

THE AVERAGE PHYSICAL PROPERTIES AND STAR FORMATION HISTORIES OF THE UV-BRIGHTEST STAR-FORMING GALAXIES AT $z \sim 3.7$

KYOUNG-SOO LEE^{1,9}, ARJUN DEY², NAVEEN REDDY^{2,10}, MICHAEL J. I. BROWN³, ANTHONY H. GONZALEZ⁴, BUELL T. JANNUZI²,
 MICHAEL C. COOPER^{5,11}, XIAOHUI FAN⁵, FUYAN BIAN⁵, EILAT GLIKMAN^{1,12}, DANIEL STERN⁶,
 MARK BRODWIN^{7,13}, AND ASANTHA COORAY⁸

¹ Yale Center for Astronomy and Astrophysics, Department of Physics, Yale University, New Haven, CT 06520, USA

² National Optical Astronomy Observatory, Tucson, AZ 85726, USA

³ School of Physics, Monash University, Clayton, Victoria 3800, Australia

⁴ Department of Astronomy, University of Florida, Gainesville, FL 32611, USA

⁵ Steward Observatory, University of Arizona, Tucson, AZ 85721, USA

⁶ Jet Propulsion Laboratory, California Institute of Technology, MS 169-327, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

⁷ Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA

⁸ School of Physical Sciences, University of California, Irvine, CA 92697, USA

Received 2010 September 14; accepted 2011 March 23; published 2011 May 11

ABSTRACT

We investigate the average physical properties and star formation histories (SFHs) of the most UV-luminous star-forming galaxies at $z \sim 3.7$. Our results are based on the average spectral energy distributions (SEDs), constructed from stacked optical-to-infrared photometry, of a sample of the 1913 most UV-luminous star-forming galaxies found in 5.3 deg² of the NOAO Deep Wide-Field Survey. We find that the shape of the average SED in the rest optical and infrared is fairly constant with UV luminosity, i.e., more UV-luminous galaxies are, on average, also more luminous at longer wavelengths. In the rest UV, however, the spectral slope β ($\equiv d\log F_\lambda / d\log \lambda$; measured at $0.13 \mu\text{m} < \lambda_{\text{rest}} < 0.28 \mu\text{m}$) rises steeply with the median UV luminosity from -1.8 at $L \approx L^*$ to -1.2 ($L \approx 4\text{--}5L^*$). We use population synthesis analyses to derive their average physical properties and find that (1) L_{UV} and thus star formation rates (SFRs) scale closely with stellar mass such that more UV-luminous galaxies are also more massive, (2) the median ages indicate that the stellar populations are relatively young (200–400 Myr) and show little correlation with UV luminosity, and (3) more UV-luminous galaxies are dustier than their less-luminous counterparts, such that $L \approx 4\text{--}5L^*$ galaxies are extinguished up to $A(1600) = 2$ mag while $L \approx L^*$ galaxies have $A(1600) = 0.7\text{--}1.5$ mag. We argue that the average SFHs of UV-luminous galaxies are better described by models in which SFR increases with time in order to simultaneously reproduce the tight correlation between the UV-derived SFR and stellar mass and their universally young ages. We demonstrate the potential of measurements of the SFR– M_* relation at multiple redshifts to discriminate between simple models of SFHs. Finally, we discuss the fate of these UV-brightest galaxies in the next 1–2 Gyr and their possible connection to the most massive galaxies at $z \sim 2$.

Key words: cosmology: observations – dust, extinction – galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: stellar content

Online-only material: color figure

1. INTRODUCTION

Constraining the main mode of star formation at high redshift is of fundamental importance to the theory of galaxy formation. Significant progress has been made in characterizing some of the general physical properties of the star-forming galaxy population at high redshift. Measurements have been made of the clustering, spectral energy distributions (SEDs), and kinematics of these galaxies (e.g., Shapley et al. 2001, 2003; Papovich et al. 2001; Pentericci et al. 2007; Lee et al. 2006, 2009; Stark et al. 2009; Vanzella et al. 2009). Despite these efforts, major questions remain about how these galaxies form their stars. We do not yet know whether their star formation history (SFH) is dominated by a series of short but intense “bursts” or is a long, continuous accretion of gas converted into stars. Constraints on the SFH of galaxy populations can be relevant to the availability of fuel for star formation and hence

also be relevant to their gas accretion history. For example, bursts may point to merger-induced processes which deliver galaxy-sized parcels of gas (e.g., Mihos & Hernquist 1996; Kolatt et al. 1999; Somerville et al. 2001), whereas continuous star formation might suggest a more stable fuel delivery system, such as “cold-mode accretion” (e.g., Keres et al. 2009; Dekel et al. 2009). Observations at $z = 0\text{--}1$ (e.g., Noeske et al. 2007; Elbaz et al. 2007) and a limited set of observations at redshift out to $z \sim 2$ (Genzel et al. 2006; Daddi et al. 2007), however, show some inconsistencies with theoretical predictions at high redshift (Davé 2008). For example, studies of faint star-forming galaxies at $z > 3$ seem to favor short duty cycles contradictory to theoretical expectations (e.g., Lee et al. 2009; Stark et al. 2009).

It is currently unclear whether such discord requires changes to the theory (e.g., implementing luminosity-dependent galaxy formation) or is the result of observational biases inherent in our incomplete census of galaxies at high redshift. This is because the most recent efforts have been focused on samples derived from the small-area, deep surveys (e.g., GOODS, HUDF; Sawicki & Thompson 2006; Bouwens et al. 2007), which, although excellent at identifying sub- L^* galaxies, lack the cosmic volume to contain significant samples of the most

⁹ Jaylee and Gilbert Mead Fellow.

¹⁰ Hubble Fellow.

¹¹ Spitzer Fellow.

¹² National Science Foundation Fellow.

¹³ W. M. Keck Postdoctoral Fellow.

luminous and massive galaxies. While the UV-faint galaxies likely dominate the cosmic star formation rate (SFR) density (e.g., Madau et al. 1996; Reddy & Steidel 2009), their measured SFRs, space densities, and clustering amplitudes (e.g., Lee et al. 2006; Bouwens et al. 2007; Reddy & Steidel 2009) suggest that they are not likely to be the progenitors of the most massive galaxies seen at later epochs (Quadri et al. 2007; Elsner et al. 2008; Marchesini et al. 2009). In contrast, the UV-brightest galaxies at $z > 3$ may provide a more plausible progenitor population for the most massive galaxies ($M_{\text{star}} \gtrsim$ a few $\times 10^{11} M_{\odot}$), especially if they are able to sustain their observed SFRs for an extended duration ($\gtrsim 1$ Gyr).

In this paper, we define a large sample (1913) of the most UV-luminous galaxies at $z > 3$ and investigate their average properties with the goal of determining their stellar masses, SFHs, and evolutionary descendants.

We use $(\Omega, \Omega_{\Lambda}, \sigma_8, h_{100}) = (0.28, 0.72, 0.9, 0.72)$. Magnitudes are given in the AB system (Oke & Gunn 1983) unless otherwise noted.

