MOSFIRE Spectroscopy of Quiescent Galaxies at $1.5 < z < 2.5$. II. Star Formation Histories and Galaxy Quenching

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

[1] Max-Planck-Institut für Extraterrestrische Physik (MPE), Giessenbachstr. 1, D-85748 Garching, Germany
[2] The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA
[3] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Abstract

We investigate the stellar populations for a sample of 24 quiescent galaxies at $1.5 < z < 2.5$ using deep rest-frame optical spectra obtained with Keck MOSFIRE. By fitting templates simultaneously to the spectroscopic and photometric data and exploring a variety of star formation histories, we obtain robust measurements of median stellar ages and residual levels of star formation. After subtracting the stellar templates, the stacked spectrum reveals the $H\alpha$ and [N II] emission lines, providing an upper limit on the ongoing star formation rate of $0.9 \pm 0.1\ M_\odot\ \mathrm{yr}^{-1}$. By combining the MOSFIRE data with our sample of Keck LRIS spectra at lower redshift, we analyze the quiescent population at $1 < z < 2.5$ in a consistent manner. We find a tight relation (with a scatter of 0.13 dex) between the stellar age and the rest-frame $U - V$ and $V - J$ colors, which can be used to estimate the age of quiescent galaxies, given their colors. Applying this age–color relation to large photometric samples, we are able to model the number density evolution for quiescent galaxies of various ages. We find evidence for two distinct quenching paths: a fast quenching that produces compact post-starburst systems and a slow quenching of larger galaxies. Fast quenching accounts for about a fifth of the growth of the red sequence at $z \sim 1.4$ and half at $z \sim 2.2$. We conclude that fast quenching is triggered by dramatic events, such as gas-rich mergers, while slow quenching is likely caused by a different physical mechanism.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: stellar content

1. Introduction

The physical mechanism responsible for quenching the star formation activity in massive galaxies remains one of the most important missing pieces in the puzzle of galaxy formation. The existence of a population of massive quiescent galaxies at $z \sim 2$ (Franx et al. 2003; Cimatti et al. 2004), when the universe was only a few Gyr old, requires remarkably rapid and efficient quenching. These galaxies are also found to be significantly more compact than their local counterparts (e.g., van Dokkum et al. 2008; Newman et al. 2012), thus raising additional questions about the process that led to their formation.

Moreover, it is not clear what these systems looked like before being quenched, and constraining the properties of their star-forming progenitors is currently a major observational goal. The physical mechanism responsible for quenching and the properties of the progenitors are tightly connected and represent two aspects of the same open question. A number of possibilities have been put forward, including evolutionary links with submillimeter galaxies (Tacconi et al. 2008; Toft et al. 2014) and compact star-forming galaxies (Barro et al. 2013; van Dokkum et al. 2015); however, observations are currently unable to discriminate among different scenarios.

A powerful method to study the past evolution of massive quiescent galaxies is the reconstruction of their star formation history (SFH), which can be done by fitting models to the observed spectral energy distribution (SED) or, when available, spectroscopy (see, e.g., Renzini 2006; Conroy 2013). This archeological approach has allowed detailed studies of the local population, revealing an antihierarchical assembly in which the most massive objects formed earlier and on shorter timescales (e.g., Thomas et al. 2005). Additionally, the stellar metallicity

can be measured from high-quality spectra: the state-of-the-art analysis of local quiescent galaxies includes measured abundances for 16 different chemical elements (Conroy et al. 2014).

The archeological approach, however, presents significant limitations when used to understand the early evolution of quiescent galaxies. First, the spectra of old stellar populations present very little evolution, a fact that produces the remarkable homogeneity of the local red sequence (Bower et al. 1992). As a result, ages become increasingly uncertain for older systems. Second, there is a degeneracy between the mass *formation* history and the mass *assembly* history. Even when a reliable SFH is found, it is not possible to know whether the various episodes were due to in situ star formation or took place in separate systems that eventually merged together. Being able to directly probe these different scenarios is fundamental for our understanding of galaxy formation.

The limitations of galaxy archeology can be overcome by observing galaxies at high redshift, where stellar ages are bounded by the age of the universe at the epoch of observation. Rest-frame optical spectra are particularly powerful when probing stellar populations younger than about 3 Gyr, due to the presence of Balmer absorption lines that are strong and evolve rapidly with age. In some sense, measuring ages and SFHs is increasingly easier at higher redshift; the difficulty lies in obtaining deep rest-frame optical spectra for faint, distant targets.

Recent studies of large, high-quality spectroscopic samples at intermediate redshift ($z < 1$) have shown that the metallicities and stellar ages of quiescent galaxies are consistent with passive evolution over the last 7 Gyr of cosmic history, and that more massive galaxies tend to have older stellar populations (Choi et al. 2014; Gallazzi et al. 2014; Chauke et al. 2018;

Wu et al. 2018b), in agreement with the results of kinematic studies (Treu et al. 2005; van der Wel et al. 2005). Extending this type of analysis to earlier epochs, when the majority of massive galaxies formed, is observationally challenging. At $1 < z < 1.5$, the use of red-sensitive CCDs enables the collection of relatively large samples (Bezanson et al. 2013; Belli et al. 2015), but at $z > 1.5$, near-infrared observations are needed, and most studies are based on stacked spectra (Whitaker et al. 2013; Newman et al. 2014; Mendel et al. 2015; Onodera et al. 2015; Fumagalli et al. 2016) or sample sizes of one to a few absorption-line spectra (van de Sande et al. 2013; Kriek et al. 2016; Glazebrook et al. 2017; Schreiber et al. 2018a). These observations revealed that the typical ages of quiescent galaxies at $z \sim 2$ are 1–2 Gyr, with a relatively large age spread at a given redshift, confirming that this is the epoch when the red sequence first formed.

In this work, we analyze the SFHs for a sample of 24 quiescent galaxies at $1.5 < z < 2.5$ for which we obtained deep near-infrared observations with the MOSFIRE instrument at Keck. We have already presented this sample in Belli et al. (2017b, hereafter Paper I), where we used the absorption-line spectra to derive the stellar kinematics. In the present work, we constrain the SFHs of these galaxies by fitting templates simultaneously to the spectroscopic and photometric data. This technique enables us to take advantage of both the wide wavelength range of photometric data and the age-sensitive features in the rest-frame optical spectra and was already successfully applied to our Keck/LRIS sample at $1 < z < 1.5$ in Belli et al. (2015).

In addition to the limited availability of deep rest-frame optical spectra, the measurement of stellar ages at high redshift suffers from systematics due to the assumptions involved in the spectral fitting. A functional form for the SFH is often assumed, but the impact of this procedure on the age measurement has not been explored in detail. In this study, we attempt a comparison of different SFH models, with the goal of understanding which physical properties are robust against different SFH assumptions.

We briefly present the spectroscopic data in Section 2. In Section 3, we discuss the spectral fits, perform a detailed comparison between different SFH models, and measure the amount of recent star formation using the Hα emission line. We then use the results of the spectral fits to calibrate the relation between age and rest-frame color in Section 4. In Section 5, we apply this relation to the large photometric catalog from the UltraVISTA survey, and we model in detail the number density evolution of the quiescent population. The main result is the identification of two distinct quenching channels, which we discuss in detail in Section 6. Finally, we summarize our results in Section 7.

Throughout the paper, we use AB magnitudes and assume a ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$.

2. Data

The present work is based on a sample of 24 deep spectra obtained at Keck with the near-infrared spectrograph MOSFIRE. In this section, we briefly present the sample; for a detailed description of the target selection, observations, and data reduction, we refer to Paper I.

We selected massive galaxies with photometric redshifts in the range $1.5 < z < 2.5$ from the 3D-*HST* catalog (Momcheva et al. 2016), prioritizing systems that are quiescent according to

their rest-frame *UVJ* colors. Since all of the targets are in the CANDELS fields (Grogin et al. 2011; Koekemoer et al. 2011), a wealth of ancillary photometric and imaging data are publicly available. The MOSFIRE observations were carried out in the *Y* and *J* bands with exposure times between 4 and 9 hr per pointing. Additional, shorter observations (1–2.5 hr) were obtained in the *H* and *K* bands in order to target the Hα emission line. A standard data reduction, which includes telluric correction and absolute flux calibration, was performed.

To explore the evolution of quiescent galaxies over a wide redshift range, in Section 4, we will augment the sample with the galaxies at $1 < z < 1.5$ observed with Keck/LRIS and presented in Newman et al. (2010) and Belli et al. (2014). This sample is well matched to the MOSFIRE one in terms of target selection and rest-frame wavelength range. As we have shown in our previous studies, our two Keck samples probe the population of massive quiescent galaxies at high redshift in an unbiased way, except for a slight lack of old and dusty objects at $z > 1.5$ (Paper I).

The combined Keck sample represents the largest homogeneous set of deep absorption-line spectroscopy of quiescent galaxies at $z > 1$. We publicly release the reduced spectra for the entire sample.[4]

3. Stellar Populations

3.1. Spectral Fitting

We analyze the data using `pyspecfit` (Newman et al. 2014), a Bayesian code that fits stellar population templates to the observed spectroscopy and photometry of a galaxy. The code employs the nested sampling algorithm (Skilling 2004) to explore the multidimensional parameter space in an efficient way.

Our spectroscopic data typically cover the rest-frame wavelength range 3700–4200 Å, but the exact interval depends on redshift and target position on the MOSFIRE slit mask. Before the fit, in each spectrum, we mask out the pixels at the expected position of the [O II] emission line, since our templates are purely stellar and do not include the contribution of ionized gas. To reduce the impact of sky lines on the spectrum, we also exclude from the fit the spectral pixels with the strongest sky emission, as measured by the observed sky spectrum.

We take the photometric data from the 3D-*HST* catalog, which contains publicly available measurements from the UV to the near-infrared. To avoid potential contamination from dust emission, we mask out the IRAC 8 μm channel. To account for systematic effects, we add in quadrature a 5% contribution to the uncertainty of each photometric point. Data points corresponding to nondetections are included in the fit with their formal uncertainty.

One limitation of optical and near-infrared spectra is the large uncertainty on the absolute flux calibration, which can be affected by slit loss and variable weather conditions. Instrumental effects can also change the overall shape of the spectral response, particularly near the limits of the observed wavelength range. Since the absolute calibration of photometric measurements is typically much more reliable, during the fit we change the overall shape of the observed spectrum, using a fourth-degree polynomial, so that it agrees with the shape given by the broadband photometry. This procedure does not affect

---

any of the spectral features of interest, which extend over narrow spectral ranges. We also scale the error spectra to ensure that the spectroscopic and photometric data are given appropriate weighting in the fit. To do this, for each galaxy, we run a first fit with the sole purpose of calculating the χ^2, which we then use to rescale the error spectrum. The scaling factors are in the range 1.7–2.3, suggesting that the uncertainty on the spectra that is produced by the data reduction is indeed too small.

We generate the synthetic spectra using the Bruzual & Charlot (2003) library and adopting a Chabrier (2003) initial mass function (IMF) and the dust extinction law from Calzetti et al. (2000). A simple stellar population (SSP) of a given age can be used to generate a realistic spectrum once the following four parameters are specified: redshift, stellar velocity dispersion, dust attenuation A_V, and stellar metallicity Z. We leave these four parameters free during the fit, setting their priors in the following way. The spectroscopic redshift can be determined by eye with reasonable accuracy, and we use a narrow Gaussian prior centered on the values reported in Paper I. We also adopt a Gaussian prior for the stellar metallicity, centered on the solar value ($Z = 0.02$) but wide enough to encompass all reasonable values ($\sigma_Z = 0.005$). For the other parameters, we use uninformative priors in the form of uniform distributions: between 50 and 500 km s^{-1} for the velocity dispersion and between 0 and 4 for A_V.

An additional free parameter is the age of the stellar population, which means that a model can be fully specified using five parameters. However, the assumption of SSP is likely too simplistic for real galaxies, and more complex SFHs, with a higher number of free parameters, are often considered. We attempt here a systematic exploration of different SFH models, with a twofold goal: first, we want to derive the stellar population properties in a robust way, ensuring that the results are not affected by the assumed form of the SFH; second, we are going to assess whether some of the SFH models are better than others in describing the spectra of real galaxies.

Each SFH model is simply a function of look-back time that is normalized to unity and depends on one or more parameters. The SFH models used in our analysis, together with the priors for each parameter, are listed in Table 1. We use the following six SFH models.

1. The SSP represents the basic form of SFH, in which all stars form in an instantaneous burst. The only free parameter is the age of the population, for which we adopt a log-uniform prior that is truncated at the age of the universe at the observed redshift.
2. Next, we consider a top-hat model, i.e., a burst with a finite duration. This model has two free parameters: the age and the duration of the burst.
3. The most popular SFH choice is an exponentially declining function, also called a *tau model*. This SFH has two free parameters: the age and the timescale τ, and the star formation rate is proportional to $\exp(-t/\tau)$. In this SFH, the current star formation rate is lower than at any time in the past, which is a reasonable assumption for galaxies that are, by selection, quiescent.
4. Another popular choice is the delayed tau model, which has the form $t \exp(-t/\tau)$. This SFH has been proposed in order to overcome the unrealistically sharp peak at early

Table 1
Priors on the SFH Model Parameters

SFH Model	Parameters	Prior
SSP	Age	LogUnif(100 Myr, t_U)
Top-hat	Age	LogUnif(100 Myr, t_U)
	Duration	LogUnif(100 Myr, t_U)
Tau	Age	LogUnif(100 Myr, t_U)
	τ	LogUnif(10 Myr, 10 Gyr)
Delayed tau	Age	LogUnif(100 Myr, t_U)
	τ	LogUnif(10 Myr, 10 Gyr)
Const + tau	Age	LogUnif(100 Myr, t_U)
	Duration	LogUnif(100 Myr, t_U)
	τ	LogUnif(10 Myr, 10 Gyr)
Piecewise[a]	A_0	LogNormal(0, 2)
	$A_{0.25}$	LogNormal(0, 2)
	$A_{0.45}$	LogNormal(0, 2)
	$A_{0.8}$	LogNormal(0, 2)
	$A_{1.4}$	LogNormal(0, 2)
	$A_{2.5}$	LogNormal(0, 2)

Note. Here t_U is the age of the universe at the spectroscopic redshift of the galaxy. LogUnif(a, b) is the log-uniform distribution between a and b. LogNormal(μ, σ) is the lognormal distribution, such that the logarithm of the variable has a mean μ and standard deviation σ.
[a] Here A_X is the amplitude of the SFH at a look-back time of X Gyr; A_4 is fixed to 1.

times featured by the tau model without introducing additional free parameters.

5. In order to allow more flexibility, we introduce a new model that consists of a constant SFH followed by an exponential decay, hereafter called the *constant + tau model*. This model has three free parameters—age, duration of the flat phase, and timescale τ for the exponential phase—and can be considered a generalization of both the top-hat and the tau models. Its main advantage is that it allows the quenching phase to be completely detached from the initial star-forming phase.
6. Finally, we consider a nonparametric SFH with a large degree of flexibility in order to explore further possibilities, such as secondary bursts or rejuvenation episodes. We adopt a piecewise description of the SFH with seven nodes at fixed, logarithmically spaced look-back times (0, 0.25, 0.45, 0.8, 1.4, 2.5, and 4 Gyr). Since the SFH models must be normalized, this corresponds to six free parameters. We choose to leave the amplitude of the first six nodes free, while the amplitude of the oldest node is arbitrarily fixed to unity. To avoid biasing the results, we keep the oldest node fixed at 4 Gyr for all galaxies, although this means that, according to the redshift, stars slightly older than the age of the universe may be allowed. The SFH is then constructed by linearly interpolating between the nodes. We adopt a lognormal prior for each amplitude, centered on 1 and with a width of 2 dex, thus allowing a large amount of freedom.

