, 464, 3742
Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2010, A&A, 524, A6
Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2016, A&A, 596, C1
Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJS, 172, 38
Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, ApJ, 788, 125
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Stockmann, M., Toft, S., Gallazzi, A., et al. 2020, ApJ, 888, 4
Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, ApJ, 749, 121
Tapia, T., Eliche-Moral, M. C., Querejeta, M., et al. 2014, A&A, 565, A31
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Toft, S., Gallazzi, A., Zirm, A., et al. 2012, ApJ, 754, 3
Toft, S., Smolčić, V., Magnelli, B., et al. 2014, ApJ, 782, 68
Treu, T., Ellis, R. S., Liao, T. X., et al. 2005, ApJ, 633, 174
Trujillo, I., Conselice, C. J., Bundy, K., et al. 2007, MNRAS, 382, 109
Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, ApJ, 650, 18
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2014, ApJL, 793, L31
van der Marel, R. P., & van Dokkum, P. G. 2007, ApJ, 668, 756
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 Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
Veale, M., Ma, C.-P., Greene, J. E., et al. 2018, MNRAS, 473, 5446
Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, A&A, 536, A105
Wellons, S., & Torrey, P. 2017, MNRAS, 467, 3887