2. DATA AND GALAXY SAMPLE

Our sample of 1913 super- L^* star-forming galaxies at $z \approx 3.7$ is selected from the 9.3 deg² Boötes field of the NOAO Deep Wide-Field Survey (NDWFS; Jannuzi & Dey 1999). The NDWFS data set used in this work includes optical B_WRI imaging (Jannuzi & Dey 1999, available from the NOAO Science Archive) taken with the Mosaic camera, near-infrared JHK_S imaging taken with the NEWFIRM camera on the Mayall 4 m Telescope (A. Gonzalez et al. 2011, in preparation), and *Spitzer* IRAC data from the *Spitzer* Deep Wide-Field Survey (SDWFS; 3.6, 4.5, 5.8, 8.0 μm ; Ashby et al. 2009, available from the *Spitzer* Heritage Archive). Where available, we also use U -band (“ U -spec” filter) data taken with the Large Binocular Telescope (LBT; F. Bian et al. 2011, in preparation) to check whether galaxies in our sample are detected. Because the optical data were taken over multiple pointings, the depth and seeing vary significantly over the entire Boötes field. In order to ensure robust photometric measurements and candidate selection, we restrict our study to areas where the seeing in the B_W , R , and I bands has a Moffat profile with an MFWHM of $\leq 1''.3$, with a best-fit Moffat parameter of $\beta < 2.5$. Using these criteria, we exclude five subfields of the NDWFS data (NDWFSJ1426p3421, NDWFSJ1428p3456, NDWFSJ1428p3531, NDWFSJ1431p3346, and NDWFSJ1434p3456). The seeing and photometric depths of the optical data in all subfields are found on the survey Web site.¹⁴ After excluding these subfields and the masked area, the effective area of our sample is ≈ 5.3 deg².

Our sample is selected by applying a Lyman break color-selection technique to the B_WRI data. The Lyman break technique is designed to identify a UV-bright star-forming galaxy with a strong spectral break (at $\lambda \leq 1216 \text{ \AA}$) resulting from absorption by the intervening Ly α forest. At redshifts $3.3 < z < 4.3$, this break lies between the B_W and R bands. Following numerous studies in the literature (e.g., Steidel et al. 1999; Giavalisco et al. 2004; Bouwens et al. 2007), we simulated the expected colors of typical star-forming galaxy SEDs at a range of redshifts and defined the following color criteria:

$$\begin{aligned} (B_W - R) &\geq 1.2 + 3.0 \times (R - I) \\ (B_W - R) &\geq 2.0 \\ (R - I) &\geq -0.3S/N(R) \geq 3S/N(I) \geq 7. \end{aligned} \quad (1)$$

In order to include only objects with robust photometric measurements, we excluded candidates lying within masked regions around bright stars (including diffraction spikes around them), large galaxies, and the edge of each subfield image. We also excluded objects that are heavily blended with adjacent sources and stars based on the SExtractor CLASS_STAR indices (Bertin & Arnouts 1996). Of 108 objects that satisfy the criterion CLASS_STAR > 0.95 , 78 objects (72%) are brighter than $I = 23.0$ where the stellarity index can be measured reliably. Furthermore, we use optical–infrared colors ($R - I$ versus $I - 3.6 \mu\text{m}$) defines a stellar locus well separated from galaxies; see Figure 8(b) in Brown et al. 2007) to further exclude stars whenever possible. We also visually examine the LBT U -band data and exclude any of the candidates which are detected; the U band samples light below the Lyman continuum break ($\lambda = 912 \text{ \AA}$) at redshifts $z \gtrsim 3.4$, and hence any candidate detected in U is unlikely to be at a redshift of interest. Very few ($< 1\%$) of the sources are identified in the U band (with detections defined as 3σ or higher, corresponding to $U \leq 26$ mag). After excluding these individual detections, we stacked the resulting sample and confirmed that the median candidate is also not detected in the (stacked) U -band image (see Figure 3). We note that less than 10% of our candidates are formally detected in IRAC channels 1 and 2 (3.6 μm), and less than 1% in channels 3 and 4. While we do not exclude any object based on the IRAC photometry, this is reassuring because the relatively shallow depth of the *Spitzer* data of the Boötes field is unlikely to warrant detection of most high-redshift sources except for very massive ones (see later). The 5σ depths of the SDWFS data are 22.74, 22.24, 20.39, and 20.25 mag (AB) for IRAC channels 1, 2, 3, and 4, respectively (Ashby et al. 2009).

The final number of candidates is 1913, corresponding to a surface density of ≈ 0.10 arcmin^{−2} over the 5.3 deg² area. As shown in Figure 1 (top right), our sample contains a significant number of very bright sources—for example, 236 sources are brighter than $I_{AB} = 23.5$; these provide ideal targets for detailed studies of galaxy properties such as their stellar populations, interstellar media, and kinematics.

3. SPECTROSCOPIC SAMPLE

In order to characterize the redshift distribution and contamination rates by low-redshift interlopers and Galactic stars, redshifts for a small sample of candidates were obtained on UT 2010 May 15 using the Deep Imaging Multi-Object Spectrograph (DEIMOS; Faber et al. 2003) on the Keck II Telescope at the W. M. Keck Observatory. Three masks, containing a total of 66 candidates, were observed with 2.0–2.5 hr each using the standard DEIMOS setup with the 600 l mm^{−1} grating ($\lambda_{\text{blaze}} = 7500 \text{ \AA}$) centered on $\sim 7600 \text{ \AA}$. Mask positions were chosen to maximize the number of primary targets (secure B_W -band dropout candidates; $S/N \geq 8$) and secondary targets (fainter candidates; $5 \leq S/N < 8$) within the DEIMOS field of view. The spectra were reduced using the DEIMOS pipeline developed by the DEEP2 survey team at UC Berkeley (M. Cooper et al. 2011, in preparation). While the details of all our spectroscopic observations will be presented elsewhere, we show typical one-dimensional spectra of our candidates in Figure 2. In addition to these observations, we compiled existing spectroscopic redshifts from past Keck observations by Spinrad, Soifer, and collaborators, as well as those obtained from the AGN and Galaxy Evolution Survey (C. Kochanek et al. 2011, in preparation).

¹⁴ <http://www.noao.edu/noao/noadeep/>

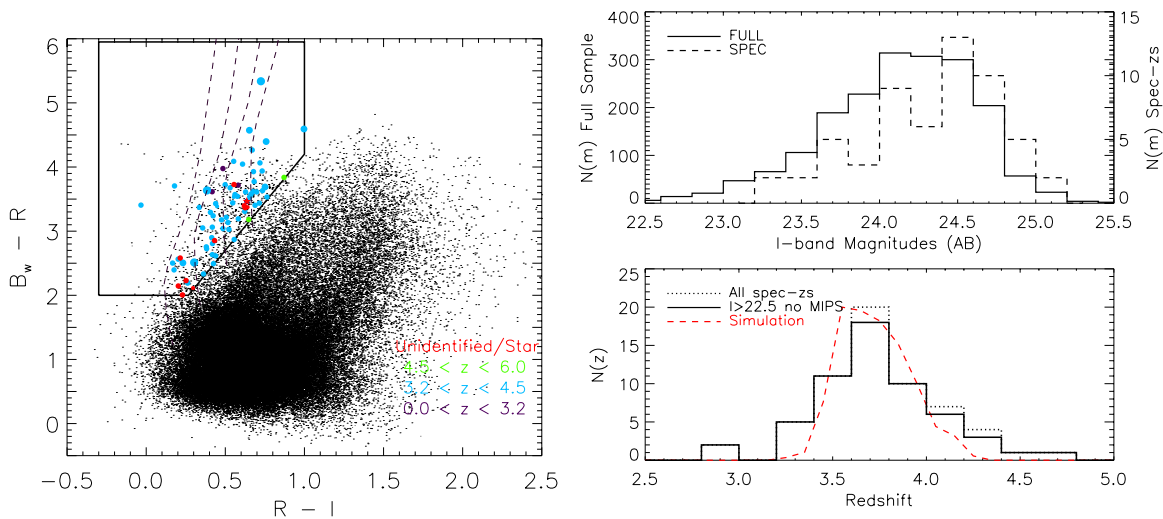


Figure 1. Left: B_WRI color-color plane with the adopted color selection criteria for $z \geq 3.3$ candidates as denoted by black lines. Black dots represent a subset of all NDWFS galaxies. Colored symbols indicate the sources with spectroscopic redshifts, demonstrating the relatively clean selection of high-redshift targets. Dashed lines show the theoretical tracks of star-forming galaxies at $z \geq 3.3$ ($z = 3.3$ at the bottom) for four reddening parameters ($E(B-V) = 0.0, 0.15, 0.30, 0.45$). Right: The top panel shows the magnitude distributions of the full sample (“FULL”: solid line) and the spectroscopic sample (“SPEC”: dashed line). The bottom panel shows the redshift distributions of the entire spectroscopic sample (dotted line) and the subset containing only the fainter galaxies with $I \geq 22.5$ (solid line). We adopt the latter as the representative redshift distribution for the full photometric sample. The expected redshift distribution of the full sample from our simulations is in good agreement with the observed one (dashed line).