The SFH represents the relative contribution of stars of different ages to a population. The total luminosity is then determined by the total number of stars in the population, i.e., by the total stellar mass. For each template, this can be calculated by multiplying its mass-to-light ratio by the observed luminosity, corrected for cosmological dimming. The stellar mass is therefore an output of the fitting procedure
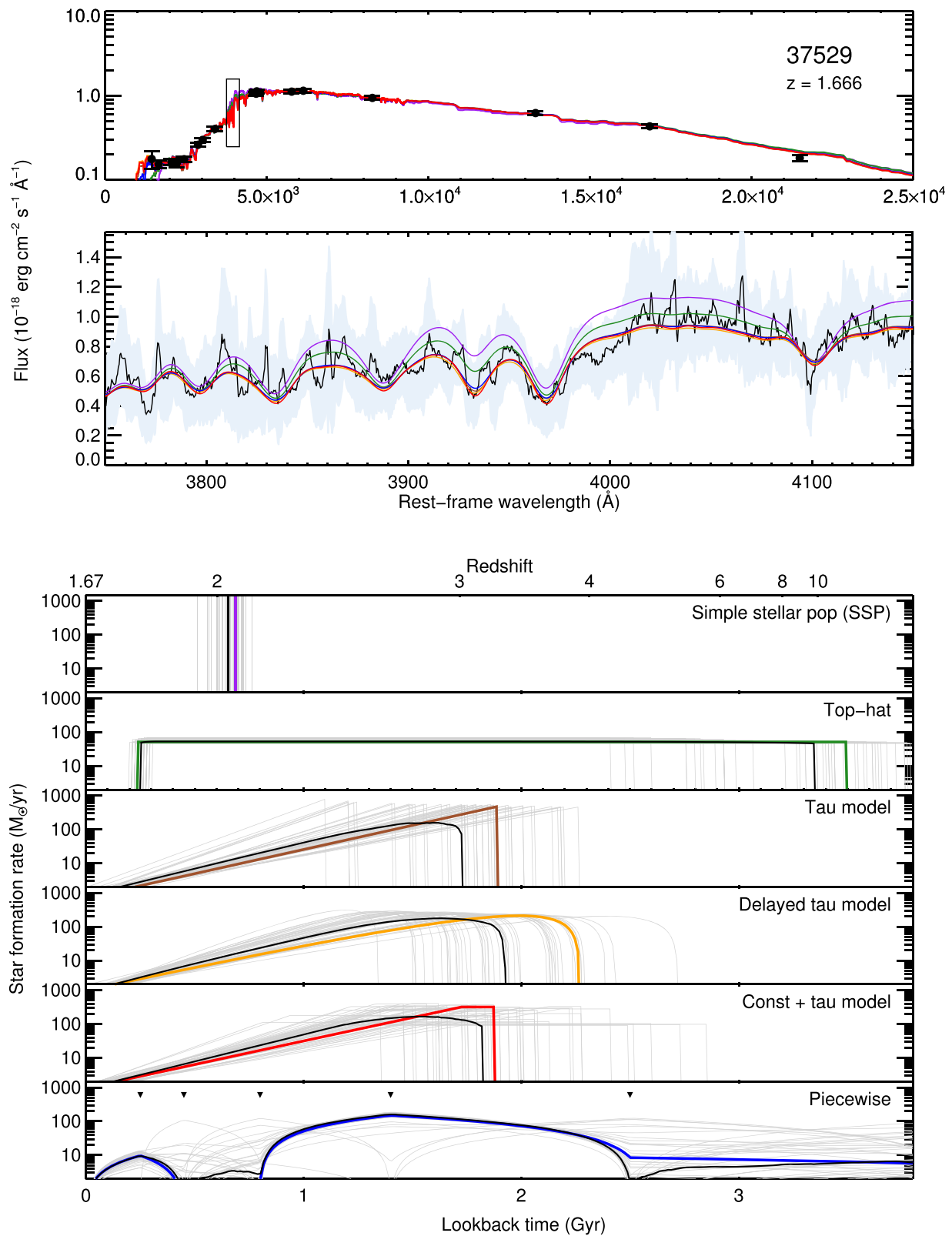
Figure 1. Example of spectral fitting for object 37529 at $z = 1.666$. The observed photometry and spectroscopy are fit using six different models for the SFH. Top: observed broadband photometry (black points) and best-fit template for each of the six models (colored lines). The black box indicates the part of the spectrum observed with MOSFIRE. Middle: observed MOSFIRE spectrum (black line) with its uncertainty (shaded region). The same six best-fit templates are also shown (colored lines). Bottom: In each panel, the best-fit SFH is shown as a colored line, where the colors correspond to the best-fit spectra shown in the top two panels; a random subset of SFHs drawn from the posterior distribution is shown in gray; and the median SFH is shown in black. The look-back time for this particular galaxy is converted into the corresponding redshift in the top x axis. In the last panel, the nodes of the piecewise SFH are marked by black triangles.
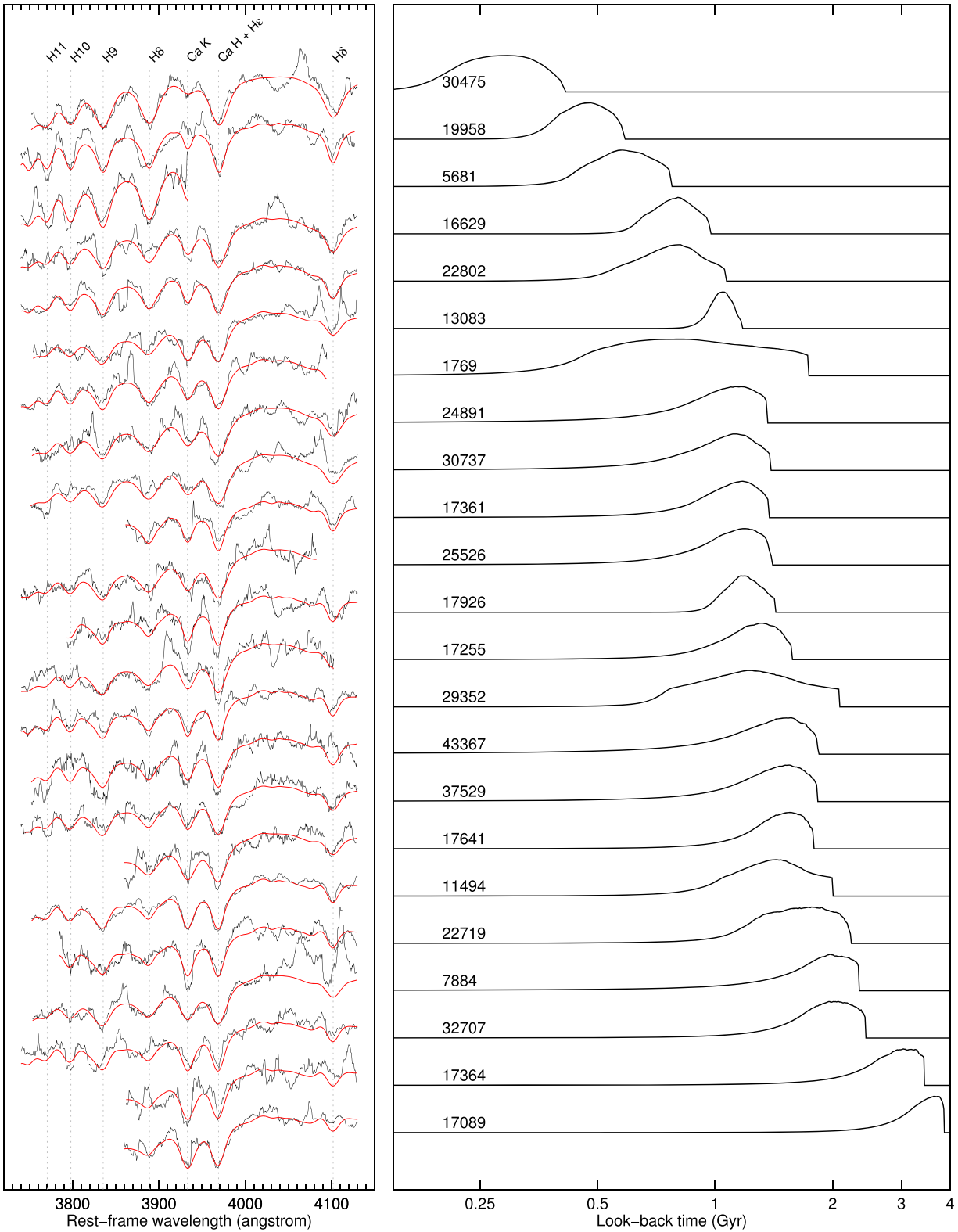
Figure 2. Rest-frame optical spectra and derived SFHs for the galaxies in the MOSFIRE sample, sorted by increasing median age. Left: For each galaxy, the MOSFIRE spectrum (black) and the best-fit template (red) are shown; the most important absorption lines are indicated by vertical dashed lines and labeled at the top. Right: median SFH assuming a constant + tau model is shown for each object.

but does not represent a free parameter of the model. The same is true for the instantaneous star formation rate, which can be easily derived from the SFH after accounting for the mass loss due to winds and supernovae. We make the simplifying

assumption of instantaneous recycling, so that a fixed return fraction R of the newly formed stars is immediately lost and does not contribute to the stellar mass. For a Chabrier IMF, the return fraction is $R = 0.41$ (Madau & Dickinson [2014]). The

Table 2
Stellar Population Properties of the MOSFIRE Sample

ID (3D-*HST*)	Field	z	t_{50}[a] (10^9 yr)	A_V[a] (mag)	Z[a]	$\log M_*/M_\odot$[a]	$\log$ sSFR[a,b] (yr^{-1})	SFR(Hα)[c] ($M_\odot$ yr^{-1})
30475[d]	UDS	1.63	0.30 ± 0.06	0.85 ± 0.10	0.013 ± 0.004	10.71 ± 0.02	−11.0 ± 1.0	<8.9
19958	COSMOS	1.72	0.48 ± 0.04	0.73 ± 0.06	0.012 ± 0.002	10.64 ± 0.01	−14.2 ± 3.0	⋯
5681	COSMOS	2.43	0.61 ± 0.07	0.64 ± 0.09	0.017 ± 0.003	10.91 ± 0.02	−12.9 ± 3.6	<14.9
16629	COSMOS	1.66	0.81 ± 0.07	0.44 ± 0.05	0.017 ± 0.003	10.59 ± 0.01	−17.0 ± 5.7	<7.2
22802	UDS	1.67	0.83 ± 0.1	0.62 ± 0.09	0.015 ± 0.004	11.01 ± 0.02	−13.2 ± 4.9	<6.8
13083	COSMOS	2.09	1.1 ± 0.1	0.20 ± 0.10	0.022 ± 0.003	10.95 ± 0.03	−20.0 ± 9.8	<14.5
1769	COSMOS	2.30	1.1 ± 0.2	0.43 ± 0.10	0.022 ± 0.005	11.16 ± 0.03	−11.1 ± 0.2	⋯
24891	UDS	1.60	1.1 ± 0.1	0.55 ± 0.10	0.015 ± 0.004	10.87 ± 0.02	−10.4 ± 0.09	<6.3
30737[d]	UDS	1.62	1.1 ± 0.1	0.51 ± 0.09	0.019 ± 0.004	11.23 ± 0.02	−10.6 ± 0.08	<15.1
17361	COSMOS	1.53	1.1 ± 0.09	0.37 ± 0.07	0.011 ± 0.002	10.74 ± 0.02	−10.9 ± 0.1	<1.7
25526	EGS	1.75	1.2 ± 0.2	0.34 ± 0.10	0.015 ± 0.005	10.69 ± 0.03	−10.9 ± 1.3	⋯
17926	EGS	1.57	1.2 ± 0.1	0.33 ± 0.09	0.012 ± 0.002	10.85 ± 0.05	−20.1 ± 10.6	<2.0
17255	COSMOS	1.74	1.3 ± 0.2	0.31 ± 0.06	0.017 ± 0.004	10.81 ± 0.03	−11.3 ± 0.4	⋯
29352	UDS	1.69	1.5 ± 0.3	0.23 ± 0.07	0.012 ± 0.003	10.82 ± 0.05	−12.6 ± 5.5	⋯
43367	UDS	1.62	1.5 ± 0.3	0.68 ± 0.10	0.017 ± 0.004	11.06 ± 0.04	−10.4 ± 0.1	<9.2
37529	UDS	1.67	1.5 ± 0.2	0.58 ± 0.07	0.014 ± 0.004	10.95 ± 0.04	−10.8 ± 0.08	<7.9
17641	COSMOS	1.53	1.5 ± 0.2	0.28 ± 0.05	0.013 ± 0.003	10.61 ± 0.04	−11.5 ± 1.8	<3.0
11494	COSMOS	2.09	1.6 ± 0.2	0.26 ± 0.06	0.022 ± 0.003	11.49 ± 0.03	−11.8 ± 0.09	<9.9
22719	EGS	1.58	1.8 ± 0.5	0.40 ± 0.09	0.012 ± 0.003	10.96 ± 0.08	−11.7 ± 7.4	<1.0
7884	COSMOS	2.11	2.0 ± 0.3	0.38 ± 0.06	0.013 ± 0.004	11.35 ± 0.02	−10.7 ± 0.08	<9.6
32707	UDS	1.65	2.1 ± 0.3	0.31 ± 0.08	0.014 ± 0.004	11.20 ± 0.04	−11.5 ± 0.10	<13.4
17364	COSMOS	1.53	3.0 ± 0.4	0.13 ± 0.07	0.015 ± 0.003	10.94 ± 0.02	−11.5 ± 0.07	<1.7
17089	COSMOS	1.53	3.4 ± 0.2	0.30 ± 0.06	0.013 ± 0.002	11.45 ± 0.02	−12.0 ± 0.07	<2.9

Note. More properties, including coordinates and morphological and dynamical measurements, are available in Table 2 of Paper I.
[a] Derived from spectral fitting adopting a constant + tau model. The value and uncertainty are given by, respectively, the median and standard deviation of the posterior distribution.
[b] Averaged over the past 100 Myr.
[c] 3σ upper limit on star formation rate, measured from the dust-corrected Hα flux.
[d] Detected in the X-ray.

relation between the SFH and star formation rate at any cosmic time t is therefore

$$\mathrm{SFH}(t) \cdot M_* = \frac{dM_*(t)}{dt} = (1 - R) \cdot \mathrm{SFR}(t). \qquad (1)$$

3.2. Results of the Spectral Fitting

Figure 1 illustrates the spectral fitting for one representative galaxy drawn from our sample. The top two panels show, in black, the observed photometry and the MOSFIRE spectrum. Six colored lines, corresponding to the best-fit templates for the six types of SFH, are overplotted. For each SFH model, the colored line in the top panel is exactly the same as the one in the second panel, because we simultaneously fit the spectroscopic and photometric data. In the bottom panel of the figure, the six types of SFHs are plotted in units of star formation rate, using Equation (1) for the conversion. In each panel, the best-fit SFH is shown as a colored line, while thin gray lines illustrate an even sampling of the posterior distribution. We also construct the median SFH by taking the median of the gray lines at each look-back time. This is shown as a thick black line and is more representative of the posterior distribution than the single best-fit curve. Throughout this work, when we quote the results of spectral fitting such as ages, star formation rates, etc., we always use the median calculated from an even sampling of the posterior distribution, not the best-fit value.

We are able to obtain meaningful fits using each of the six SFH models for every galaxy in our sample except for one, ID

35616. Despite the fact that the MOSFIRE spectrum of this galaxy has a remarkably high signal-to-noise ratio, none of the six models of SFH is able to reproduce the photometry; for example, adopting a tau model, we obtain a reduced χ^2 (for the photometry only) of 4.3, which is a clear outlier given that the rest of the sample has a mean χ^2 of 1.1. Since the *HST* imaging shows that this galaxy is clearly interacting with a close companion (see Figure 2 of Paper I), it is likely that some of the photometric measurements are contaminated. Furthermore, this object is also the only one in the sample with strong Hα emission, as we discuss below in Section 3.5. For these two reasons, we exclude 35616 from the remainder of our analysis.

The 23 objects for which we obtain a good spectroscopic and photometric fit are shown in Figure 2. For each galaxy, we plot the MOSFIRE spectrum with the best-fit template, and the median SFH that we obtain when adopting a constant + tau model. The galaxies are sorted by their median age, from the youngest (top) to the oldest (bottom). The sample clearly spans a variety of stellar ages. A few galaxies, for which the fit indicates ages smaller than 1 Gyr, present strong higher-order Balmer lines, indicative of populations dominated by B and A stars. The vast majority of the galaxies in the sample, however, show moderate Balmer absorption lines and a relatively strong 4000 Å break, indicating the presence of both young and old stars. These galaxies can easily be sorted in age by looking at the relative strength of the two Ca II absorption lines (see, e.g., Sánchez Almeida et al. 2012). If a significant population of stars younger than 1 Gyr is present, then the blend of the Ca H and Hϵ lines is deeper than Ca K. A further test of the relative
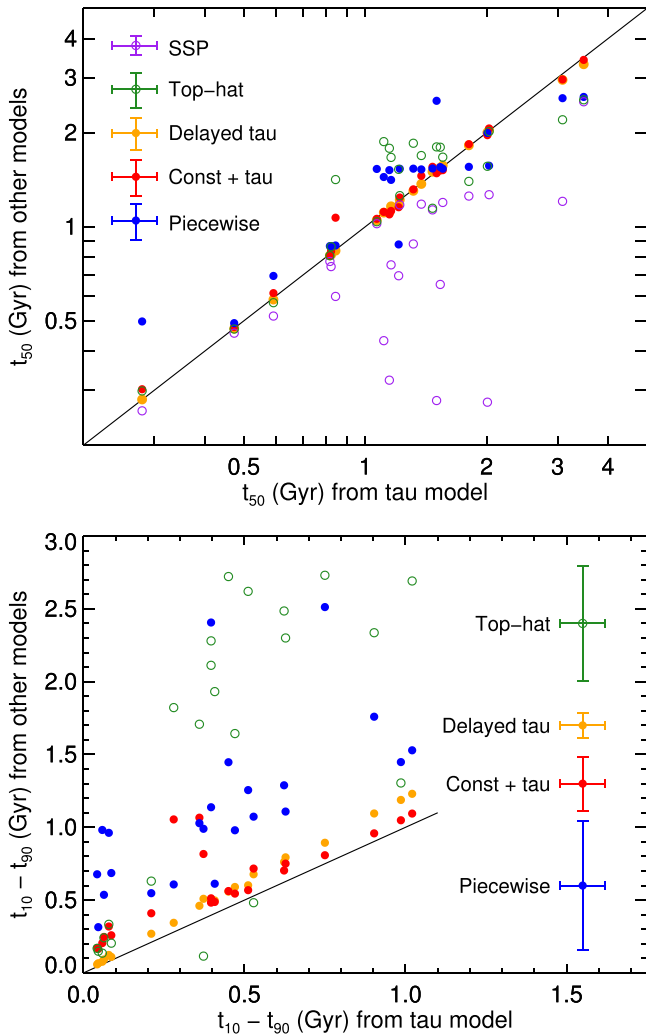
Figure 3. Top: median stellar age t_{50} measured using different SFH models vs. the value obtained adopting the tau model. Bottom: similar comparison for the quantity $t_{10} - t_{90}$, a proxy for the length of the star formation episode. This quantity cannot be measured for SSPs. In both panels, each point corresponds to the median value of the posterior, and the median uncertainty for each SFH model is shown by the error bars in the legend.

ordering of these spectra comes from the Hδ absorption line, which is stronger for younger populations. This qualitative assessment shows that the results of the fits are consistent with the rest-frame optical spectra and not dominated by the photometry.