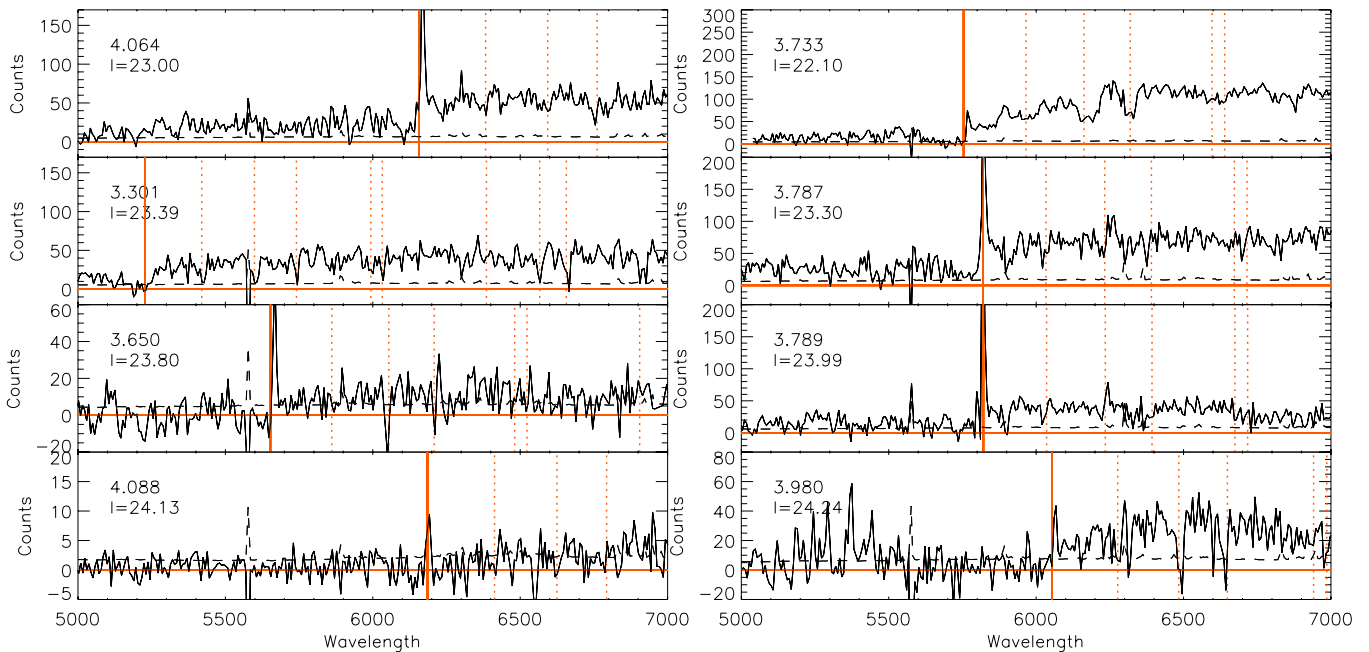


Figure 2. Example DEIMOS spectra of B_W -band dropout candidates. The spectra are ordered in increasing I -band magnitudes from top to bottom. Solid vertical lines mark the location of $\text{Ly}\alpha$, while dotted vertical lines mark, from left to right, $\text{Si II } 1260 \text{ \AA}$, $\text{O I} + \text{Si II } 1302 \text{ \AA}$, $\text{C II } 1335 \text{ \AA}$, $\text{Si IV } 1394, 1403 \text{ \AA}$, $\text{Si II } 1527 \text{ \AA}$, and $\text{C IV } 1548, 1551 \text{ \AA}$ in absorption, respectively.

In total, 78 sources from our sample have been observed spectroscopically. Of these, redshifts could not be determined for 12 sources due to optical faintness or a lack of useful lines. Two of the spectroscopically confirmed sources are *Spitzer* MIPS $24 \mu\text{m}$ sources at $z = 4.230$ and 3.675 (Desai et al. 2008), respectively. The former is a narrow-line active galactic nucleus (AGN) exhibiting C IV and He II, and the latter is a broad-line QSO. Additionally, three sources have been identified as QSOs from our DEIMOS spectra ($z = 3.835, 4.230, 4.050$), and one source is a Galactic star. The total number of confirmed interlopers (i.e., those other than star-forming galaxies at $z \sim 3.7$) is five AGNs and one star. Hence, the rate

of contamination ranges from 8% (if all 12 unidentified sources are at high redshift: $6/78$) to 23% (if all 12 sources are at low redshift: $(6+12)/78$). In Figure 1 (bottom right), we show the two redshift distributions of our spectroscopic samples, one for the full sample including the MIPS sources and the other showing only $I \geq 22.5$ galaxies without MIPS detections. We further note that the MIPS $24 \mu\text{m}$ sources are fairly bright ($I \lesssim 22.5$) compared to the majority of our candidates. Hence, we adopt the latter as the representative redshift distribution $N(z)$ because the majority of our sample is fainter than $I = 22.5$. Thus, the redshift distribution is well described by the mean $\langle z \rangle = 3.7$ and a standard deviation of 0.4.

Table 1
Average Photometry of Star-forming Galaxies at $z \sim 3.7$

		Sample					
	Δm^a	1	2	3	4	5	6
Range	...	$19.46 \leq I^b < 23.06$	$23.06 \leq I < 23.66$	$23.66 \leq I < 23.96$	$23.96 \leq I < 24.26$	$24.26 \leq I < 24.46$	$24.46 \leq I < 24.96$
M_{1700}^c	...	-23.30	-22.51	-22.18	-21.89	-21.63	-21.43
N^d	...	74	254	322	464	307	495
U		>29.84	>30.51	>30.63	>30.83	>30.61	>30.87
B_W	0.09	25.97 ± 0.18	26.93 ± 0.27	27.01 ± 0.28	27.42 ± 0.34	27.86 ± 0.42	28.73 ± 0.63
R	0.09	23.01 ± 0.11	23.68 ± 0.06	23.94 ± 0.06	24.17 ± 0.07	24.35 ± 0.08	24.52 ± 0.08
I^e	0.09	22.59 ± 0.11	23.38 ± 0.04	23.71 ± 0.05	24.00 ± 0.06	24.26 ± 0.07	24.46 ± 0.07
J	0.14	22.21 ± 0.18	23.21 ± 0.08	23.45 ± 0.08	23.93 ± 0.09	24.12 ± 0.17	24.38 ± 0.19
H	0.09	21.92 ± 0.23	22.96 ± 0.11	23.31 ± 0.13	23.63 ± 0.15	23.84 ± 0.29	24.09 ± 0.20
K_s	0.09	21.56 ± 0.20	22.75 ± 0.22	22.97 ± 0.18	23.38 ± 0.25	23.47 ± 0.52	23.80 ± 0.40
$3.6 \mu\text{m}$	0.24	21.70 ± 0.24	22.65 ± 0.09	22.84 ± 0.08	23.22 ± 0.07	23.47 ± 0.11	23.76 ± 0.10
$4.5 \mu\text{m}$	0.25	21.74 ± 0.23	22.60 ± 0.11	22.78 ± 0.09	23.11 ± 0.09	23.48 ± 0.14	23.74 ± 0.13
$5.8 \mu\text{m}$	0.38	21.57 ± 0.29	22.53 ± 0.32	22.68 ± 0.30	23.01 ± 0.36	23.44 ± 0.65	23.61 ± 0.65
$8.0 \mu\text{m}$	0.51	21.67 ± 0.30	22.44 ± 0.26	22.61 ± 0.31	23.31 ± 0.39	23.78 ± 0.72	23.71 ± 0.66

Notes.

^a Aperture correction factor applied to the photometry measured in a $5''$ diameter to obtain the total magnitude listed in the table.

^b The Kron-like total magnitude in the I band (SExtractor MAG_AUTO: Kron 1980; Bertin & Arnouts 1996).

^c Absolute magnitude at $1700 \text{ \AA}$ assuming redshift $z = 3.7$.

^d The number of the photometric candidates used for stacking.

^e I -band photometry measured in a $5''$ diameter with the aperture correction (see Section 3).

To further validate the adopted redshift distribution, we carried out photometric simulations to estimate the expected redshift distribution of our dropout sample based on the broadband color selection. The simulation consisted of (1) creating artificial star-forming galaxies over a range of redshifts, UV luminosities, spectral slopes, morphologies, and sizes, (2) inserting them into the real images, and (3) color selecting in an identical manner in which the real galaxies were selected. We refer interested readers to Ferguson et al. (2004) and Lee et al. (2006) for more details. The expected redshift distribution is shown in Figure 1 (bottom right; dashed line), in good agreement with the observed one.

4. MEASUREMENTS OF THE AVERAGE SPECTRAL ENERGY DISTRIBUTIONS

We investigate the average SED properties of the UV-selected star-forming galaxies by stacking our imaging in the $B_W RIJHK_S [3.6][4.5][5.8][8.0]$ passbands on the positions of the 1913 galaxies in our sample. For a galaxy at $z = 3.7$ —the median redshift of the sample—these bands probe a rest-frame wavelength range of $0.1\text{--}2 \mu\text{m}$. In order to investigate whether their UV brightness (a proxy for SFR) is correlated with other galaxy properties such as stellar mass or extinction, we define six subsamples ordered by their I -band magnitudes. In Table 1, we summarize the details of each subsample. Given the relatively narrow redshift distribution ($z = 3.7 \pm 0.4$), the I -band flux provides a fairly accurate representation of the observed UV luminosity measured at a rest frame of $1700 \text{ \AA}$. At a fixed I -band flux, the UV luminosity of a source at $z = 3.7$ will be higher (lower) than that at $z = 3.3$ (4.1) by ≈ 0.08 dex (assuming that the UV spectrum has the form $f_\lambda \propto \lambda^{-1.45}$).

In order to create average SEDs of galaxies within each luminosity bin, we first constructed stacked two-dimensional images for each band. To create clean stacked images in each band, unbiased by outliers and sky subtraction uncertainties, we adopted the following procedure: (1) We first convolved all the images in a given band to a common point-spread

function (PSF). (2) Then, from these convolved images we extracted $1' \times 1'$ image cutouts centered on each candidate. (3) Next, in order to obtain an unbiased estimate of the sky brightness level, we generated an object mask in order to mask all sources present in each cutout, including the central one (the IRAF objmasks task was used to detect objects above the threshold, then several pixels were added to grow the mask in order to ensure that all the light from sources and their halos was effectively excluded). We then estimated the sky level using the modal value of the unmasked pixels in each frame and subtracted this value from the image (i.e., to create an image with zero sky). (4) Since the image cutouts come from survey data with different zero points, we then rescaled each cutout to a common zero-point magnitude. (5) Finally, we constructed a median stacked image for each band from all the scaled, sky-subtracted cutouts of candidates within a given luminosity bin. In constructing the median, we used object masks to exclude pixels from all objects other than the central one. Using the mean instead of the median does not change our results significantly. Before the final step of stacking, we visually checked all the cutout images. Whenever the central source was significantly blended with a foreground object, we excluded the source from stacking in order to avoid contaminating the stacking signal with interlopers. The effective number of sources used in stacking was reduced by 5%–10%, depending on passbands due to confusion.

This procedure was repeated for each passband (the $UB_W RIJHK_S$ and IRAC bands). Because the blending and crowding vary greatly with passbands, slight adjustments were made to the IRAF objmasks parameters to ensure that all light from the surrounding sources was well masked without masking the majority of the pixels in each cutout. The latter would result in substantial compromise in the effective depth of the stacked image, as most pixels were masked and thus not used in the stacking. In Figure 3, we show the stacked images of all six subsamples in the passbands where flux was detected. The stacked LBT U -band data yielded only an upper limit since no detection was made. The upper limit ranges over $AB = 29.8\text{--}30.9$ for all of our subsamples (see Table 1).

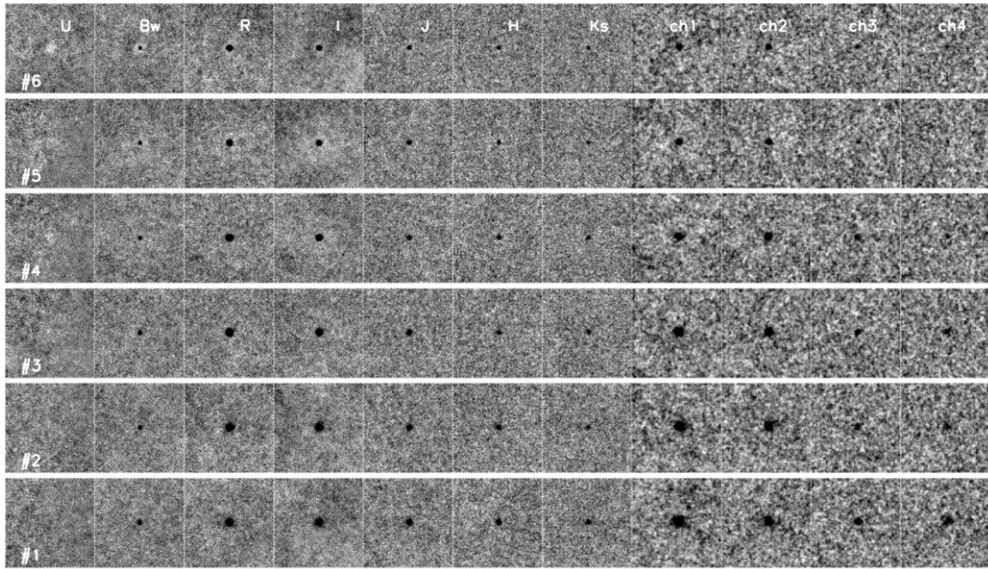


Figure 3. Stacked images in the $UBWRIJHK_S[3.6][4.5][5.8][8.0]$ passbands are shown for all six subsamples (the brightest bin is at the bottom). Each image has a dimension of $1' \times 1'$ centered around the photometric sample. The central source is clearly detected in all passbands out to IRAC $8.0 \mu\text{m}$ with the exception of the U band.