In addition to the SFH, our spectral fits also determine the stellar mass, dust attenuation, and stellar metallicity, which are listed in Table 2 for the entire sample. The stellar metallicities span the range between half solar and slightly extrasolar, but their posterior distributions are broad, meaning that this quantity cannot be robustly measured by the spectral fits. We choose to include it as a free parameter anyway to avoid biasing the results for the other properties.

3.3. The Effect of the SFH Model on the Results

The example in Figure 1 shows that the six SFH models can produce different results when applied to the same object, in terms of both the best-fit template and the recovered physical properties, such as the stellar age. One of the goals of the present analysis is to compare different SFH models and assess the extent to which the choice of the model affects the results of the spectral fit.

Since each SFH model has a different number and type of parameters, we need to define nonparametric quantities in order to be able to compare the results among one another. Following Pacifici et al. (2016), we define the characteristic time t_X as the look-back time at which $X\%$ of the stellar mass was already formed: $\int_{t_X}^{\infty} \mathrm{SFH}(t_{lb})dt_{lb} = X/100$, where t_{lb} is the look-back time. We start by considering t_{50}, i.e., the median stellar age. We calculate the value of t_{50} using the six SFH models and show the results for the entire sample in the top panel of Figure 3. Since the tau model is the most widely used, we take it as a reference and plot, for each galaxy, the median ages obtained with the other models versus the median age obtained from the tau model. Clearly, the SSP model systematically underestimates the median age. This is a well-known issue, due to the fact that the SSP model effectively returns the stellar population age weighted by light instead of mass. Interestingly, the other five models yield consistent results, showing that the galaxies in our sample span an order of magnitude in stellar ages, from 300 Myr to 3 Gyr. The delayed tau and constant + tau models are in extremely good agreement with the tau model, while the top-hat and piecewise models present a slightly larger scatter. The piecewise model shows an abundance of points around $t_{50} \sim 1.5$ Gyr, which is likely an artificial effect due to the positioning of the nodes (see Figure 1).

Next, we consider the quantity $t_{10} - t_{90}$, which is the time it took to assemble the central 80% of the stellar mass of a galaxy, and can be considered as a proxy for the length of the star formation activity. The bottom panel of Figure 3 shows that different SFH models yield wildly different estimates for $t_{10} - t_{90}$. In particular, the top-hat produces very long bursts, often extending all the way to the limit set by the age of the universe. The tau model, instead, produces very short bursts, which is not surprising given its simple parameterization. The piecewise model yields intermediate results with large uncertainties, which are likely more realistic given its high degree of flexibility. The SSP yields zero by definition and is not shown in the figure. We conclude that the length of the SFH is not as robust as the median age, and its measurement depends strongly on the assumed SFH model.

Another important property of quiescent galaxies is their current level of star formation activity. For each SFH model, we divide the star formation rate (given by Equation (1)) by the stellar mass, thus obtaining the specific star formation rate (sSFR). The SSP and top-hat models, by definition, have zero current star formation rate. The delayed tau and constant + tau models are very similar to the tau model; therefore, we only compare the piecewise and tau models in the top panel of Figure 4. For most of the galaxies, we are only able to obtain upper limits, indicating that the spectra and photometry are not able to constrain the current sSFR. But if we use the sSFR averaged over the previous 100 Myr, shown in the bottom panel, then we are able to obtain measurements with meaningful error bars for most of the galaxies. Moreover, the piecewise and tau models are in excellent agreement, confirming that the spectra are indeed able to detect the level of star formation in the recent but not immediate past. The values of sSFR we obtain range between 10^{-12} and 10^{-10} yr^{-1}, corresponding to mass-doubling times longer than three times the age of the universe at $z \sim 2$, consistent with the quiescent nature of the sample.
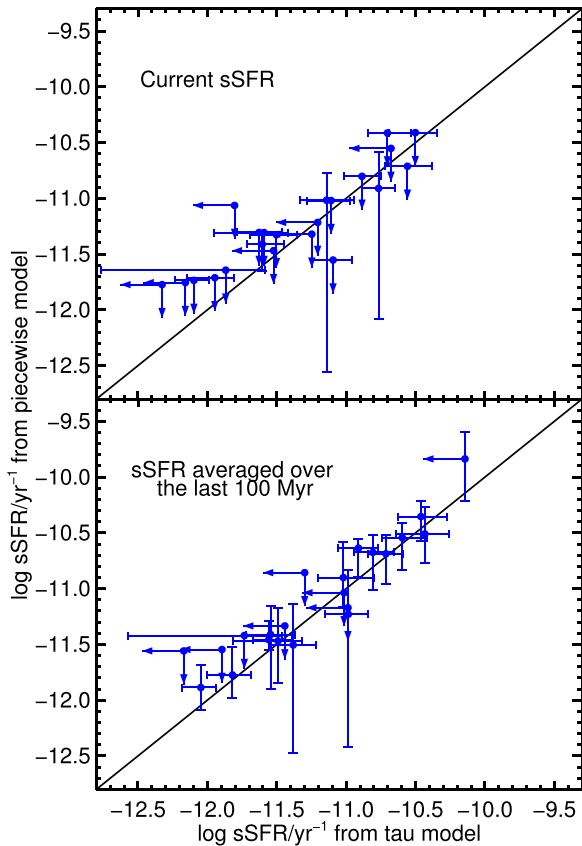
Figure 4. Top: comparison of the current sSFR obtained using the piecewise and tau models. Bottom: same comparison but for the sSFR averaged over the past 100 Myr. Arrows indicate 95% upper limits.
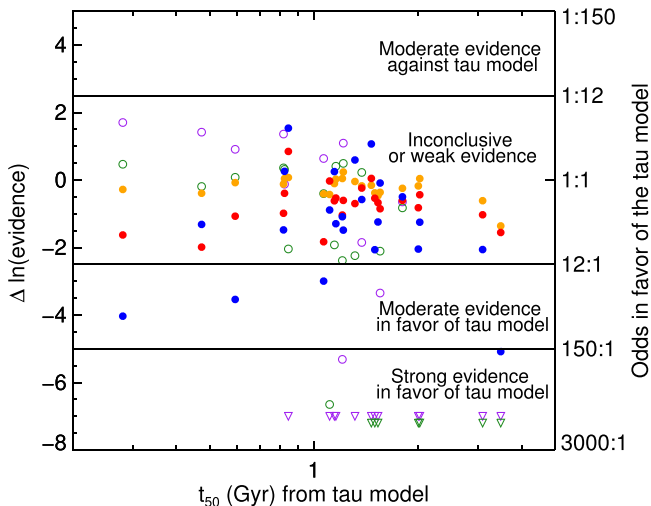

Figure 5. Logarithmic difference between the evidence for a given model and the evidence for the tau model as a function of median stellar age. Each color represents one of the five SFH models, as listed in Figure 3. Triangles indicate upper limits. The right-hand y axis gives the equivalent odds ratio, which is equal to $\exp(\Delta \ln \text{evidence})$. Indicative regions of increasing evidence are marked following Trotta (2008).

The most robust measurements are therefore the median age t_{50} and the sSFR averaged over the last 100 Myr, which are listed for each galaxy in Table 2. When looking at parameters that are not directly related to the SFH, different models give consistent results. The stellar mass is particularly robust, with a difference between the piecewise and tau models of only 0.02 ± 0.04 dex (mean and standard deviation, respectively). Discrepancies in dust extinction (-0.04 ± 0.10 mag) and stellar metallicity ($7\% \pm 15\%$) are also small.

3.4. Optimal SFH Models

We now ask the question of whether the data favor any of the six models over the others. To answer this, we take advantage of the nested sampling algorithm employed by pyspecfit, which, in addition to the posterior distribution, also returns the model *evidence*. This quantity is the integral of the likelihood, weighted by the prior, over the full parameter space, and its value is proportional to the support of the data for a given model. In practice, model evidence is only relative: the ratio of the evidence of two models can be interpreted as the ratio of the odds in favor of one model against the other. Once again, we take the tau model as reference, and for each galaxy, we define the quantity $\Delta \ln \text{evidence}$ as the logarithmic difference between the evidence for a given model and the evidence for the tau model.

We show the distribution of $\Delta \ln \text{evidence}$ for the entire sample as a function of galaxy age in Figure 5. This quantity can be directly translated to the odds that the tau model explains the data better than another model, and these are shown on the right-hand y axis. We also mark the regions of the plot corresponding to weak, moderate, and strong evidence according to a popular, though somewhat arbitrary, scale (see, e.g., Trotta 2008). We find that the delayed tau and constant + tau models fare as well as the tau model. On the other hand, the SSP and top-hat models are strongly disfavored. It is particularly interesting to note that the SSP actually fares slightly better than the tau model for young galaxies but is unable to well describe older galaxies, for which a more extended SFH is clearly needed. Similar conclusions apply to the top-hat model.

The piecewise model is comparable to the tau model except at young ages, where it performs moderately worse, presumably because of its rather coarse binning. However, since the evidence is particularly sensitive to the adopted priors, the comparison of nonparametric and parametric models must be performed with care. For example, the maximum allowed stellar age for the piecewise model is fixed to 4 Gyr (see Section 3.1), while for the parametric models, we adopt the age of the universe at the observed redshift. To test whether this has any effect on the results, we run an additional spectral fit using a piecewise SFH where the last node, instead of being fixed at 4 Gyr, is equal to the age of the universe at the observed redshift. The resulting evidence is, on average, unchanged, and individual points move in Figure 5 by $|\Delta \ln \text{evidence}| < 1.5$. We also test for the width of the priors, which we vary by one order of magnitude in each direction (i.e., adopting LogNormal (0, 1) and LogNormal(0, 3) on each of the SFH nodes; see Table 1). The quantity $\Delta \ln \text{evidence}$ tends to decrease in both cases, on average by 1.0 and 0.5, respectively, and the result shown in Figure 5 is qualitatively unchanged. The physical properties, such as median ages and stellar masses, are remarkably stable. For an in-depth discussion of the effects of the prior when using nonparametric models, we refer to Leja et al. (2018).

Keeping in mind that the evidence is weighted by the size of the parameter space, we conclude that the extra degrees of

freedom of the piecewise models (six free parameters, as opposed to the two of the tau model) are not warranted, since the final results are not measurably better than those obtained with simpler models. This means that the secondary bursts recovered by the piecewise model in some cases, such as the one shown in Figure 1, are possible but not required by the data. We caution, however, against a simplistic interpretation of this result: in the context of SFH models, having more degrees of freedom than required by the data is not necessarily a poor choice (see, e.g., Simha et al. 2014). For example, if one is interested in the measurement of a specific parameter, such as the median stellar age, it may be a good idea to allow more freedom than strictly required by the data.

Our analysis can be summarized as follows: the tau, delayed tau, and constant + tau models yield consistent results, whereas the SSP and top-hat models show clear limitations in reproducing the data. The piecewise model is the most flexible one but does not describe the data significantly better than the simple tau model. It seems that the constant + tau model strikes a balance between simplicity and flexibility, and for this reason, we choose it as our "fiducial" model.

3.5. The Hα and [N II] Emission Lines

Our additional, shallower H- and K-band spectroscopic observations target the Hα emission for 19 out of 24 galaxies in the sample. We find only one clear detection of Hα, for 35616, the object we removed from the main analysis. The [N II] line is also detected, and, although Hα is contaminated by a strong sky line, we are able to derive a flux ratio [N II]/Hα ∼ 0.8, which is high and suggests a contribution from active galactic nucleus (AGN) activity (e.g., Kewley et al. 2013), consistent with the fact that this is one of the three galaxies in the sample that are detected in the X-ray. About 10% of quiescent galaxies at high redshift harbor this type of emission (Belli et al. 2017a). Since 35616 is in a close interaction, it is possible that the AGN activity is triggered by gas that has been tidally stripped from the companion.

For the remaining galaxies in the sample, we scale the observed spectrum so that the median level of the continuum around Hα matches the best-fit template from the spectral fit to ensure a proper flux calibration. We then subtract the best-fit template to remove the contribution of the stellar absorption. We calculate upper limits on the nondetections by integrating the error spectrum in a window of $500\,\mathrm{km\,s^{-1}}$ (corresponding to a line width $\sigma \sim 200\,\mathrm{km\,s^{-1}}$) centered on the Hα wavelength. After correcting the line fluxes for the extinction due to dust, adopting the attenuation A_V obtained from the spectral fit, we convert to values of star formation rate (based on a Chabrier IMF) following Kennicutt (1998). The 3σ upper limits on the star formation rates are listed in Table 2. The upper limits range from 1 to 15 $M_\odot\,\mathrm{yr^{-1}}$; the highest values correspond to cases in which strong sky lines enter the Hα spectral window. For each galaxy, the limit on the star formation rate derived from Hα is higher than the value predicted by the spectral fit.

To obtain a deeper measurement, we shift each of the 18 spectra to the rest frame and stack them. The result is shown in the top panel of Figure 6: both Hα and [N II] are clearly detected. The bottom panel shows what we obtain when stacking the spectra without subtracting the stellar template. In this case, the emission from [N II] is detected, while Hα is
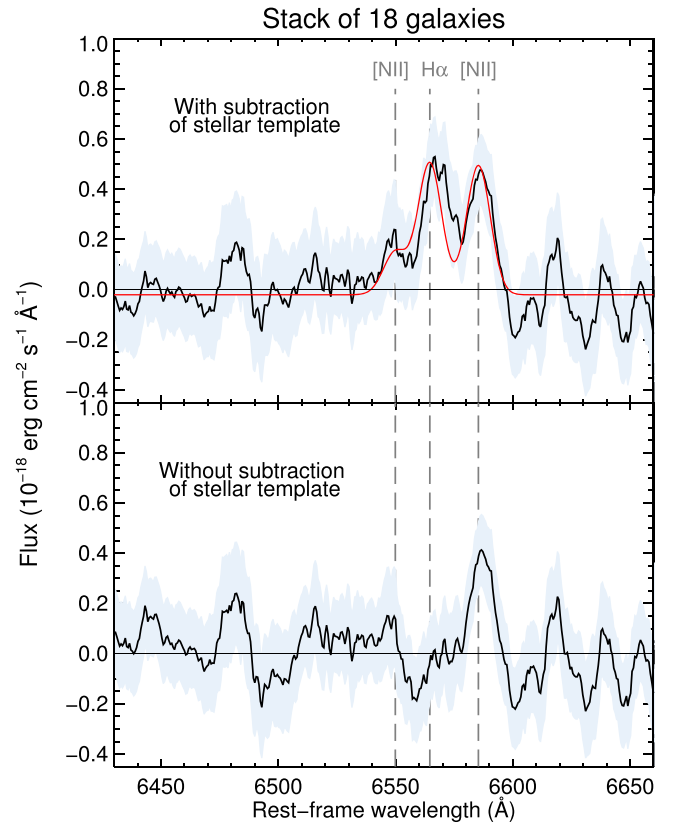

Figure 6. Stacked Hα spectrum for 18 galaxies in our sample. The blue shaded region marks the uncertainty, and the red line is the Gaussian fit to the emission lines. The top panel shows the result when the best-fit stellar template is subtracted from each spectrum before stacking; the bottom panel shows the result obtained when simply subtracting the median continuum from each spectrum.

marginally detected in absorption. Without a detailed template of the stellar continuum, it would not be possible to measure such a small amount of Hα emission.

We model the emission lines Hα, [N II] 6548, and [N II] 6584 as three Gaussians. We fix the wavelengths to their theoretical values and the flux ratio of the two [N II] lines to 3 and set the velocity dispersion of the three lines to be identical. The best fit of this model to the continuum-subtracted stack is shown in red in Figure 6. We obtain an Hα flux of $(6.7 \pm 0.8) \cdot 10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$, a flux ratio of [N II]/Hα = 1.0 ± 0.1, and a velocity dispersion of $233 \pm 17\,\mathrm{km\,s^{-1}}$. One galaxy, 30475, has a marginally detected [N II] line (it is also one of the three galaxies of the sample detected in the X-ray). Removing it from the stack has a minimal impact on the results. Further splitting the sample into two random subsets does not change the results of the fit in a statistically significant way, confirming that the ionized emission is not due to any individual object. We also test the effect of the SFH model on the subtraction of the stellar template: adopting the best-fit piecewise model instead of the constant + tau model, we obtain nearly identical emission-line measurements.