We measured the aperture photometry for each stacked image and thus derived the average SED of each subsample. The photometry was measured in a circular aperture of $5''$ in diameter, then aperture corrected to derive total magnitude. For the optical passbands, which were convolved to a common Moffat profile PSF (MFWHM = $1''.35$, $\beta = 2.5$; see Brown et al. 2007), and for the near-infrared data, which were also convolved to a common PSF (MFWHM = $1''.35$ for H and K_S ; MFWHM = $1''.6$ for J , $\beta = 2.5$; A. Gonzalez et al. 2011, in preparation), the aperture correction factors were estimated analytically. For the *Spitzer* IRAC data, we use the values published in Ashby et al. (2009). The aperture correction factors are given in Table 1 in units of magnitude.

To estimate the impact of the intrinsic spread of the range of galaxy luminosities/magnitudes on the measured median value for the sample, we carried out Monte Carlo simulations. Each time, a random 30% of the sample was chosen and the stacking and photometric measurements were carried out as outlined above. We repeated this procedure 500 times to construct the flux distribution of a given subsample. The intrinsic scatter was estimated from this distribution assuming that the random errors follow a normal distribution (i.e., they can be scaled from the photometric error of the full stack), and was then added in quadrature to the photometric measurement error to estimate the total scatter in the stacked photometry (Table 1).

As shown in Figure 4, the average rest-frame optical SEDs of galaxies in different UV luminosity bins are strikingly similar to one another in their overall shapes. All of them share the flatness in the rest-frame optical portion of the spectrum ($f_\nu \propto \lambda^0$) and exhibit a weak Balmer break, which is located between the H and K_S bands at redshift $z = 3.1\text{--}4.1$. Furthermore, it is evident that the UV-brighter galaxies are also brighter in the rest optical by roughly the same scale factor as in the UV. Despite the overall similarities, however, there are some measurable differences. In the rest-frame UV portion of the SED, the UV slope β ($\equiv d\log F_\lambda / d\log \lambda$; measured in the observed RIJ bands) is steeper with increasing UV luminosity. In the next section, we discuss the physical implications of our results using stellar population model fitting of the averaged SEDs.

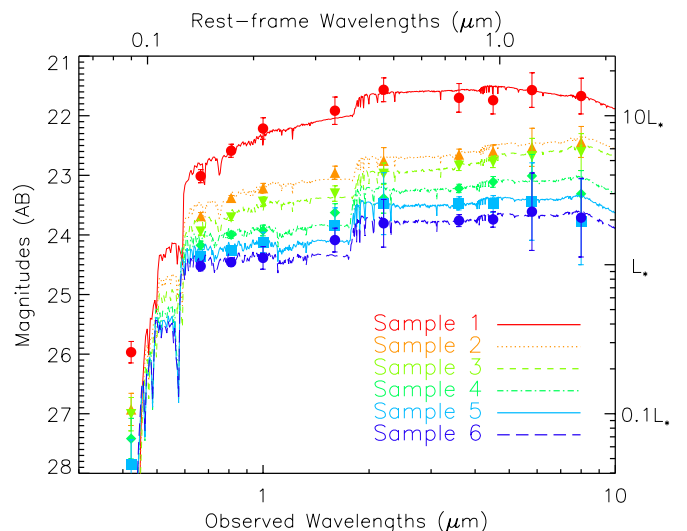


Figure 4. Average SEDs of the six UV subsamples shown with the best-fit population synthesis model, assuming a CSF history. The SEDs were obtained by convolving the best-fit model with the redshift distribution $N(z)$ to account for redshift uncertainties. The B_WRIJHK_S and IRAC photometry probe out to the rest-frame $2 \mu\text{m}$ at $3.3 < z < 4.3$. On the right y-axis, we also show the characteristic UV luminosity L^* at $1700 \text{ \AA}$ $z \sim 3$ (Steidel et al. 1999), corresponding to the I -band magnitudes (at $\lambda_{\text{obs}} \approx 8100 \text{ \AA}$, the third data point from left for each SED). The overall shapes of the SEDs are remarkably similar to one another, suggesting regularity of their formation processes (see the text).

5. AVERAGE PHYSICAL PROPERTIES

5.1. Stellar Population Synthesis Analyses

We determine the average physical properties, specifically the stellar mass, age, and reddening, of our samples by fitting the Bruzual & Charlot (2003) stellar population synthesis models to the stacked photometry. We assume the Chabrier (2003) initial mass function (IMF) and the Calzetti et al. (2000) form for the effective dust extinction law and we consider only solar metallicity populations. We model the SFHs to be constant, exponentially declining ($\text{SFR} \propto e^{-t/\tau}$) with a wide range of τ

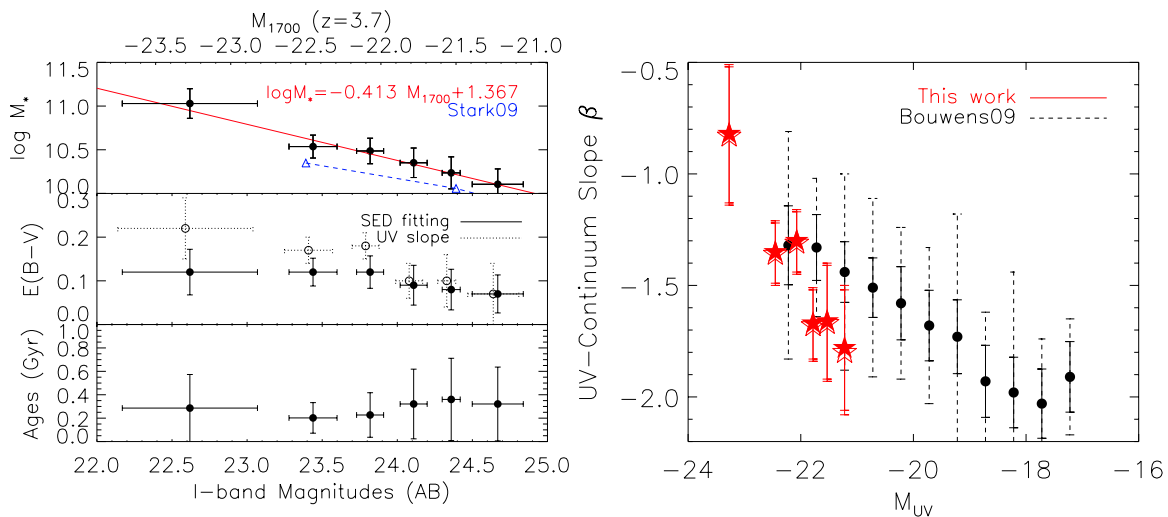


Figure 5. Left: Best-fit parameters for the six UV bins are shown. As shown in the top left panel, the mean stellar masses of galaxies rise steeply with UV luminosity (or I -band magnitudes), implying that galaxies with higher SFRs are also more massive. We also show the formal best fit to our measurements (solid line) and that of Stark et al. (2009, dashed line). The middle left panel shows that dust reddening also increases with UV luminosity. The best-fit extinction values from the SED fitting (filled circles) and those computed directly from the UV slope (open circles; offset slightly for clarity) using the Meurer et al. (1999) correlation are shown. The bottom left panel shows the inferred best-fit ages indicating young ages for all six subsamples and showing little or no correlation with UV luminosity. Right: Median UV slopes β at different UV luminosities found in this work (filled stars) are compared with those of Bouwens et al. (2009, shown in filled circles). For this work, the UV slopes were measured from the stacked (median) RIJ photometry, assuming a median redshift of $z = 3.7$. The RIJ photometry corresponds to the rest-frame wavelength range of 1300–2600 Å at $z = 3.3$ –4.0. As for the Bouwens et al. (2009) points measured from a single UV color, both mean (solid error bars) and intrinsic (dashed error bars) 1σ scatter are shown. The two measurements are in quantitative agreement and show that the average dust extinction of galaxies increases with UV luminosity over a wide range of $M_{UV} = -23$ to -17 .