The high [N II]/Hα ratio we obtain for the stacked spectrum rules out star formation activity as the main source of emission lines. In the local universe, quiescent galaxies commonly feature weak emission lines due to the ionization from old stellar populations (Cid Fernandes et al. 2011). However, the
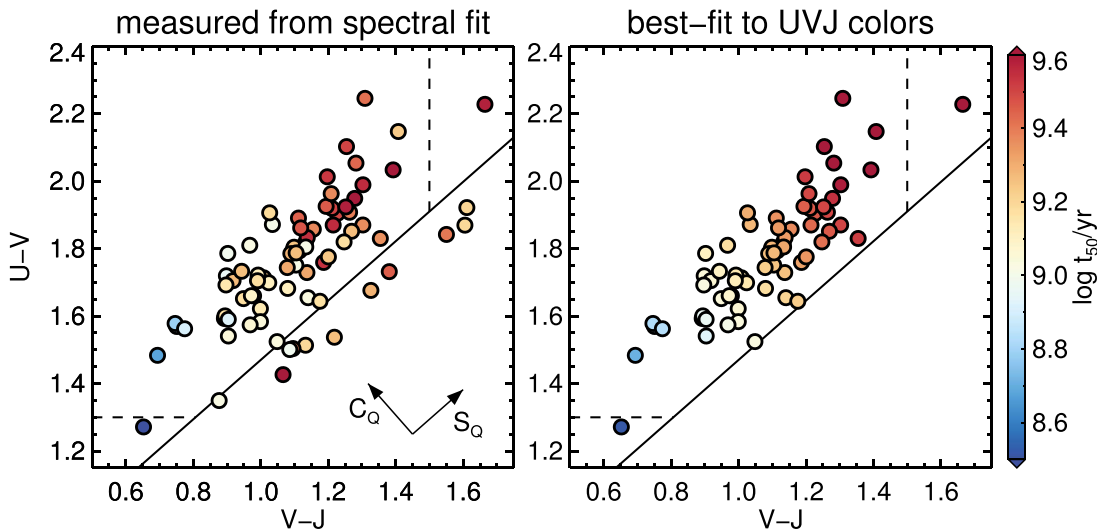
Figure 7. Left: *UVJ* diagram for both the MOSFIRE and LRIS samples, color-coded by the median stellar age t_{50} as measured from fitting models to the spectroscopic and photometric data. The rotated coordinates S_Q and C_Q are shown in the bottom right corner. Right: same as left panel but color-coded by the median age inferred from the *UVJ* colors using Equation (3). In this panel, only galaxies within the quiescent region, defined by the diagonal black line, are shown. The dashed lines mark the edges of the quiescent population proposed by Muzzin et al. (2013b), which we do not use in our analysis.

maximum Hα equivalent width predicted for this mechanism is 3 Å, while we measure 4.7 ± 0.6 Å in our stack. Therefore, we conclude that the most likely source for the ionized gas emission is low-level AGN activity. Ongoing star formation could still contribute to the observed flux, which therefore can be used to derive a strict upper limit to the star formation rate. After correcting the observed Hα flux for the average value of dust extinction in the sample ($A_V = 0.43$), we obtain an star formation rate upper limit of 0.9 ± 0.1 $M_\odot$ yr^{-1}, which confirms the truly quiescent nature of the sample. Dividing by the average stellar mass, we obtain log sSFR ~ -11.1, which is fully consistent with the values derived from the spectral fits and shown in Figure 4.

4. The Age–Color Relation

In this section, we will develop a simple technique to estimate stellar ages for galaxies without using spectroscopic data by taking advantage of the relatively tight relation between median age and rest-frame colors. This will enable the study of a much larger photometric sample, which will be presented in the next section.

In order to probe a larger redshift range and increase the sample size, for the present analysis, we combine the MOSFIRE data with the LRIS observations at $1 < z < 1.5$ from our previous studies (Newman et al. 2010; Belli et al. 2014, 2015). We perform the spectroscopic fits to the LRIS sample as described in Section 3, adopting a constant + tau model. We remove objects with low signal-to-noise ratio or poor fits, as well as objects that are outside the CANDELS field of view and therefore not present in the 3D-*HST* catalog; this leaves 56 galaxies. The results of the spectral fitting for these objects are given in Appendix A. Combining the two Keck samples, we obtain a homogeneous data set of 79 galaxies spanning the redshift range $1 < z < 2.5$ and the stellar mass range $10.5 < \log M_*/M_\odot < 11.5$.

In the left panel of Figure 7, we show the distribution of the median stellar ages t_{50} for the combined Keck sample on the *UVJ* diagram, i.e., the rest-frame $U - V$ versus $V - J$ plane.

This is a useful diagram in which quiescent and star-forming systems tend to form two distinct sequences (e.g., Wuyts et al. 2007). We take the rest-frame colors from the 3D-*HST* catalog, which are therefore independent of our Keck spectra. The figure shows that the distribution of stellar ages is remarkably regular, and that the red sequence is a sequence in age. This is not a new result: we already showed this using our LRIS sample at $1 < z < 1.5$ in Belli et al. (2015), and an independent study by Domínguez Sánchez et al. (2016) confirmed this finding. At $z \sim 2$, Whitaker et al. (2013) and Mendel et al. (2015) used stacked spectra to show that the blue end of the sequence consists of younger galaxies. However, our sample presents two important advantages over previous studies: it has a relatively large number of individual measurements, so that we do not have to rely on stacking, and it is based on a robust estimate of the stellar age that does not depend on the SFH assumptions, as we have shown in Section 3. We point out that we also detect the age trend along the red sequence in the MOSFIRE sample alone, as shown in Appendix B.

The diagonal black line marks our definition of the quiescent region, which we take from Muzzin et al. (2013b). The original definition included a vertical boundary at $V - J = 1.5$ and a horizontal one at $U - V = 1.3$, shown as dashed lines. Recent studies have suggested that these additional constraints exclude galaxies that are quiescent and dusty (e.g., Whitaker et al. 2015) and that have been quenched abruptly (Merlin et al. 2018), respectively. Interestingly, the age trend does seem to continue beyond the very blue and very red ends of the quiescent box. For these reasons, we do not consider these edges and simply use the diagonal line to define the quiescent population. The smooth age trend breaks down just outside this diagonal line. This happens for a variety of reasons, including a higher dust attenuation and a stronger contribution of the [O II] emission line to the *U*-band flux of star-forming galaxies. The latter effect depends strongly on the templates used to derive the rest-frame colors: when adopting purely stellar templates, as we did in Belli et al. (2015), these objects fall inside the quiescent box. For consistency, throughout this work, we use
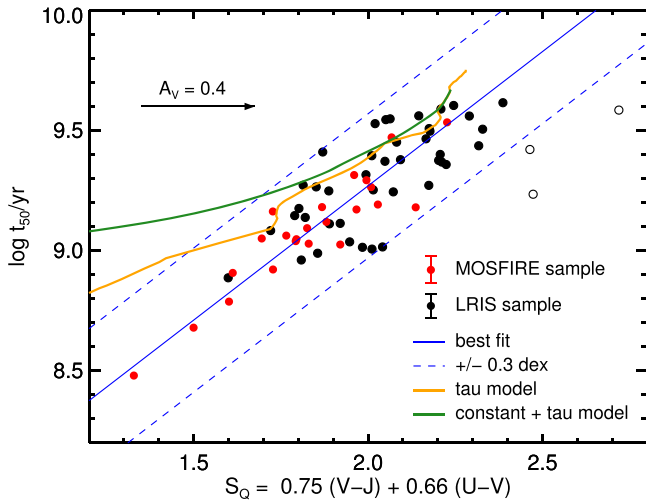
Figure 8. Measured median stellar age as a function of the S_Q coordinate for both the MOSFIRE (red) and LRIS (black) samples. The median age uncertainties are shown by the error bars in the legend; open symbols are outliers and excluded from the fit. The solid blue line shows the best fit given in Equation (3), and the dashed blue lines mark the region within 0.3 dex from the best fit. The tracks for two simple models of SFHs are also shown: a tau model with $\tau = 100$ Myr (orange) and a constant star formation of 2 Gyr followed by an exponential decline with $\tau = 100$ Myr (green). The models stop 5.7 Gyr (the age of the universe at $z \sim 1$) after the onset of star formation. The horizontal arrow shows the change in S_Q due to a dust attenuation $A_V = 0.4$, which is the mean value measured for the galaxies in the Keck sample.

rest-frame colors derived by exclusively fitting the broadband photometry and using the same templates (from Brammer et al. 2008).

To simplify the quantitative analysis of the age trend, we introduce a rotated system of coordinates on the *UVJ* diagram following Fang et al. (2018). The rotated axes S_Q and C_Q, illustrated in the left panel of Figure 7, are respectively parallel and perpendicular to the boundary of the quiescent box as defined in Muzzin et al. (2013b), which has a slope of 0.88. Thus, the new coordinate system corresponds to a rotation of $\theta_Q = \arctan 0.88 = 41°\!.3$ and is described by the following transformations:

$$S_Q = 0.75\,(V - V) + 0.66\,(U - V),$$
$$C_Q = -0.66\,(V - J) + 0.75\,(U - V). \qquad (2)$$

The angle θ_Q is slightly different from that used by Fang et al. (2018), $\theta = 34°\!.8$, which is appropriate for the sequence of star-forming galaxies. In order to distinguish the two sets of coordinates, we add the subscript "Q" to ours to indicate that these apply to the quiescent population. It is easy to see that the stellar ages are only a function of S_Q and do not depend on C_Q. Selecting only the objects in the quiescent region, we show the relation between the logarithmic stellar age and rotated coordinate S_Q in Figure 8. Most galaxies lie on a remarkably tight relation. To better understand this relation, we also show the theoretical tracks for two simple, dust-free SFHs: one is a tau model, and the other is a model where the star formation rate is constant for 2 Gyr and then falls off exponentially. Both models are significantly offset from the observed distribution, particularly for ages younger than 1 Gyr. This is not surprising, since dust extinction has a strong impact on the value of S_Q. Our spectroscopic fits reveal the presence of some dust in most

quiescent galaxies, with a mean value of $\langle A_V \rangle = 0.4$, which corresponds to a reddening effect of $\langle \Delta S_Q \rangle = 0.33$, shown by the horizontal arrow in the figure. We conclude that the observed tight relation between stellar age and rest-frame colors cannot be predicted by a simple model and is actually due to a combination of stellar population aging and evolution of the dust content.

The model tracks reach a maximum value of S_Q around 2.3. The observed galaxies that have colors much redder than this must be affected by substantial dust extinction. This is particularly true for the three objects (shown as open symbols) that are redder than $S_Q \sim 2.4$, are clearly outliers, and have dust attenuation values among the highest in the sample, $A_V \sim 0.8$–1. Excluding these objects, we fit a linear relation to the data, obtaining

$$\log t_{50}/\mathrm{yr} = 7.03 + 1.12 \cdot S_Q$$
$$= 7.03 + 0.84\,(V - J) + 0.74\,(U - V), \qquad (3)$$

which is shown in blue in the figure. Virtually all galaxies lie within ± 0.3 dex of this relation; the only exceptions are the three dusty galaxies. It is therefore reasonable to limit the validity of Equation (3) to $S_Q < 2.4$. All galaxies redder than this value are particularly dusty, and a proper estimate of their stellar ages must rely on direct spectroscopic observations. Excluding the outliers, the standard deviation of the age residuals is 0.13 dex, or 35%. This means that it is possible, in most cases, to predict the median stellar age of a quiescent galaxy at high redshift just by measuring the rest-frame $U - V$ and $V - J$ colors. The scatter of this relation is significantly larger than the typical age uncertainty (0.05 dex), meaning that spectroscopic data are still needed for a precise measurement. Nonetheless, Equation (3) is a very effective tool for estimating stellar ages from photometric observations. We show this in the right panel of Figure 7, where we color-code the Keck sample by the stellar ages inferred from the *UVJ* colors using the best-fit relation, finding a striking agreement with the values measured from the spectroscopy, shown in the left panel.

To check that this age–color relation holds for the full population of quiescent galaxies, we plot the age residuals as a function of redshift, stellar mass, stellar metallicity, and dust attenuation in Figure 9. The residuals are remarkably flat in redshift and stellar mass and show a small trend with metallicity and dust attenuation. This is expected, since both metallicity and dust can redden the observed SED of a galaxy, but the intensity of this effect is small, pointing once more at a correlation between different galaxy properties that has the effect of reducing the scatter in color for a given stellar age. However, given the poor constraint on the stellar metallicity obtained with our spectral fits, it is possible that part of the remaining scatter in the age–color relation is driven by metallicity.

In principle, from the bottom panel of Figure 9, it is possible to derive a correction term that depends on the dust attenuation A_V and would lead to an even tighter relation between color and age. However, the physical properties shown in the figure are derived from the fit to both spectroscopy and photometry; in the absence of spectra, the degeneracy between dust and age would lead to a substantially larger uncertainty on the dust
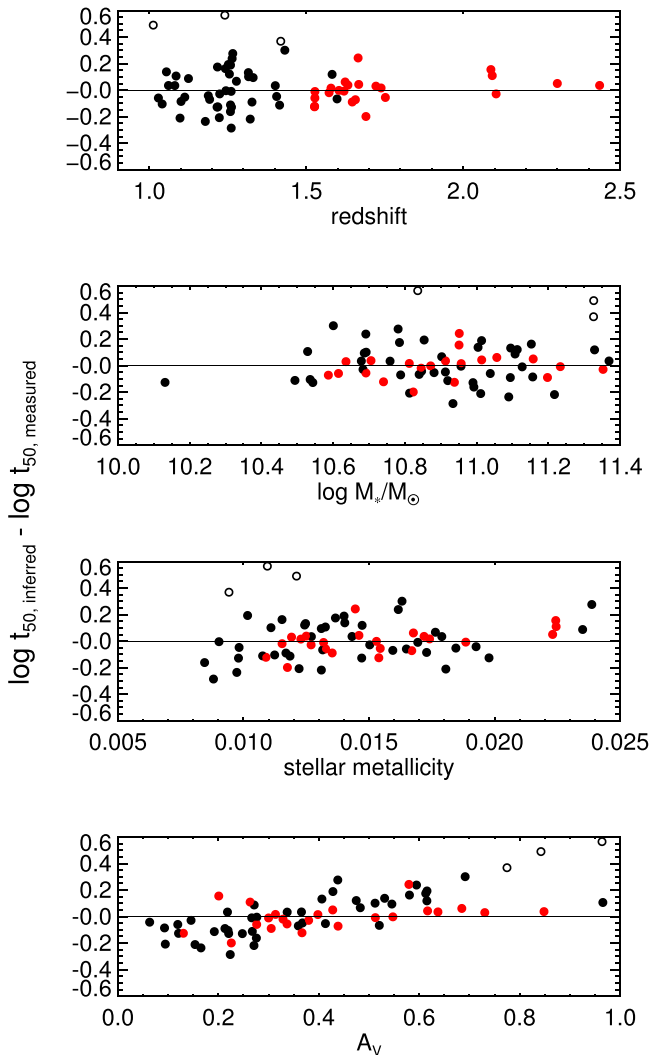
Figure 9. Logarithmic difference between inferred and measured median ages as a function of redshift, stellar mass, stellar metallicity, and dust attenuation A_V (top to bottom). All physical properties are derived from the spectroscopic fit. Symbols are the same as in Figure 8.

attenuation (see, e.g., Belli et al. 2015), thus limiting the accuracy of such a correction.

5. Number Density Evolution of the Quiescent Population

Now that we have a method to estimate the stellar ages without the need for spectroscopic data, we can explore the evolution of quiescent galaxies in large photometric surveys. The goal of this section is to measure number densities and study their evolution with cosmic time, a task currently not practical with spectroscopic samples due to their relatively small size and complex selection function.

We consider the galaxy population in the UltraVISTA survey (McCracken et al. 2012), which covers a wide area (1.62 deg^2) and thus enables the measurement of number densities with a high degree of precision. We take the Muzzin et al. (2013a) catalog of sources in the UltraVISTA field and follow the sample selection performed by Muzzin et al. (2013b) for studying the galaxy mass function: we exclude all objects that are fainter than $K_S = 23.4$, flagged as stars, or located in bad regions or that have contaminated photometry. We also

restrict our analysis to the redshift range in which the age–color relation has been calibrated, $1 < z < 2.5$, and to galaxies with $\log M_*/M_\odot > 10.8$, which is the 95% mass-completeness threshold at $z \sim 2.5$ (Muzzin et al. 2013b). The resulting sample consists of 5335 galaxies, and its distribution on the *UVJ* diagram is shown in Figure 10, split into four redshift bins. Galaxies are color-coded according to their inferred median age, which we derive using Equation (3). Star-forming galaxies or quiescent galaxies outside the calibrated color range are shown as black dots. The four redshift bins have been defined so that each bin spans the same comoving volume. This means that the number of galaxies shown in the figure can be directly interpreted as a number density. For example, it can be easily seen that the population of massive quiescent galaxies grows substantially between $z \sim 2.5$ and 1.