(A color version of this figure is available in the online journal.)

values (10 Myr–5 Gyr), or linearly increasing. The stellar age is allowed to vary between 50 Myr and the age of the universe at redshift z where the fitting is performed (i.e., ≈ 1.66 Gyr at $z \approx 3.7$). The minimum age of 50 Myr is set based on dynamical timescale arguments when considering the measurements of half-light radii and velocity dispersions of galaxies at high redshift (Bouwens et al. 2004; Erb et al. 2006). While these assumptions for age ranges, IMF, metallicity, and extinction may not be accurate descriptions of the properties of high-redshift galaxies (and may even vary with galaxy mass and/or luminosity), our main interest lies in comparing the relative average properties of galaxies in different UV luminosity bins. Our current study is limited by the nature of our data, which is confined to broadband photometry; more details will emerge only from deep spectroscopic observations, which are now within grasp for the bright candidates within our sample.

Using a Monte Carlo approach, we estimated the uncertainties of the derived physical properties that result from the finite width of the redshift distribution. We randomly drew a redshift from the observed $N(z)$ distribution (right bottom of Figure 1), assigned it to the observed average SED, found the best-fit parameters at each time, and repeated this procedure 100 times. Although we used the observed spectroscopic redshift distribution for these analyses, we also simulated the expected redshift distribution given our adopted UV color selection criteria and a large range of theoretical galaxy SEDs. The two distributions agree with each other reasonably well (Figure 1), and we have verified that using one or the other does not affect our main results.

5.2. Luminosity, Mass, and Age

Given the narrow redshift distribution, the galaxies in our sample range in rest-frame UV luminosity from $M_{1700} \approx -21.43$ to -23.30 (see Table 1). Our sample therefore consists

primarily of $L \gtrsim L^*$ galaxies and greatly increases the number of candidates for the UV-brightest galaxies. Based on our population synthesis analyses, we find that the average galaxies in our sample span roughly a decade in stellar mass, from $\log(M_*/M_\odot) \approx 10$ –11, and the best-fit population synthesis ages of 200–400 Myr in the UV bins considered here.

We summarize our results in Figure 5 assuming a constant SFH model, which facilitates direct comparison with other studies in the literature. However, we note that linearly rising SFH models generally provide an equally good fit to the stacked SEDs. The average stellar masses of our samples are tightly correlated with their UV luminosities, as expected from the homologous SED shapes in the different luminosity bins (Figure 4). Assuming the median redshift $z = 3.7$ to convert the I -band AB magnitude to the observed UV luminosity, we find that the correlation is well described by

$$\log M_*/M_\odot = -0.413 M_{1700} + 1.367$$

(top panel of Figure 5). Since the uncertainties in stellar mass fully account for both random and intrinsic scatter in the photometry within the sample, and the width of the redshift distribution, the intrinsic correlation is likely even tighter than measured here. The UV luminosity range of the NDWFS sample overlaps with the luminous end of the sample (about 70–80 galaxies) probed by Stark et al. (2009) at $M_{UV} \lesssim -21.2$ at similar redshifts. Our estimation of the median stellar masses at these UV bins is slightly higher than theirs, as shown in Figure 5 (dashed line; we adopted the stellar mass when CSF is assumed), but considering the large scatter and small number of galaxies included in these bins, the two measurements are in qualitative agreement with each other. We further compare our results with those of Shapley et al. (2005), who measured the stellar mass and UV luminosity of moderately UV-luminous star-forming galaxies at $z \sim 2.3$. While they found no correlation between

UV luminosity and stellar mass, in contrast to our results, their estimation of the median stellar mass is very consistent with ours, $\log M_*/M_\odot = 10.0 \pm 0.2$, at $M_{\text{UV}} \sim -21.0$, which is our faintest bin.¹⁵ Further comparison is not possible at this time, as only ≈ 20 galaxies in their sample are bright enough to be in ours.

Unlike stellar mass, we find no correlation between the median age of the stellar population and UV luminosity. The best-fit age values (200–400 Myr in all cases) suggest that the majority of galaxies are dominated by young star-forming populations—not just in their rest-frame UV light, but also in their rest-frame optical-to-near-infrared light—that are considerably younger than the age of the universe at $z \sim 3.7$ (≈ 1.7 Gyr; e.g., Papovich et al. 2001). Assuming exponentially declining SFHs would generally make best-fit ages even younger. The young ages are inferred from the weak Balmer breaks between the H and K_S bands, and are thus robust against different SFHs assumed for the SED fitting. Finally, our results show that there is a positive correlation between reddening and observed UV luminosity, reflecting the increasingly steeper UV slope with luminosity observed in the stacked SEDs (Figure 4). This extends to higher luminosity and redshift trends observed in other Lyman break galaxy populations (e.g., Papovich et al. 2001; Bouwens et al. 2009).

5.3. Dust Content

We confirm the previous observations that more UV-luminous star-forming galaxies at high redshift are dustier (see, e.g., Reddy et al. 2010; Shapley et al. 2001; Bouwens et al. 2010 at $z \sim 2, 3$, and $\gtrsim 4$, respectively) as shown in Figure 5 (both panels). In addition to the reddening parameters derived from population synthesis analyses, we can also directly measure the UV slope $\beta \equiv d\log F_\lambda / d\log \lambda$ from the stacked SEDs (RIJ bands: rest 1300–2600 Å), after correcting for the intergalactic absorption (Madau 1995) that affects the short-wavelength bands. From our UV-brightest bin to the least luminous bin, the slope changes from $\beta = -1.16 \pm 0.31$ (sample 1) to -1.37 ± 0.14 (sample 3) to -1.78 ± 0.28 (sample 6). Shifting the median redshift by ± 0.2 does not significantly alter our results ($\lesssim 10\%$ in the estimated UV slope β). Using the Calzetti extinction law and the correlation between the UV slope and reddening (Meurer et al. 1999; Calzetti et al. 2000), these values correspond to the reddening $E(B - V) = 0.22 \pm 0.07, 0.17 \pm 0.03, 0.07 \pm 0.07$.¹⁶ The difference between the extinction found by the SED population synthesis fitting and direct measurements of the UV slope is likely due to the fact that the stellar ages were allowed to vary freely for the former but not for the latter.¹⁵

Our measurements are largely complementary to the UV slope determination made by Bouwens et al. (2009). While they also measured the UV slope of UV-luminous ($L \gtrsim L^*$) galaxies, the majority of the galaxies in their luminous samples come from the GOODS data (0.1 deg^2 in total), and thus the total number of

galaxies included in these bins is much smaller than ours (by a factor of $\gtrsim 10$). In Figure 5 (right) we show these results together. The two measurements are in quantitative agreement with each other, although there are slight discrepancies at $L \approx L^*$. This may be in part due to the longer wavelength baseline used for our measurements (RIJ ; out to $1 \mu\text{m}$) compared with theirs ($i'z$) where the UV slope β steepens (see, e.g., Figure 4 near the rest-frame $1 \mu\text{m}$). Furthermore, inclusion of the R band in the determination of the UV continuum slope may result in, at least for some galaxies, contamination of the flux from $\text{Ly}\alpha$ emission, and, as a result, may cause the UV continuum to appear bluer than it is. Indeed, some ($\gtrsim 30\%$) of our spectroscopic samples have been observed with $\text{Ly}\alpha$ emission. However, a much larger spectroscopic sample is needed to make statistical estimates of the fraction of galaxies with $\text{Ly}\alpha$ emission as well as the distribution of the equivalent widths, and thus how this will affect the UV slope of each stack. Finally, we note that the selection bias of the photometric candidates also contributes to measurement differences; for a fixed I -band (detection band) flux, galaxies with bluer $R - I$ colors generally have a higher completeness, and thus the median UV slope of the sample will be bluer (steeper) than the intrinsic one (see Figure 3 of Bouwens et al. 2009, for example). However, correcting for this effect statistically is not straightforward in our case, as we use the stacked photometry for the J band to determine the UV slope. Without knowing the intrinsic distribution of $I - J$ colors (which is the quantity we want to measure), it is statistically difficult to correct the UV slope measured from the RIJ photometry based on the $R - I$ colors alone. However, this explanation is consistent with the fact that the most discrepant measurements are where our selection completeness is the lowest ($\approx L^*$).

When these effects are properly accounted for, the two measurements of the UV continuum slope are in good agreement. The overall trend is clear; the average UV continuum slope increases with UV luminosity over a wide range of luminosities ($0.02L^* - 6L^*$).

6. DISCUSSION

We have identified the largest photometric sample to date of UV-bright star-forming galaxies at $z \sim 3.7 \pm 0.4$ by applying a standard color-selection technique to the broadband data from the NDWFS. Our final sample consists of 1913 galaxies in the luminosity range of $L^* - 5L^*$ distributed over an area of 5.3 deg^2 . We have divided the sample into six bins of UV luminosity (i.e., observed I -band magnitude) and constructed the average SEDs in each bin. As described above, we confirm a strong trend between UV luminosity and stellar mass. In this section, we discuss the implications of these results for the nature of the high-redshift star-forming galaxy population.

6.1. The Mass Dependence of the Star Formation Rate

The fact that the more UV-luminous galaxies are considerably more massive than less UV-luminous galaxies implies a strong correlation between the instantaneous SFR (derived from the rest-frame UV luminosity) and stellar mass (derived from the population synthesis fits and primarily mid-infrared photometry). Given the young mean stellar ages (200–400 Myr) inferred for the (luminosity-weighted) stellar populations in these galaxies (Section 4), the tight correlation between the current forming rates of stars and existing stellar content also implies that the formation of the bulk of their stellar content is likely associated with the current episode of star formation rather than any

¹⁵ This is excluding five galaxies in their sample with very red UV-to-optical color $\mathcal{R} - K_S > 3.5$. Similar sources at $z \sim 3.7$ would be bright enough to be detected ($>5\sigma$) in the IRAC ($3.6 \mu\text{m}$) band. Such sources are unlikely to be common in the galaxies in our sample, as less than 5%–10% of our candidates are formally detected in the SDWFS data.

¹⁶ $A(1600 \text{ Å}) = 2.31(\beta - \beta_0)$ (Equation (9) in Calzetti et al. 2000) is used for this conversion where they adopted $\beta_0 = -2.1$ for the intrinsic UV spectral slope, assuming a CSF history for 1 Gyr (Leitherer et al. 1999). The intrinsic slope does not change significantly for the same SFH for 300 Myr, the median age found for our sample.

previous ones. Furthermore, we infer that the average SFHs of super- L^* galaxies are smoothly varying and long (i.e., at least comparable to the width of the redshift distribution $N(z)$, or $\gtrsim 400$ Myr) rather than peaky/bursty, as the latter would lead to little or no correlation between UV luminosity (or SFR) and stellar mass.

The long, continuous star formation implied by our results is apparently at odds with the relatively young ages suggested by the weak Balmer break observed in all UV bins. Within the conventional “ τ model,” similarly young ages and a tight SFR- M_* relation can be achieved only if the formation of galaxies began at the same point in time. Such a scenario, however, in which the majority of galaxies (of all luminosities) have such a synchronized formation redshift (at $z_f \approx 4.2\text{--}4.5$) and SFH is unphysical! To maintain a more physical picture, we argue that the average SFH of the UV-bright Lyman break galaxy population needs to be one in which the SFR increases with time. Even models in which the SFR increases mildly with time would keep the age of the stellar population relatively young, since the light from the galaxy is always dominated by recently formed stars, while maintaining a tight SFR- M_* correlation. We note that similar arguments have been made in the literature to explain the same dilemma found for $z \sim 2$ star-forming galaxies (Maraston et al. 2010), to explain the observed change in UV luminosity and stellar mass growth with redshift (Papovich et al. 2011) and to model reionization (Finlator et al. 2011).

6.2. Implications for Star Formation Histories

The observed relation between the SFR and stellar mass (Ψ - M_* scaling law, hereafter) holds a direct constraint on the rate at which the SFR increases with time. While a more robust determination needs to be based on photometric measurements of individual galaxies with spectroscopic redshifts (to accurately measure both the mean and the scatter of the relation), we can make a rough estimate based on our stacked photometric results. For simplicity, we assume that all galaxies in our sample lie at $z = 3.7$ and adopt the Kennicutt (1998) calibration¹⁷ to convert the median I -band magnitude to UV SFRs for each bin, after correcting for extinction (derived from the best-fit SED model of each UV bin). The logarithmic slope of the Ψ - M_* relation ranges from 0.8–1.1 and depends sensitively on the adopted extinction parameter. When the slope is fixed to unity, the best fit yields the relation $\Psi(z \sim 3.7) = (240 \pm 30 M_\odot \text{ yr}^{-1}) [M_*/10^{11} M_\odot]$ implying that these galaxies double their stellar masses every 290 ± 40 Myr.¹⁸ These results are very similar to that found for star-forming galaxies at $z \sim 2$ (Daddi et al. 2007) where the Ψ - M_* relation is well described by $\Psi(z \sim 2) = (200 M_\odot \text{ yr}^{-1}) (M_*/10^{11} M_\odot)^{0.9}$, and are somewhat higher than those found for IRAC-selected star-forming galaxies at $z \sim 3$ (Magdis et al. 2010) for which the scaling relation

of $\Psi(z \sim 3) = (330 M_\odot \text{ yr}^{-1}) (M_*/10^{11} M_\odot)^{0.91}$ is found. Our results suggest that the specific SFR (sSFR), on average, remains roughly constant, or changes only mildly at best, with stellar mass (see also Daddi et al. 2007; Dunne et al. 2009; Pannella et al. 2009; Magdis et al. 2010; Kajisawa et al. 2010; Reddy et al. 2006, 2010).

While our stacking results provide useful insight into how the sSFR changes with stellar mass, especially in a luminosity regime where no other measurements are currently available, we caution readers that the current data at hand are not adequate to constrain the full distribution of sSFR at given stellar mass bins (see, e.g., Reddy et al. 2006, 2010, based on individual measurements) since one needs to properly factor into the error budget the full redshift distribution and K -correction. Furthermore, to infer the scaling law between sSFR and stellar mass, we made an additional assumption that all galaxies participate in the tight Ψ - M_* relation with no significant outliers. In reality, it is still possible that there may be a small fraction of very massive galaxies (say, $M_* = 10^{11} M_\odot$) that are not very UV luminous (say, $I = 24.5$) as the star formation starts to fade. While such a population would not bias our primary measurements (i.e., the median stellar mass in a given UV bin) as long as their fraction is low, it could lower the median sSFR at a fixed stellar mass.

6.2.1. Simple Prescriptions for the Star Formation History

In order to develop some intuition regarding the potential of the Ψ - M_* relation to constrain the underlying SFH of a galaxy population, we investigated a few simple toy models in which the SFR is independent of mass but solely a function of time, and which start forming stars at a common redshift. While such models naturally result in a constant sSFR at any subsequent redshift of observation, different models will yield different normalizations at different redshifts, which can potentially be used to discriminate between models. In particular, we can compare the predictions of the toy models to the observed Ψ - M_* relationships at $z \approx 3.7$ (this work) and at $z \approx 2$ (from the work of Daddi et al. 2004, 2007; Reddy et al. 2006).