Since the quiescent population forms a sequence of increasing stellar age, the simplest model to explain the observations is one where all galaxies appear at the blue end and, as they age, move along the sequence toward the red end. This simple hypothesis has strong implications for galaxy quenching, as we will discuss below. It is therefore important to test whether the data are in agreement with this model. Combining number densities and stellar ages, we are now able to perform this test.

We start by selecting the population of quiescent galaxies, using the red box in Figure 10. Then we select, among the quiescent galaxies, those that have young stellar ages. Following popular terminology, we refer to these as *post-starburst* galaxies, since they are observed a short time after their main phase of star formation (independently of whether that was a burst). Looking at the age–color relation for the Keck sample, shown in Figure 8, we find that an age threshold of 800 Myr gives a pure selection, meaning that virtually all galaxies with an inferred age below this threshold happen to have a spectroscopic age that is also below this value. We also limit the definition to galaxies that are older than 300 Myr, the youngest age we measure in the Keck sample, simply because the age–color relation has not been tested beyond this limit. The selection for post-starburst galaxies is therefore $300\,\mathrm{Myr} < t_{50} < 800\,\mathrm{Myr}$ and is shown as a blue box in Figure 10.

We count the number of objects in each of the selection regions and divide by the comoving volume probed in each redshift bin, obtaining number density measurements. These are listed in Table 3 and shown in the top panel of Figure 11 as a function of redshift for the quiescent galaxies (red line) and the subset of post-starburst galaxies (thick blue line). Uncertainties are calculated by combining in quadrature the Poisson error with the cosmic variance, which is calculated following Moster et al. (2011). Since the value of 800 Myr is arbitrary, we vary the post-starburst selection by moving the threshold to 1 Gyr and 600 Myr. The corresponding number densities are plotted as thin blue lines, and the effect of this change on the edge of the selection box is shown in the first panel of Figure 10. We find that post-starburst galaxies have a roughly constant number density over the entire redshift range. However, as the overall population of quiescent galaxies grows rapidly with cosmic time, the subset of post-starburst galaxies becomes increasingly less relevant, going from being 34% of the total quiescent population at $z \sim 2.5$ to just 4% at $z \sim 1$.

Next, we measure the rate of growth of the red sequence by taking the difference of the number density of quiescent
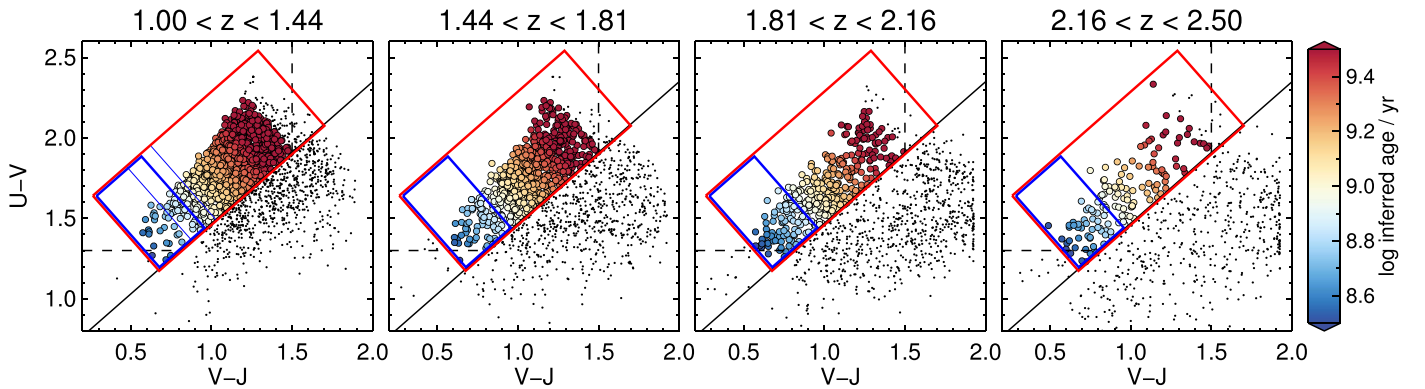
Figure 10. Distribution of the rest-frame *UVJ* colors for all galaxies with $\log M_*/M_\odot > 10.8$ in the UltraVISTA survey, split into four redshift bins spanning equal comoving volume. The points are color-coded by inferred median ages using Equation (3). Black dots represent galaxies for which the inferred age cannot be calculated, either because they are outside the range in which the age–color relation has been calibrated, or because they are star-forming objects. The red box indicates the selection region for quiescent galaxies, and the blue box indicates that for post-starburst galaxies, defined as having $300\,\mathrm{Myr} < t_{50} < 800\,\mathrm{Myr}$. In the first panel, thin blue lines show how the edge of the selection box changes when adopting a maximum age for post-starburst galaxies of 600 Myr or 1 Gyr.

Table 3
Galaxy Number Densities in the UltraVISTA Survey

	$1.00 < z < 1.44$	$1.44 < z < 1.81$	$1.81 < z < 2.16$	$2.16 < z < 2.50$
All quiescent galaxies	-3.69 ± 0.04	-3.98 ± 0.05	-4.27 ± 0.06	-4.63 ± 0.07
$300\,\mathrm{Myr} < t_{50} < 1\,\mathrm{Gyr}$	-4.79 ± 0.06	-4.77 ± 0.06	-4.71 ± 0.07	-4.96 ± 0.08
$300\,\mathrm{Myr} < t_{50} < 800\,\mathrm{Myr}$	-5.05 ± 0.07	-4.96 ± 0.07	-4.87 ± 0.07	-5.10 ± 0.09
$300\,\mathrm{Myr} < t_{50} < 600\,\mathrm{Myr}$	-5.41 ± 0.09	-5.32 ± 0.09	-5.13 ± 0.08	-5.30 ± 0.10

Note. The quantities listed correspond to $\log N/\mathrm{Mpc}^{-3}$, where N is the number density. Only galaxies with $\log M_*/M_\odot > 10.8$ are considered.

galaxies in adjacent bins divided by the cosmic time interval between the centers of the bins. This is plotted in red in the bottom panel of Figure 11. We then compare this to the flux of galaxies going through the post-starburst region. This is given by the number density of post-starburst galaxies divided by the time it takes them to cross the post-starburst selection region. Assuming a passive evolution, galaxies will take 500 Myr to cross the post-starburst region, going from an age of 300 Myr to an age of 800 Myr. We show this number density growth with a thick blue line in the figure. It is clear that the flux of galaxies through the post-starburst region is not sufficient to explain the observed growth of the quiescent population. At $z \sim 2.2$, the observed growth rate is about twice that inferred from the post-starburst population. At $z \sim 1.4$, this ratio has increased to ~ 5 times. Moreover, this result does not depend on the age threshold used in the definition of the post-starburst population, as shown by the thin blue lines in Figure 11. Changing the age threshold has a large impact on the measured number densities but not on the growth rates, because of the corresponding change in the time spent in the post-starburst region. This is an indirect validation of our inferred age technique.

Since the flux of post-starburst galaxies is much smaller than the observed growth of the quiescent population, we must conclude that many galaxies appear on the red sequence when their median ages are already older than 1 Gyr. On the *UVJ* diagram, this would correspond to crossing the diagonal line dividing quiescent and star-forming systems without ever entering the post-starburst region. This is the main result of our study and has important consequences for the study of galaxy quenching. Since post-starburst galaxies are both quiescent and

young, they must have been quenched in a very rapid way, on timescales of the order of a few hundred Myr. However, our result shows that most massive galaxies did not undergo this phase and instead must have been quenched slowly, to allow their stellar populations to age while they were still star-forming.

The growth rates that we measured are based on ages inferred from photometry alone, which cannot be as accurate as spectroscopic measurements. However, given the large galaxy population used in the derivation, the precision in the inferred age of a single object is not relevant. What matters is that we are not missing post-starbust galaxies that are shifted outside the selection box by heavy dust reddening. Figure 8 shows that no such galaxies are present in the Keck sample. The outliers in the age–color relation are dusty but also old and are so rare (see black points in Figure 11) that their impact on the number density calculation is limited. On the other extreme of the color distribution, some galaxies in the UltraVISTA survey are bluer than the youngest objects in the Keck sample and have not been considered in the number density calculation. This is also a rare population, and extending the selection region to include these objects does not change the results significantly. We also assumed, in our derivation, that galaxies evolve passively during the post-starburst phase. We note that this is a conservative assumption: any deviation from passive evolution, such as frosting or rejuvenation, will make the colors bluer and increase the visibility time for the post-starburst galaxies, therefore decreasing their contribution to the growth of the red sequence and strengthening our result.

In our analysis, we neglected the contribution of mergers to the number density evolution of quiescent galaxies. This is a
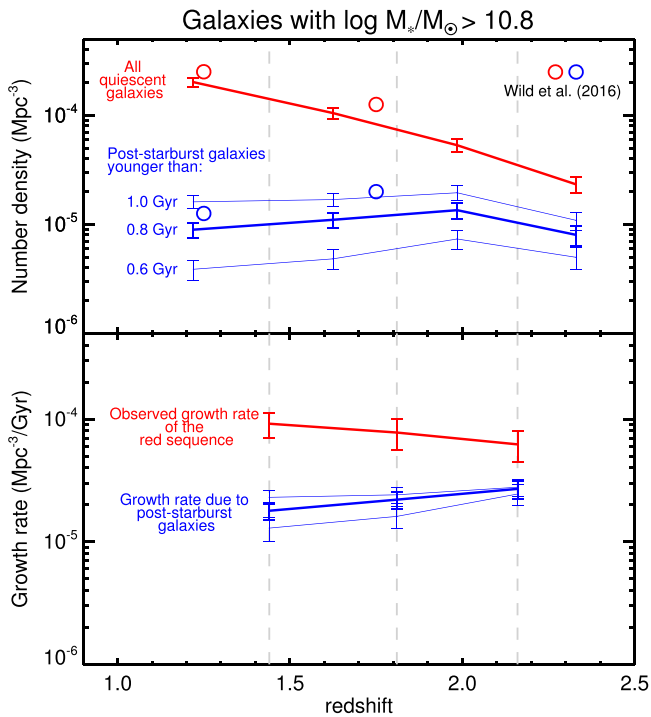
Figure 11. Top: number density of quiescent galaxies (red) and post-starburst galaxies (blue) with $\log M_*/M_\odot > 10.8$ in the UltraVISTA survey. Circles mark the values reported by Wild et al. (2016). Bottom: observed growth rate of the quiescent population (red) and rate of growth due to the flux of post-starburst galaxies (blue). In both panels, the thin blue lines correspond to changes in the threshold age used to select post-starburst galaxies. The vertical dashed lines mark the redshift bins, which are the same as those in Figure 10.

reasonable approximation, given that in the time spanned by our redshift bins, only one in 10 galaxies will undergo a major merger (e.g., Man et al. 2016b). In Appendix C, we carry out a quantitative assessment of merging, which causes two distinct and opposite effects. On the one hand, massive quiescent galaxies merge together, thus decreasing the number density of the quiescent population; on the other hand, quiescent systems with $\log M_*/M_\odot < 10.8$ can grow in mass entirely via gas-poor mergers, therefore increasing the number density of the population considered in our analysis. In Appendix C, we show that these two effects are one to two orders of magnitude smaller than the observed growth of the quiescent population over $1 < z < 2.5$ and can therefore be neglected.

Our result differs from that of a previous study by Whitaker et al. (2012), who found that the growth of the red sequence at $z > 1$ is consistent with the aging of post-starburst galaxies. However, Whitaker et al. based their selection of post-starburst galaxies on the results of SED fitting, obtaining a window on the *UVJ* diagram that is substantially wider than ours. In another study of post-starburst galaxies, Wild et al. (2016) adopted a selection based on a principal component analysis of galaxy SEDs. While this selection is, in principle, completely independent from our method, it gives very similar results in terms of distribution of *UVJ* colors (Wild et al. 2014). Not surprisingly, we find number densities for both the quiescent and post-starburst populations in agreement with those measured by Wild et al. (2016), shown as circles in Figure 11. Wild et al. correctly stated that all quiescent galaxies must have gone through the post-starburst phase at $z \sim 2$ if the visibility time for that phase was 250 Myr. Given

our high-quality Keck spectra, we can now robustly measure this timescale, which is significantly larger than 250 Myr, and therefore conclude that at most half of the quiescent galaxies at $z \sim 2$ were post-starburst at some point in the past. Finally, studies at lower redshift found that the contribution of post-starburst galaxies to the growth of the red sequence declines strongly from $z \sim 1$ to 0 (Wild et al. 2009; Dressler et al. 2013; Rowlands et al. 2018). This is in qualitative agreement with Figure 11, which shows that post-starburst galaxies become less important at later cosmic times.

6. Discussion: Fast and Slow Quenching

6.1. Two Quenching Channels

Our analysis of the rest-frame colors and number densities demonstrates that not all quiescent galaxies go through the post-starburst phase after being quenched. This implies a variety of quenching timescales: some galaxies undergo a fast quenching, becoming post-starburst and then evolving along the red sequence, while the remaining ones must follow a slow-quenching path that never produces the blue colors typical of post-starburst galaxies. We illustrate these two paths on the *UVJ* diagram in Figure 12 using two simple toy models. The fast-quenching track, shown in blue, is exemplified by a tau model with a short timescale, $\tau = 100$ Myr, while for the slow track, in green, we take $\tau = 1$ Gyr. We also include the effect of dust reddening, which we model in an identical way for the two tracks: the dust attenuation is $A_V = 2$ at the beginning, declines following the evolution of the star formation rate, and then reaches a plateau at $A_V = 0.4$, the typical value found for our quiescent galaxies.

The toy models are effective at illustrating the salient features of the fast and slow tracks. Clearly, the red sequence on the *UVJ* diagram can be joined at any point along its diagonal edge, a region often called the green valley. The exact position of the crossing point will depend on both the quenching timescale and the dust attenuation. The qualitative difference between the two tracks is, however, mainly due to the quenching timescale. For $\tau = 1$ Gyr, for example, even a dust-free model will not enter the post-starburst region. In order to reproduce the colors of the most extreme post-starburst galaxies, which are even bluer than our selection box, it is necessary to have little dust and extremely short timescales, $\tau < 100$ Myr. The rarity of these very blue quiescent galaxies, evident in Figure 12, implies that galaxies move across that region very rapidly, consistent with an extrapolation of our age–color relation.

Our number density analysis leads to a quantitative result: in order to explain the observed color distribution of quiescent galaxies, the two quenching channels must be in place simultaneously and have similar quenching rates at $z \sim 2$. This is consistent with the qualitative result of our earlier study of the LRIS sample (Belli et al. 2015), in which we proposed that post-starburst galaxies follow a fast-quenching path based on the derived SFHs and the observation that they have systematically smaller sizes than older quiescent galaxies. Other studies, at both low and high redshift, have reached similar conclusions (e.g., Schawinski et al. 2014; Moutard et al. 2016; Wild et al. 2016; Carnall et al. 2018; Forrest et al. 2018; Wu et al. 2018a). In these works, the two quenching channels are often characterized not only by different timescales but
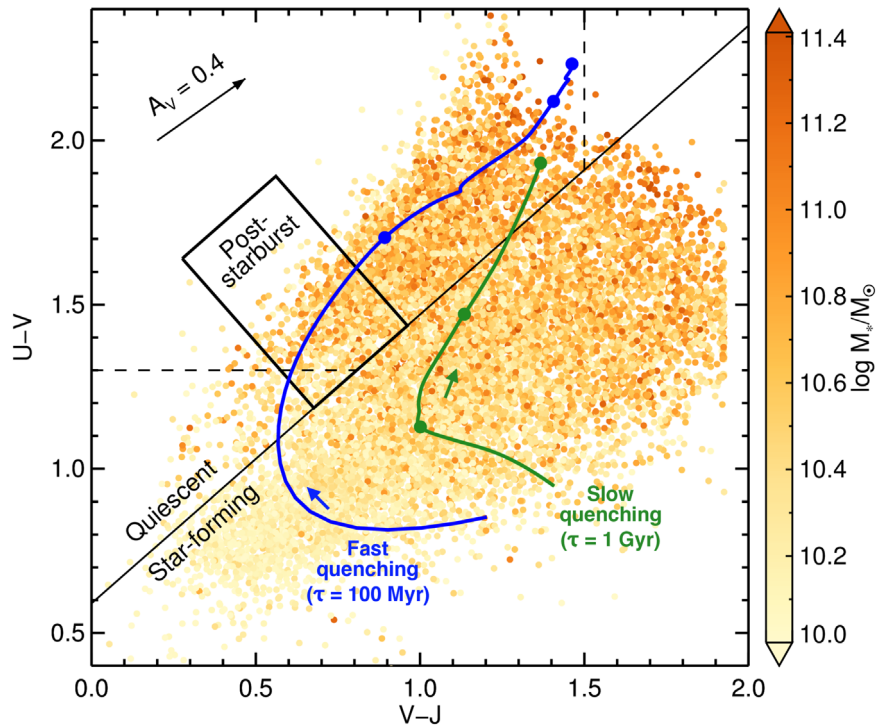
Figure 12. Illustration of the fast- and slow-quenching paths on the *UVJ* diagram. The distribution of all galaxies in the UltraVISTA catalog with $1.5 < z < 2.0$ and $\log M_*/M_\odot > 10$ is shown in the background, color-coded by stellar mass. The diagonal line represents our adopted division into star-forming and quiescent galaxies; the dashed lines are additional constraints used by Muzzin et al. (2013b), which we do not use. The black rectangle marks the region that we use to select post-starburst galaxies. The colored lines show two tau models, one with $\tau = 100$ Myr (blue) and the other with $\tau = 1$ Gyr (green); the filled points on the tracks mark even intervals of 1 Gyr. In both cases, we adopt a simple model for the dust attenuation, which evolves following the SFR, with the constraints of $A_V = 2$ at the beginning and $A_V = 0.4$ at the end. The arrow in the top left corner shows the effect of dust attenuation on the *UVJ* colors.

also by different morphological or structural properties, as we discuss below.

6.2. Physical Properties of Post-starburst Galaxies

Given their importance in the context of galaxy quenching, we want to explore in more detail the properties of post-starburst systems. For this analysis, we make use of the catalog from the 3D-*HST* survey (Brammer et al. 2012), which covers a much smaller area than UltraVISTA but is based on high spatial resolution imaging and grism spectroscopy from *HST*. In order to explore a variety of physical properties, we assemble different catalogs that have been publicly released and include grism redshifts (Momcheva et al. 2016), rest-frame colors and stellar population properties (Skelton et al. 2014), size and morphological properties (van der Wel et al. 2014), and measurements of the local overdensities (Fossati et al. 2017). Given the limited survey area, the number of post-starburst galaxies is small; to increase the statistics, we slightly extend the mass range, selecting all objects with $\log M_\odot/M_* > 10.6$. Still, we find only four galaxies with rest-frame colors in the post-starburst region in the redshift range $1 < z < 1.5$; therefore, we limit the analysis to the range $1.5 < z < 2.5$, where the cosmic volume probed is larger. With this selection, we obtain a sample of 554 quiescent galaxies, of which 65 fulfill the post-starburst color selection.

We plot some of the physical properties for this sample of quiescent galaxies in Figure 13 as a function of the rest-frame color S_Q, defined in Equation (2). Using the age–color relation, we can then convert this into an inferred stellar age, shown on the top axis. The post-starburst galaxies are those with an inferred age between 300 and 800 Myr. The figure shows that this population has clearly distinct properties compared to the older quiescent systems. Post-starburst systems tend to have lower stellar masses, consistent with the mass function study performed by Wild et al. (2016). To compare the structural properties, we calculate a mass-normalized size for each object by taking the measured effective radius and dividing it by the size typical for quiescent galaxies at that stellar mass and redshift, which we take from the van der Wel et al. (2014) mass–size relation. The distribution of normalized sizes shows that post-starburst galaxies are more compact than the overall quiescent population. This was first shown by Whitaker et al. (2012) and successively confirmed by a number of studies (Belli et al. 2015; Yano et al. 2016; Almaini et al. 2017; Maltby et al. 2018; Wu et al. 2018a). The distribution of Sérsic indices does not vary strongly with stellar age, with a marginal increase toward the post-starburst region, indicating that these systems are likely spheroidal (Almaini et al. 2017; Maltby et al. 2018). Finally, in the last panel, we show the distribution of the local overdensity, finding that roughly 80% of post-starburst galaxies are found in an overdense environment ($\delta > 0$), as opposed to 25% of the older quiescent galaxies. Despite residing in overdensities, post-starburst galaxies are not likely to be satellites: using the Fossati et al. (2017) catalog, we find that three quarters of the post-starburst galaxies shown in Figure 13 are centrals with a probability larger than 90%. Our results are consistent with the excess of post-starburst galaxies that has been found in clusters at $z < 1$ (Moutard et al. 2018; Socolovsky et al. 2018). However, these post-starburst galaxies
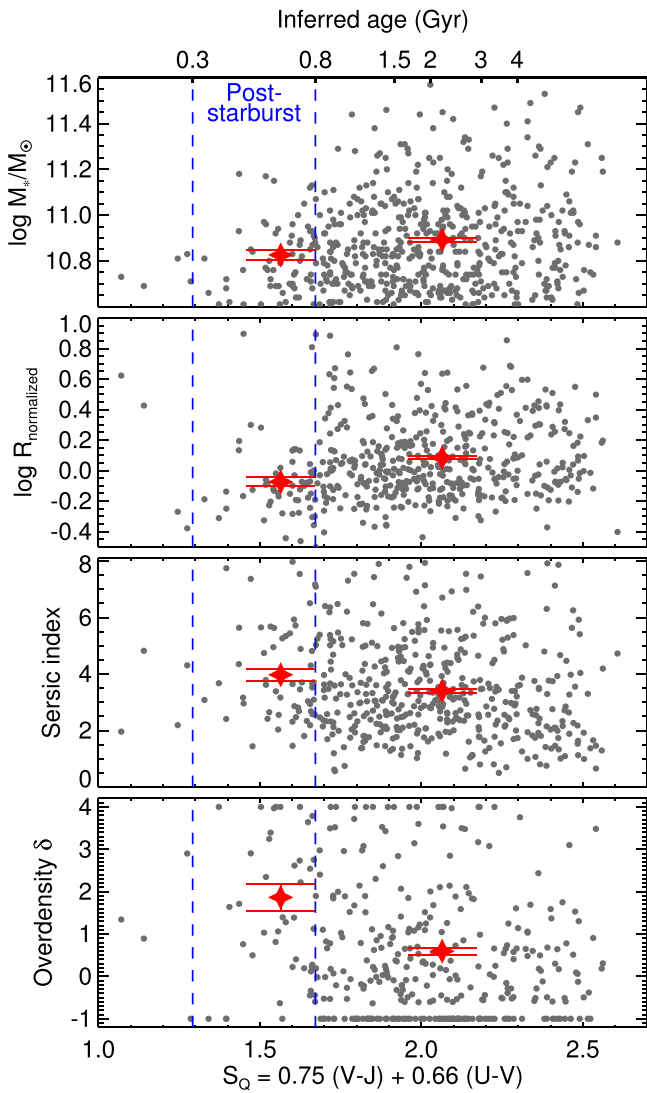
Figure 13. Properties of quiescent galaxies from the 3D-*HST* survey, with $\log M_*/M_\odot > 10.6$ and $1.5 < z < 2.5$, as a function of their rest-frame color S_Q, which we convert to inferred stellar age on the top axis using Equation (3). From top to bottom, we show the distribution of stellar mass, mass-normalized effective size, Sérsic index, and overdensity δ within a radius of 0.75 Mpc. The average values for the post-starburst population (300 Myr $< t_{50} < 800$ Myr) and the population of old quiescent galaxies ($t_{50} > 800$ Myr) are shown in red.

at lower redshifts are less massive ($\log M_*/M_\odot < 10.5$) and are likely to be satellites.

The emerging picture is one in which the fast-quenching channel produces post-starburst galaxies that are preferentially compact and spheroidal, have lower stellar masses, and lie in higher density environment compared to the bulk of the quiescent population. The fact that post-starburst galaxies have such distinct physical properties points to a physical difference between fast and slow quenching, rather than to a single physical process characterized by a wide distribution of quenching timescales. One pressing question remains: What are the physical mechanisms responsible for the two quenching channels?

6.3. Fast Quenching

Prime candidates for the fast-quenching mechanism are gas-rich processes that naturally explain both the short timescales

and the compact sizes found in post-starburst galaxies. Major mergers (Hopkins et al. 2006) or violent disk instabilities (Dekel & Burkert 2014) can drive nuclear inflows and cause centrally concentrated starbursts. These processes have been explored in detail in recent theoretical work, specifically because they can explain why massive quiescent galaxies at high redshift are generally very small compared to their local counterparts. Using a suite of hydrodynamical simulations, Zolotov et al. (2015) showed that mergers or violent disk instabilities can trigger a *compaction* event, in which a massive stellar core is formed in a short time, with a resulting decrease in the galaxy effective size. Quenching follows naturally due to a combination of gas exhaustion and stellar feedback; AGN feedback may also be involved, particularly for the maintenance of quiescence over longer timescales. However, these simulations predict a combined timescale for compaction and quenching of the order of 2 Gyr (see also Tacchella et al. 2016), which is substantially longer than what we measure for the post-starburst objects and more similar to our slow-quenching channel. This discrepancy may indicate the need for stronger feedback or additional physical mechanisms. A hint about the important role played by AGN feedback in this critical phase comes from the high incidence of ionized gas emission (not due to star formation) found in the post-starburst region of the *UVJ* diagram (Belli et al. 2017a).

Compaction can be triggered by a number of different processes, including major or minor mergers, counterrotating streams, perturbation due to giant clumps, and tidal interactions (Zolotov et al. 2015). The large fraction of post-starburst galaxies residing in overdensities suggests a prominent role for mergers and interactions with nearby galaxies (see also, e.g., Brodwin et al. 2013; Wu et al. 2014). We note that post-starburst systems are mostly centrals and therefore not affected by so-called environmental quenching, which has been proposed for explaining the abundance of quiescent low-mass systems in dense environments, and applies only to satellites (Peng et al. 2010). A link with environmental quenching, however, cannot be ruled out at $z < 1$, where the post-starburst population consists of lower-mass galaxies (Maltby et al. 2018).

Compaction is not the only possible explanation for the small sizes of quiescent galaxies at high redshift. The alternative possibility is that these systems formed at very high redshift, when star-forming disks were significantly smaller, and then evolved passively retaining their compact structures (Wellons et al. 2015). However, the young ages we measure for post-starburst galaxies rule out this scenario. We conclude that the observed properties of post-starburst galaxies require a compaction phenomenon, defined in its broadest sense as the formation of quiescent galaxies that are more compact than their star-forming progenitors. To establish exactly how and why this happens, more theoretical and observational studies are clearly needed.

In Section 5, we found that the importance of the fast-quenching channel increases with redshift, up to $z \sim 2.5$. At even higher redshift, it is likely that fast quenching represents, in practice, the only available track, simply because of the young age of the universe. The earliest quiescent galaxy with a spectroscopic confirmation is a massive system at $z = 3.717$ that features strong Balmer absorption lines, typical of post-starburst systems (Glazebrook et al. 2017). Interestingly, this object happens to have a massive close companion
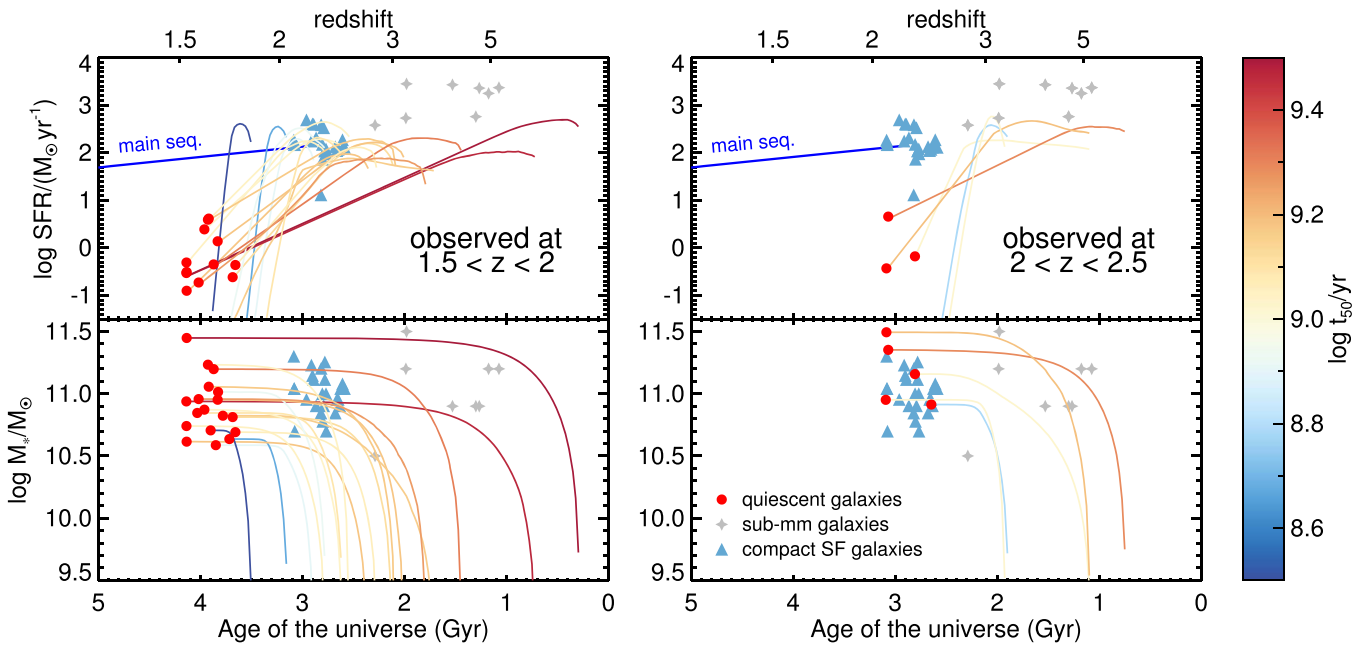
Figure 14. Reconstructed SFHs (top) and stellar mass growth (bottom) for the quiescent galaxies in our MOSFIRE sample, adopting the constant + tau model. The left and right panels show two redshift bins, corresponding to galaxies observed at $1.5 < z < 2$ and $2 < z < 2.5$, respectively. The lines are colored according to the median stellar age t_{50}, and the red points mark the observed redshift. In the top panels, the blue line indicates the evolution of the main sequence for a fixed stellar mass of $10^{11} M_\odot$ from Whitaker et al. (2014). The sample of submillimeter galaxies from Toft et al. (2014) is shown as gray diamonds, while the compact star-forming galaxies from van Dokkum et al. (2015) are shown as blue triangles.

(Simpson et al. 2017; Schreiber et al. 2018b). It is possible that the very first massive quiescent galaxies, which reside in overdense environments, were all quenched rapidly via major mergers.

6.4. Slow Quenching

Normal star-forming galaxies have depletion times of about 1 Gyr (e.g., Genzel et al. 2010), which means that a continuous replenishment of the gas reservoir via gas inflows is needed to continue the star formation activity over a cosmological timescale. If such replenishment is halted for any reason, the galaxy slowly consumes the remaining cold gas and quenches over long timescales. Such gas starvation is therefore the simplest explanation for the slow-quenching path. Unfortunately, there are many different physical mechanisms that can produce gas starvation. The most commonly invoked scenario is one where the gas cannot cool when the halo grows beyond a threshold mass due to the formation of a stable virial shock (Birnboim & Dekel 2003; Kereš et al. 2005), possibly in conjunction with heating by AGN radio-mode feedback (Bower et al. 2006; Croton et al. 2006). There are, however, other possibilities: the gas may cool but then be made stable against collapse by the deep potential well of a large stellar bulge (Martig et al. 2009; Genzel et al. 2014), or the gas inflows could stop even before reaching the halo (Feldmann & Mayer 2015).

In the local universe, the stellar metallicity of quiescent galaxies is substantially larger than that of star-forming objects of the same mass, which is a signature of starvation (Peng et al. 2015). This is consistent with the increasing importance of the slow-quenching channel at later cosmic epochs.

6.5. Identifying the Star-forming Progenitors

Another way to constrain the quenching mechanisms is to link quiescent galaxies with their star-forming progenitors. Using the SFHs from our spectral fits, we can estimate physical properties such as stellar mass and star formation rate for the progenitors of the quiescent galaxies in our sample. We caution that this analysis is strongly affected by the assumption of the SFH model and should be considered as purely qualitative.

In Figure 14, we show the inferred past evolution in both star formation rate (top) and stellar mass (bottom) for our MOSFIRE sample, split into two redshift bins. The tracks shown in the figure correspond to the constant + tau model for the SFH and are color-coded according to the median stellar age. We compare the inferred evolution with the properties of spectroscopically confirmed submillimeter galaxies (Toft et al. 2014) and compact star-forming galaxies (van Dokkum et al. 2015), two populations that have been proposed as possible progenitors of massive quiescent galaxies. The masses and star formation rates of compact star-forming galaxies are consistent with those of the progenitors of quiescent galaxies, at least for those observed at $1.5 < z < 2$, and are also consistent with the star-forming main sequence (e.g., Whitaker et al. 2014). On the other hand, the submillimeter galaxies have substantially higher values of star formation rate, which are not consistent with the SFH tracks. However, this discrepancy can be easily explained by the typical duty cycle of submillimeter galaxies, which may be as short as 40 Myr (Toft et al. 2014); such a short-lived phase would leave virtually no signature on the spectra observed more than a Gyr later. Figure 14 also shows that the peak value of star formation rate is roughly the same for all quiescent galaxies, around 100–300 $M_\odot$ yr^{-1}, and does not depend on the median stellar age. This means that the SFHs of post-starburst galaxies do not require a burst of star formation

above the main sequence; they simply require rapid formation and quenching.

In the compaction scenario, galaxies first shrink and then quench. The population of compact star-forming galaxies is thought to represent the intermediate stage after compaction and before quenching (Barro et al. 2013; van Dokkum et al. 2015). Measurements of the kinematics have been used to draw a link between compact star-forming galaxies and quiescent systems. The integrated velocity dispersion of the gas in compact star-forming galaxies is in broad agreement with that of the stellar populations in quiescent galaxies (Barro et al. 2014; Nelson et al. 2014). A comparison of the spatially resolved kinematics, though, would be even more compelling. Recent observations showed that the gas in compact star-forming galaxies is actually rotating rapidly in both the ionized (van Dokkum et al. 2015; Wisnioski et al. 2018) and molecular (Barro et al. 2017; Tadaki et al. 2017) components. This is not in agreement with the kinematics of the most massive quiescent galaxies at $z \sim 0$, which tend to be pressure-supported. At high redshift, however, quiescent galaxies seem to be very different from local ellipticals.

Our previous analysis of the MOSFIRE sample, which focused on the stellar kinematics, suggested that the rotational support of quiescent galaxies is significantly larger at high redshift (Paper I). This surprising result is consistent with detailed studies of a few gravitationally lensed systems at $z > 2$ (Newman et al. 2015, 2018b; Toft et al. 2017) and a study of a large sample at $z < 1$ (Bezanson et al. 2018). The fact that massive quiescent galaxies at high redshift present a large degree of rotational support has two important consequences in the context of quenching. First, it suggests that the giant ellipticals observed at $z \sim 0$ were originally quiescent disks, which lost their angular momentum likely because of dry mergers. This means that, contrary to the standard view, the morphological transformation does not need to happen simultaneously with the quenching process in order to form ellipticals. Second, it offers an additional constraint on the quenching process, which needs to produce a rotating stellar system. Since starvation leaves the morphology and kinematics untouched, the slow-quenching channel should produce preferentially rotating quiescent systems. However, gas-rich mergers can also produce rotating remnants (Wuyts et al. 2010). It is therefore not clear whether kinematics can be used for a clean identification of the progenitors.

The question of whether the two quenching channels require distinct populations of progenitors is still open. It is possible that all star-forming galaxies can potentially undergo fast quenching after a close interaction or merger. On the other hand, post-starburst galaxies show smaller stellar masses compared to older quiescent systems, and the tracks shown in Figure 12 also indicate a mass segregation between fast and slow quenching. If this were the case, star-forming galaxies above a certain stellar mass (which depends on redshift) would be the ideal candidates to be progenitors of the slow-quenching track.

7. Summary and Conclusions

7.1. The SFHs of $z \sim 2$ Massive Quiescent Galaxies

We analyzed the deep near-infrared spectra of 24 massive quiescent galaxies at $1.5 < z < 2.5$ obtained with Keck/MOSFIRE. The high quality of the spectroscopic data, which we now release publicly, allowed us to perform a study of the stellar population properties of these objects to an unprecedented level of detail.

We took full advantage of all of the available data by fitting templates simultaneously to the observed spectroscopy and broadband photometry. We explored six different SFH models with the goal of understanding how this impacts the derived stellar population properties. We found that the SSP and top-hat models are not able to reproduce the observed data for at least some galaxies in our sample. The remaining four models were all equally successful, including the commonly used tau model, which has only two free parameters. Despite the degeneracy among different SFH models, the median stellar ages are found to be robust, together with the sSFR averaged over the past 100 Myr. However, the duration of the star formation depends strongly on the SFH shape and therefore cannot currently be measured robustly.

In order to break the degeneracy between different SFH models, significantly deeper spectra would be needed. This may be possible for individual objects, but for larger samples, we will have to wait for the next generation of ground-based telescopes. A complementary approach is to use the stellar metallicity to indirectly probe the formation timescale (Kriek et al. 2016), which, however, requires a wider wavelength range than that available for our sample.

7.2. Ionized Gas in Quiescent Galaxies

By stacking shallower spectra obtained at redder wavelengths, we detected the Hα and [N II] emission lines. The Hα line is faint and would not be detected without a proper subtraction of the underlying stellar absorption. We measured a high line flux ratio, [N II]/H$\alpha = 1.0 \pm 0.1$, which rules out star formation as the sole origin of the ionization. About 10% of quiescent galaxies at high redshift feature emission lines with a similar flux ratio, which is likely due to AGN-driven outflows (Belli et al. 2017a). Our finding suggests that low-level AGN activity is even more common in quiescent galaxies but in most cases is too weak to be noticed. This is also confirmed by the sample of gravitationally lensed quiescent galaxies in Newman et al. (2018a), where the lensing magnification allows us to observe faint emission lines in most of the objects. Such low-level AGN feedback may play a role in either quenching the star formation or maintaining the galaxies as quiescent.

Assuming that the dust-corrected Hα flux is entirely due to star formation yields a strict upper limit on the star formation rate, which is 0.9 ± 0.1 $M_\odot \, \mathrm{yr}^{-1}$. This corresponds to an average $\log \mathrm{sSFR} \sim -11.1$, confirming the results of the spectral fitting. This low value of star formation is in broad agreement with other studies that used less-direct methods (Fumagalli et al. 2014; Man et al. 2016a; Gobat et al. 2017) and suggests a very effective quenching in massive galaxies already at $z \sim 2$.

7.3. The Age–Color Relation

By combining the MOSFIRE observations with our previous LRIS sample at lower redshift, we found that the majority of quiescent galaxies at $1 < z < 2.5$ lie on a remarkably tight relation between stellar age and UVJ colors, which is independent of redshift and stellar mass. Such a relation deviates from that predicted by a simple model of passive evolution due to the contribution of dust reddening, which

evolves rapidly during the post-starburst phase. The use of a large sample of high-quality spectra is therefore critical for determining the age–color relation. We quantified this relation and its scatter, which is only 0.13 dex in age, and we gave an explicit calibration that can be used to infer the approximate stellar age of any high-redshift quiescent galaxy, given its rest-frame colors, without the need for spectroscopic data. This is a simple yet powerful method that allowed us to model the evolution of the quiescent population and that can also be used for other applications, such as selecting specific targets for spectroscopic follow-up.

Based on our observations and on previous studies (e.g., Whitaker et al. 2015; Merlin et al. 2018), we argued that the lower and upper edges commonly used to define the quiescent region in the *UVJ* diagram are not justified and artificially exclude very young and very dusty quiescent galaxies. The young quiescent galaxies, also called post-starburst, represent a key evolutionary phase in the formation of the red sequence. The nature of the dusty objects (which are the only galaxies in our sample that do not lie on the age–color relation), instead, is currently not known: it is possible that they consist of edge-on disks and/or interacting systems.

It would be interesting to extend the age–color relation to higher redshifts; however, the few spectroscopic measurements available at $z > 3$ show that *UVJ* colors may become less effective in dividing quiescent and star-forming galaxies (Schreiber et al. 2018a).

7.4. Galaxy Quenching

Combining the age–color relation, which we derived from our spectroscopic sample, with a large, volume-limited sample drawn from the UltraVISTA survey, we were able to model the evolution of the red sequence at $1 < z < 2.5$ in a quantitative and self-consistent way. The main result of this analysis is the identification of two quenching channels: fast quenching produces compact post-starburst objects, while slow quenching contributes to the growth of the red sequence by adding objects with stellar populations that are already old. Fast quenching is increasingly more important at high redshift, contributing about half of the growth of the quiescent population at $z \sim 2$.

Post-starburst objects are clearly distinct from the bulk of the red sequence and feature lower stellar masses, compact sizes, spheroidal morphologies, and higher overdensities, as also shown by a number of recent studies. Together with their rapid formation timescale, these properties well fit a scenario in which violent, gas-rich events such as mergers are responsible for their formation. The slow-quenching channel, instead, must be due to a smooth decline in the gas reservoir, which may be caused by fuel starvation. However, there are many physical processes that may be responsible for starvation, including AGN feedback, virial shock heating, and cosmological or gravitational effects.

Appendix A
Results of the Spectral Fitting for the LRIS Sample

The stellar population properties, measured via spectral fitting, for the LRIS sample are listed in Table 4.

Table 4
Stellar Population Properties of the LRIS Sample

ID (3D-HST)	Field	z	t_{50}[a] (10^9 yr)	A_V[a] (mag)	Z[a]	$\log M_*/M_\odot$[a]	$\log$ sSFR[a,b] (yr^{-1})
17678	GOODS-N	1.60	0.77 ± 0.07	0.52 ± 0.07	0.013 ± 0.002	10.84 ± 0.01	−11.8 ± 3.7
23303	COSMOS	1.33	0.91 ± 0.05	0.55 ± 0.05	0.013 ± 0.002	10.69 ± 0.01	−14.7 ± 6.9
33158	EGS	1.40	1.0 ± 0.2	0.76 ± 0.12	0.016 ± 0.003	10.92 ± 0.02	−9.9 ± 0.1
28523	COSMOS	1.58	1.0 ± 0.05	0.62 ± 0.04	0.015 ± 0.002	11.33 ± 0.01	−24.9 ± 8.6
4748	GOODS-N	1.27	1.0 ± 0.1	0.44 ± 0.05	0.024 ± 0.003	10.78 ± 0.03	−22.3 ± 9.0
17533	EGS	1.26	1.0 ± 0.08	0.59 ± 0.05	0.016 ± 0.003	10.69 ± 0.02	−10.4 ± 0.06
21628	COSMOS	1.43	1.0 ± 0.1	0.69 ± 0.09	0.016 ± 0.003	10.60 ± 0.02	−11.0 ± 2.4
1433	GOODS-N	1.40	1.1 ± 0.1	0.31 ± 0.09	0.016 ± 0.005	10.52 ± 0.02	−10.2 ± 0.07
4275	COSMOS	1.22	1.1 ± 0.09	0.61 ± 0.07	0.014 ± 0.004	10.79 ± 0.02	−10.5 ± 0.05
41108	GOODS-S	1.22	1.2 ± 0.4	0.12 ± 0.05	0.020 ± 0.004	10.13 ± 0.05	−11.3 ± 2.5
22782	EGS	1.08	1.3 ± 0.2	0.22 ± 0.06	0.018 ± 0.004	10.68 ± 0.04	−11.4 ± 4.8
36048	EGS	1.28	1.3 ± 0.1	0.48 ± 0.05	0.018 ± 0.004	10.90 ± 0.03	−10.4 ± 0.05
12167	EGS	1.19	1.4 ± 0.3	0.36 ± 0.06	0.016 ± 0.004	10.79 ± 0.04	−13.1 ± 3.8
19419	COSMOS	1.26	1.4 ± 0.3	0.27 ± 0.06	0.011 ± 0.002	10.49 ± 0.04	−16.6 ± 8.2
23909	EGS	1.22	1.5 ± 0.2	0.22 ± 0.06	0.015 ± 0.005	10.54 ± 0.04	−11.3 ± 1.0
9636	COSMOS	1.08	1.6 ± 0.2	0.42 ± 0.07	0.018 ± 0.004	10.26 ± 0.02	−10.2 ± 0.06
796	COSMOS	1.15	1.6 ± 0.2	1.12 ± 0.06	0.020 ± 0.003	11.12 ± 0.03	−10.6 ± 0.07
16571	COSMOS	1.10	1.6 ± 0.2	0.95 ± 0.09	0.026 ± 0.005	10.65 ± 0.04	−10.7 ± 0.08
30221	COSMOS	1.24	1.7 ± 0.3	0.96 ± 0.06	0.011 ± 0.003	10.84 ± 0.03	−28.2 ± 16.8
8962	COSMOS	1.09	1.8 ± 0.4	0.97 ± 0.06	0.013 ± 0.003	10.53 ± 0.04	−12.4 ± 2.5
23084	COSMOS	1.04	1.8 ± 0.2	0.22 ± 0.04	0.011 ± 0.002	10.54 ± 0.02	−11.3 ± 0.08
30145	COSMOS	1.40	1.8 ± 0.3	0.34 ± 0.07	0.013 ± 0.003	10.76 ± 0.03	−15.4 ± 12.7
19719	COSMOS	1.14	1.8 ± 0.2	0.53 ± 0.05	0.014 ± 0.003	10.94 ± 0.02	−10.2 ± 0.05
19796	COSMOS	1.26	1.8 ± 0.09	0.28 ± 0.03	0.009 ± 0.001	10.99 ± 0.02	−11.5 ± 0.06
28619	COSMOS	1.25	1.9 ± 0.2	0.62 ± 0.04	0.010 ± 0.003	10.85 ± 0.02	−11.3 ± 0.08
25627	COSMOS	1.10	1.9 ± 0.3	0.15 ± 0.04	0.018 ± 0.002	11.01 ± 0.04	−15.0 ± 4.3
20342	COSMOS	1.10	2.0 ± 0.2	0.61 ± 0.07	0.014 ± 0.003	11.05 ± 0.02	−10.5 ± 0.07
38867	EGS	1.11	2.1 ± 0.2	0.41 ± 0.04	0.018 ± 0.003	10.88 ± 0.02	−10.7 ± 0.04
26884	EGS	1.24	2.3 ± 0.1	0.58 ± 0.04	0.012 ± 0.002	11.15 ± 0.02	−11.2 ± 0.06
27877	EGS	1.05	2.3 ± 0.3	0.53 ± 0.05	0.014 ± 0.003	11.00 ± 0.02	−11.6 ± 0.7
5087	EGS	1.41	2.4 ± 0.3	0.37 ± 0.06	0.010 ± 0.002	10.91 ± 0.03	−11.2 ± 0.07
10703	COSMOS	1.26	2.4 ± 0.3	0.47 ± 0.04	0.012 ± 0.003	11.11 ± 0.02	−11.4 ± 0.07
17070	GOODS-N	1.24	2.4 ± 0.1	0.28 ± 0.03	0.009 ± 0.001	10.96 ± 0.01	−30.6 ± 22.8
42466	GOODS-S	1.42	2.5 ± 0.3	0.19 ± 0.06	0.012 ± 0.002	10.92 ± 0.02	−12.4 ± 8.5
8381	COSMOS	1.41	2.5 ± 0.4	0.45 ± 0.07	0.010 ± 0.002	10.92 ± 0.03	−10.3 ± 0.08
30838	COSMOS	1.32	2.5 ± 0.4	0.51 ± 0.06	0.011 ± 0.002	10.69 ± 0.03	−11.2 ± 0.07
20106	EGS	1.06	2.6 ± 0.5	0.93 ± 0.08	0.013 ± 0.004	11.07 ± 0.04	−10.3 ± 0.09
19083	COSMOS	1.26	2.6 ± 0.2	0.22 ± 0.04	0.009 ± 0.001	10.93 ± 0.01	−11.2 ± 0.04
43042	GOODS-S	1.42	2.6 ± 0.5	0.77 ± 0.05	0.009 ± 0.005	11.33 ± 0.03	−33.6 ± 26.1
4527	COSMOS	1.26	2.7 ± 0.3	0.43 ± 0.06	0.014 ± 0.003	11.01 ± 0.02	−12.3 ± 10.1
2543	COSMOS	1.33	2.8 ± 0.3	0.21 ± 0.07	0.012 ± 0.002	11.09 ± 0.02	−11.9 ± 2.5
16164	EGS	1.19	2.9 ± 0.4	0.58 ± 0.08	0.018 ± 0.004	10.95 ± 0.04	−10.5 ± 0.07
11783	COSMOS	1.26	2.9 ± 0.3	0.27 ± 0.03	0.017 ± 0.003	11.13 ± 0.02	−11.2 ± 0.04
27472	EGS	1.22	3.1 ± 0.5	0.15 ± 0.06	0.015 ± 0.004	10.68 ± 0.04	−11.6 ± 0.1
13041	GOODS-N	1.32	3.2 ± 0.4	0.41 ± 0.05	0.012 ± 0.002	11.09 ± 0.03	−11.6 ± 0.09
14060	COSMOS	1.19	3.2 ± 0.2	0.06 ± 0.03	0.019 ± 0.002	10.85 ± 0.02	−12.1 ± 0.1
30393	EGS	1.18	3.4 ± 0.5	0.17 ± 0.04	0.010 ± 0.002	11.09 ± 0.05	−11.8 ± 0.09
31555	COSMOS	1.32	3.5 ± 0.3	0.27 ± 0.05	0.013 ± 0.002	11.22 ± 0.03	−11.1 ± 0.05
31613	GOODS-S	1.22	3.5 ± 0.5	0.09 ± 0.05	0.012 ± 0.002	10.81 ± 0.04	−11.8 ± 0.08
22126	EGS	1.06	3.6 ± 0.6	0.37 ± 0.09	0.014 ± 0.003	11.37 ± 0.05	−11.9 ± 2.7
16822	EGS	1.26	3.6 ± 0.4	0.25 ± 0.05	0.010 ± 0.002	10.99 ± 0.04	−11.6 ± 0.07
18045	EGS	1.01	3.8 ± 0.6	0.84 ± 0.08	0.012 ± 0.003	11.33 ± 0.05	−11.1 ± 0.08
24456	EGS	1.10	3.9 ± 0.4	0.09 ± 0.04	0.017 ± 0.002	11.16 ± 0.04	−12.0 ± 0.08
32114	EGS	1.03	4.0 ± 0.5	0.12 ± 0.05	0.016 ± 0.002	11.04 ± 0.04	−11.9 ± 0.09
29863	EGS	1.12	4.1 ± 0.3	0.27 ± 0.05	0.024 ± 0.003	11.11 ± 0.03	−11.4 ± 0.05
30244	COSMOS	1.01	4.3 ± 0.3	0.16 ± 0.04	0.010 ± 0.001	11.08 ± 0.02	−10.6 ± 0.03

Notes. More properties, including coordinates and morphological and dynamical measurements, are available in Table 2 of Belli et al. (2014).
[a] Derived from spectral fitting adopting a constant + tau model. The value and uncertainty are given by, respectively, the median and standard deviation of the posterior distribution.
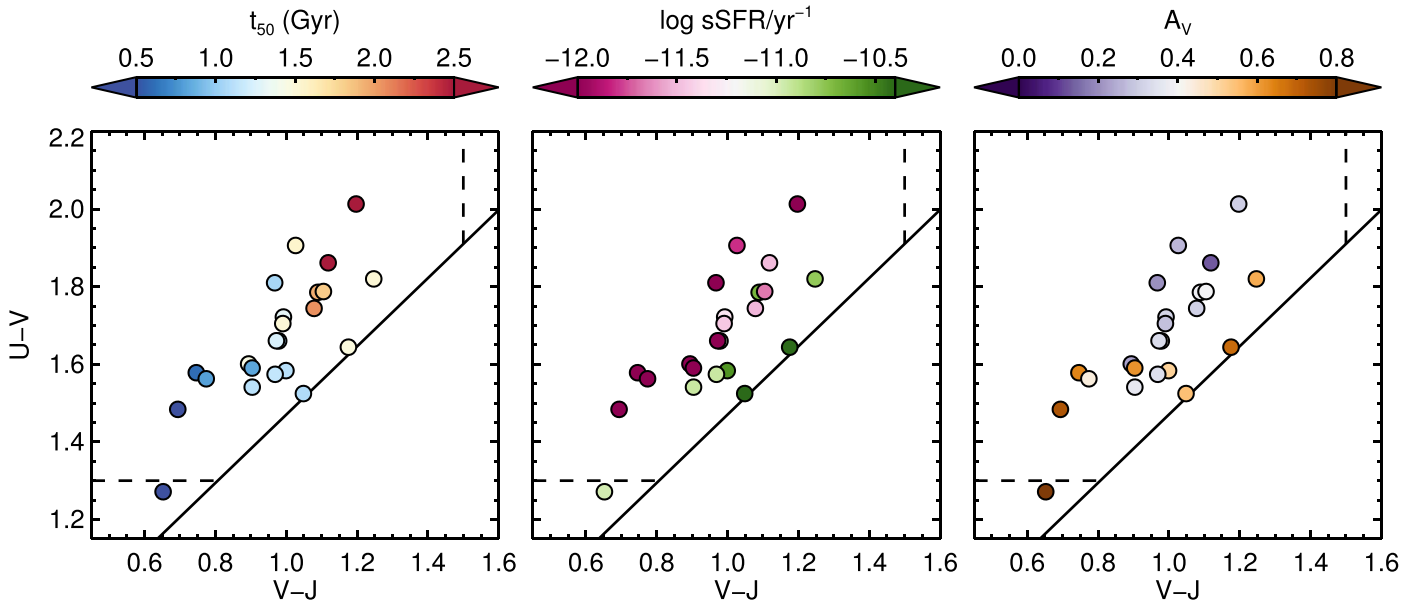[b] Averaged over the past 100 Myr.

Figure 15. Distribution on the *UVJ* diagram of the median stellar age t_{50}, the sSFR averaged over the last 100 Myr, and the dust attenuation A_V for the MOSFIRE sample. The stellar population properties are derived from the spectral fits assuming the constant + tau model; the rest-frame colors are from the 3D-*HST* catalog. The black line marks the separation between quiescent and star-forming galaxies, and the dashed lines are the additional selection criteria used by Muzzin et al. (2013b).

Appendix B
Stellar Populations of the MOSFIRE Sample on the *UVJ* Diagram

In Figure 15, we explore the distribution of the stellar population properties of the MOSFIRE sample on the *UVJ* diagram. The stellar population properties are derived assuming a constant + tau model, while the rest-frame colors are taken from the 3D-*HST* catalog, which are therefore independent from our MOSFIRE data and our own spectral fits. The age–color relation is clearly detected in the first panel. The distribution of sSFR (averaged over the last 100 Myr) also shows a clear trend, but one that is perpendicular to that of the stellar ages. Finally, the dust extinction seems to be higher in both younger and more star-forming galaxies. In Section 3.3, we demonstrated that these three quantities do not change significantly when adopting a different SFH model; therefore, the results shown in Figure 15 are robust. These results are qualitatively consistent with our study at lower redshift (Belli et al. 2015).

Appendix C
Assessing the Impact of Merging on the Number Density Evolution of Quiescent Galaxies

In Section 5, we inferred the rate at which massive galaxies are quenched by measuring the number density evolution of the population of galaxies that are quiescent and more massive than $10^{10.8} M_\odot$, which we will call the *reference population*. In our derivation, however, we neglected the contribution from galaxy merging, which has two distinct and opposite effects on the number density of the reference population: (1) when two galaxies belonging to the reference population merge together, the number density decreases; and (2) when less massive quiescent galaxies increase their stellar mass via gas-poor merging and cross the threshold of $10^{10.8} M_\odot$, thus entering the reference population, the number density increases.

A realistic modeling of merging and its effect on the number density evolution is outside the scope of the present work. However, we can obtain an approximate estimate of the magnitude of this effect by making the following simplifying assumptions.

1. We only consider major mergers, defined as having a mass ratio larger than 1:4. This is justified, given that minor mergers contribute four times less than major mergers to the mass growth of massive galaxies (Man et al. 2016a).

2. Mergers involve a range of mass ratios, but to simplify the calculations, we only consider a fixed mass ratio value. Observations of pair fractions showed that the distribution of mass ratios at high redshift is virtually flat in logarithmic space (Newman et al. 2012), consistent with theoretical predictions (Hopkins et al. 2010). Using this distribution, it is then possible to calculate the average value, which for major mergers is 1:1.85, corresponding to a mass growth for the central galaxy of $\Delta \log M_* \approx 0.2$. We therefore assume that in each major merger, the *central* galaxy increases its stellar mass by this amount, while the *satellite* galaxy disappears.

3. There are substantial uncertainties on the measurement of merger rates. Following the results of Newman et al. (2012), we assume a major merger fraction $f_{mg} = 10\%$, constant with redshift and stellar mass, and an effective merger timescale $\tau_{mg} = 1$ Gyr. This means that at any given moment, 10% of galaxies are observed to be in a major merger, and each merger is visible for 1 Gyr.

4. The mass function of quiescent galaxies, Φ_{qsc}, in this mass and redshift range can be approximated fairly well by a Schechter function with fixed slope $\alpha = -0.4$, characteristic mass $M^* = 10^{10.8} M_\odot$, and normalization that varies with redshift (Muzzin et al. 2013b).

We can now calculate the contribution of merging to the number density evolution of quiescent galaxies. We start by

defining the (redshift-dependent) number density of the reference population:

$$\tilde{\Phi} \equiv \int_{10.8}^{+\infty} \Phi_{\text{qsc}}(\log M)\, d\log M. \tag{4}$$

This number density decreases every time two of the objects in the reference population merge together. In our simple model, this happens every time the satellite galaxy is above the mass threshold of $10^{10.8} M_\odot$, which means that the central must be above $10^{11} M_\odot$. The number density of galaxies leaving the reference population over a timescale τ_{mg} is then

$$f_{\text{mg}} \cdot \int_{11.0}^{+\infty} \Phi_{\text{qsc}}(\log M)\, d\log M = 0.49\, \tilde{\Phi}\, f_{\text{mg}}. \tag{5}$$

This is actually an upper limit, because it includes cases in which a massive quiescent galaxy merges with a star-forming satellite.

At the same time, the reference population will also grow due to less massive objects that cross the mass threshold. In order for this to happen, the central must have a mass in the range $10^{10.6} M_\odot < M_* < 10^{10.8} M_\odot$ and be quiescent, in order for this not to count as a quenching event. The number density of the reference population will then increase by

$$f_{\text{mg}} \cdot \int_{10.6}^{10.8} \Phi_{\text{qsc}}(\log M)\, d\log M = 0.61\, \tilde{\Phi}\, f_{\text{mg}}. \tag{6}$$

We can then calculate the rate at which the reference population number density changes because of these two effects:

$$-\frac{0.49\, \tilde{\Phi}\, f_{\text{mg}}}{\tau_{\text{mg}}} + \frac{0.61\, \tilde{\Phi}\, f_{\text{mg}}}{\tau_{\text{mg}}} = 0.12\, \tilde{\Phi}\, \frac{f_{\text{mg}}}{\tau_{\text{mg}}}. \tag{7}$$

If we use our estimate of $\tilde{\Phi}$ (shown in the top panel of Figure 11 and listed in Table 3), we obtain $0.3 \cdot 10^{-6}\, \text{Mpc}^{-3}\, \text{Gyr}^{-1}$ at $z \sim 2.5$ and $2 \cdot 10^{-6}\, \text{Mpc}^{-3}\, \text{Gyr}^{-1}$ at $z \sim 1$. These values are about two orders of magnitude below the observed growth of the quiescent population shown in the bottom panel of Figure 11. Even allowing for the substantial uncertainties in our assumptions (such as the merger timescale, merger rate, and mean mass ratio), it is clear that merging has a much smaller effect than quenching on the evolution of the number density of quiescent galaxies at $z > 1$.

We point out that the two opposite effects of merging on the number density have similar magnitude, and therefore almost cancel out, because the mass threshold chosen for our study happens to match the knee of the mass function. However, the absolute value of each of these two competing effects is small, $(1.5–10) \cdot 10^{-6}\, \text{Mpc}^{-3}\, \text{Gyr}^{-1}$, and not sufficient to significantly affect our results. We thus conclude that merging can be safely neglected in our analysis.

ORCID iDs

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

References

Almaini, O., Wild, V., Maltby, D. T., et al. 2017, MNRAS, 472, 1401
Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, ApJ, 765, 104
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2017, ApJL, 851, L40
Barro, G., Trump, J. R., Koo, D. C., et al. 2014, ApJ, 795, 145
Belli, S., Genzel, R., Förster Schreiber, N. M., et al. 2017a, ApJL, 841, L6
Belli, S., Newman, A. B., & Ellis, R. S. 2014, ApJ, 783, 117
Belli, S., Newman, A. B., & Ellis, R. S. 2015, ApJ, 799, 206
Belli, S., Newman, A. B., & Ellis, R. S. 2017b, ApJ, 834, 18
Bezanson, R., van der Wel, A., Pacifici, C., et al. 2018, ApJ, 858, 60
Bezanson, R., van Dokkum, P., van de Sande, J., Franx, M., & Kriek, M. 2013, ApJL, 764, L8
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Brodwin, M., Stanford, S. A., Gonzalez, A. H., et al. 2013, ApJ, 779, 138
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Chabrier, G. 2003, PASP, 115, 763
Chauke, P., van der Wel, A., Pacifici, C., et al. 2018, ApJ, 861, 13
Choi, J., Conroy, C., Moustakas, J., et al. 2014, ApJ, 792, 95
Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687
Cimatti, A., Daddi, E., Renzini, A., et al. 2004, Natur, 430, 184
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, ApJ, 780, 33
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Dekel, A., & Burkert, A. 2014, MNRAS, 438, 1870
Domínguez Sánchez, H., Pérez-González, P. G., Esquej, P., et al. 2016, MNRAS, 457, 3743
Dressler, A., Oemler, A., Jr., Poggianti, B. M., et al. 2013, ApJ, 770, 62
Fang, J. J., Faber, S. M., Koo, D. C., et al. 2018, ApJ, 858, 100
Feldmann, R., & Mayer, L. 2015, MNRAS, 446, 1939
Forrest, B., Tran, K.-V. H., Broussard, A., et al. 2018, ApJ, 863, 131
Fossati, M., Wilman, D. J., Mendel, J. T., et al. 2017, ApJ, 835, 153
Franx, M., Labbé, I., Rudnick, G., et al. 2003, ApJL, 587, L79
Fumagalli, M., Franx, M., van Dokkum, P., et al. 2016, ApJ, 822, 1
Fumagalli, M., Labbé, I., Patel, S. G., et al. 2014, ApJ, 796, 35
Gallazzi, A., Bell, E. F., Zibetti, S., Brinchmann, J., & Kelson, D. D. 2014, ApJ, 788, 72
Genzel, R., Förster Schreiber, N. M., Lang, P., et al. 2014, ApJ, 785, 75
Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, MNRAS, 407, 2091
Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, Natur, 544, 71
Gobat, R., Daddi, E., Strazzullo, V., et al. 2017, A&A, 599, A95
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Hopkins, P. F., Bundy, K., Croton, D., et al. 2010, ApJ, 715, 202
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1
Kennicutt, R. C., Jr. 1998, ApJ, 498, 541
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kewley, L. J., Dopita, M. A., Leitherer, C., et al. 2013, ApJ, 774, 100
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, Natur, 540, 248
Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2018, arXiv:1811.03637
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maltby, D. T., Almaini, O., Wild, V., et al. 2018, MNRAS, 480, 381
Man, A. W. S., Greve, T. R., Toft, S., et al. 2016a, ApJ, 820, 11
Man, A. W. S., Zirm, A. W., & Toft, S. 2016b, ApJ, 830, 89
Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, ApJ, 707, 250
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
Mendel, J. T., Saglia, R. P., Bender, R., et al. 2015, ApJL, 804, L4
Merlin, E., Fontana, A., Castellano, M., et al. 2018, MNRAS, 473, 2098
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Moutard, T., Arnouts, S., Ilbert, O., et al. 2016, A&A, 590, A103
Moutard, T., Sawicki, M., Arnouts, S., et al. 2018, MNRAS, 479, 2147
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, ApJS, 206, 8
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, ApJ, 777, 18
Nelson, E., van Dokkum, P., Franx, M., et al. 2014, Natur, 513, 394
Newman, A. B., Belli, S., & Ellis, R. S. 2015, ApJL, 813, L7
Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018a, ApJ, 862, 125
Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018b, ApJ, 862, 126
Newman, A. B., Ellis, R. S., Andreon, S., et al. 2014, ApJ, 788, 51

Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162
Newman, A. B., Ellis, R. S., Treu, T., & Bundy, K. 2010, ApJL, 717, L103
Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, ApJ, 808, 161
Pacifici, C., Oh, S., Oh, K., Lee, J., & Yi, S. K. 2016, ApJ, 824, 45
Peng, Y., Maiolino, R., & Cochrane, R. 2015, Natur, 521, 192
Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193
Renzini, A. 2006, ARA&A, 44, 141
Rowlands, K., Wild, V., Bourne, N., et al. 2018, MNRAS, 473, 1168
Sánchez Almeida, J., Terlevich, R., Terlevich, E., Cid Fernandes, R., &
　　Morales-Luis, A. B. 2012, ApJ, 756, 163
Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, MNRAS, 440, 889
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
Simha, V., Weinberg, D. H., Conroy, C., et al. 2014, arXiv:1404.0402
Simpson, J. M., Smail, I., Wang, W.-H., et al. 2017, ApJL, 844, L10
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Skilling, J. 2004, in AIP Conf. Proc. 735, Bayesian Inference and Maximum
　　Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, &
　　U. von Toussaint (Melville, NY: AIP), 395
Socolovsky, M., Almaini, O., Hatch, N. A., et al. 2018, MNRAS, 476, 1242
Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 458, 242
Tacconi, L. J., Genzel, R., Smail, I., et al. 2008, ApJ, 680, 246
Tadaki, K.-i., Kodama, T., Nelson, E. J., et al. 2017, ApJL, 841, L25
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ,
　　621, 673
Toft, S., Smolčić, V., Magnelli, B., et al. 2014, ApJ, 782, 68

Toft, S., Zabl, J., Richard, J., et al. 2017, Natur, 546, 510
Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005, ApJL, 622, L5
Trotta, R. 2008, ConPh, 49, 71
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2005, ApJ, 631, 145
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, MNRAS, 449, 361
Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, ApJL, 811, L12
Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104
Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, ApJ, 745, 179
Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, ApJL, 770,
　　L39
Wild, V., Almaini, O., Cirasuolo, M., et al. 2014, MNRAS, 440, 1880
Wild, V., Almaini, O., Dunlop, J., et al. 2016, MNRAS, 463, 832
Wild, V., Walcher, C. J., Johansson, P. H., et al. 2009, MNRAS, 395, 144
Wisnioski, E., Mendel, J. T., Förster Schreiber, N. M., et al. 2018, ApJ, 855, 97
Wu, P.-F., Gal, R. R., Lemaux, B. C., et al. 2014, ApJ, 792, 16
Wu, P.-F., van der Wel, A., Bezanson, R., et al. 2018a, ApJ, 868, 37
Wu, P.-F., van der Wel, A., Gallazzi, A., et al. 2018b, ApJ, 855, 85
Wuyts, S., Cox, T. J., Hayward, C. C., et al. 2010, ApJ, 722, 1666
Wuyts, S., Labbé, I., Franx, M., et al. 2007, ApJ, 655, 51
Yano, M., Kriek, M., van der Wel, A., & Whitaker, K. E. 2016, ApJL,
　　817, L21
Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327