We begin by considering four toy models which share the common characteristic that SFRs are smoothly rising (or constant) with cosmic time.

1. Exponentially Rising (EXP): $\Psi(t) = \Psi_0 \exp[(t - t_0)/\tau]$.
2. Power-law Rising (POW): $\Psi(t) = \Psi_0(t/t_0)^\alpha$.
3. Constant Star Formation (CSF): $\Psi(t) = \Psi_0$.
4. Delayed Star Formation (DLY): $\Psi(t) = \Psi_0 \exp[-(t - t_0)/\tau](t/t_0)$.

The simplicity of these model parameterizations makes them an attractive starting point before considering more complex models (see also Renzini 2009; Noeske et al. 2007). All models have only one free parameter (or, in the case of the CSF model, none), which facilitates direct comparison between different models. In what follows, we give a brief description of our toy models.

Each galaxy is assumed to have formed at $z = z_f$ and followed a given SFH with cosmic time. All models are normalized at $t = t_0$ to have a star formation rate $\Psi(t_0) = \Psi_0$ where time $t(z)$ is defined as the time elapsed since the formation redshift z_f . The location of the same galaxy on the Ψ - M_* plane with cosmic time can be followed at each time step or z . Similar trajectories can be computed for other galaxies of the same SFH but with different normalizations Ψ_0 . The Ψ - M_* scaling relation at a fixed redshift is then obtained by locating all galaxies of

¹⁷ In light of our discussion (Section 6.1) on rising SFH, it is worth mentioning the uncertainty of this commonly used conversion on different SFHs: $\text{SFR}(M_\odot \text{ yr}^{-1}) = 1.4 \times 10^{-28} L_\nu(\text{ergs}^{-1} \text{ Hz}^{-1})$. For a fixed IMF, the conversion factor is lower by 10%–15% (7%) at a population age of 100 (300) Myr for SFHs that rise linearly with time. As for exponentially declining τ models, the factor is higher by less than 5% for $\tau \geq 300$ Myr at population ages older than 100 Myr. Considering the mild dependence of the conversion factor on population age and different SFH models, our estimation of the Ψ - M_* scaling law should be insensitive to the range of stellar population parameters of galaxies within a given UV bin.

¹⁸ The SFR is the rate of increase in stellar mass, or $\Psi \equiv dM_*/dt$ (modulo mass loss). Hence, a linear relation of $M_* = \tau\Psi$ implies $M_*(t) \propto e^{t/\tau}$ in order for galaxies to maintain the scaling relation with cosmic time. The mass-doubling time is given as $\tau \ln 2$.

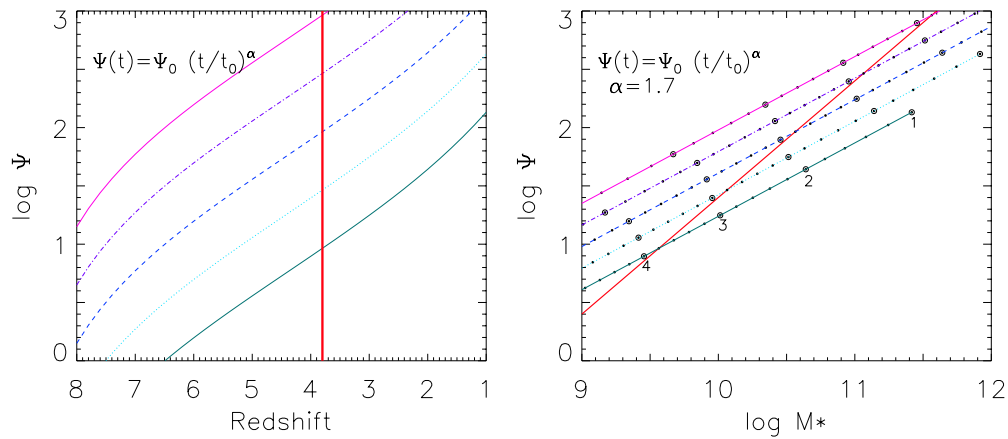


Figure 6. Left: SFHs of a power-law model at five different amplitudes assuming formation redshift $z_f = 9$ and $\alpha = 1.7$. The stellar mass at redshift z , $M_*(z)$, to the zeroth order is the time integral of the SFR Ψ from the formation redshift z_f to z . The vertical line indicates roughly where the median redshift of the B_W -band dropouts lies. Right: Ψ - M_* scaling relation for the same SFHs. Each of the five parallel lines indicates the time trajectory of a galaxy following the SFH shown in the left panel. As the galaxy assembles mass, each trajectory travels from left to right in the figure with cosmic time. For the model with the lowest amplitude Ψ , the redshifts of various cosmic epochs are marked. The Ψ - M_* scaling relation at a given cosmic epoch can be obtained by connecting different trajectories (shown for $z = 3.8$ as a steeper solid line).

different normalizations at the same time step $t(z)$. In Figure 6, we illustrate our calculation for the POW model. The Ψ - M_* scaling law obtained with any smoothly rising SFH is predicted to have a logarithmic slope of unity, given our assumption of synchronized formation redshifts for all galaxies. Reproducing a shallower or steeper slope requires either the formation redshift z_f or star formation timescales (given as α or τ in our toy models) to have a mass dependence similar to that discussed by Noeske et al. (2007). For example, requiring galaxies with a higher normalization Ψ_0 to have higher formation redshift would yield a shallower slope, which may be in qualitative agreement with the hierarchical picture of galaxy formation. We defer such modeling to future work as the current measurements (derived from stacked broadband data) lack the constraining power for such precision.

For the current model, we fix the formation redshift to $z_f = 8$ and the redshifts at which galaxies are “observed” to $z = 3.7$ and $z = 2$ for comparison with the measured relationships. The choice of formation redshift is uncertain, but it is unlikely to be critical to the current discussion. The age estimates from the population synthesis fitting suggest a minimum formation redshift of $z \geq 4.7$, and studies of the global SFR show evidence of galaxy formation at least to $z \gtrsim 7$ (e.g., Papovich et al. 2011). Because the elapsed time is very small at high redshifts, and since we are mainly concerned with the observed SFRs at $z \sim 2$ – 3.7 , choosing $z_f = 10$ instead of $z_f = 8$ makes little difference to the results. The inclusion of a range of formation redshift and more realistic redshift distribution $N(z)$ (e.g., such as that shown in Figure 1) will indeed introduce scatter into the observed Ψ - M_* scaling relation. While we postpone more detailed analyses to future work, we note that the amplitude of such scatter is likely moderate ($\lesssim 0.1$ dex). Given the narrow observed redshift distribution (at least for the $z \approx 3.7$ sample) and the fact that we are investigating Ψ - M_* relations derived from stacked data, it is unlikely that these issues will significantly affect our results.

6.2.2. Comparison to Observations

Following this simple procedure, we investigate how each model fares against the observational constraints. First, we compute the Ψ - M_* scaling relation at $z \sim 3.7$ and determine

the best-fit parameter (given as α or τ) by matching the model predictions to our measurements (see Figure 7(a) for the time evolution predicted by each model). Then, using the best-fit parameter, we compute the scaling relation at redshift 2 to 6 in order to follow the redshift evolution of each model. As the observational constraints (of individual measurements) at other cosmic epochs are limited at the present time, we mainly compare our models to the measurements at $z \sim 3.7$ (this work), BzK galaxies at $z \sim 2$ (Daddi et al. 2007; Pannella et al. 2009), and BX/BM star-forming galaxies at $z \sim 2$ (Reddy et al. 2006). We note that the scaling relation from individual measurements by Daddi et al. (2007) is similar to that measured from the stacked radio data (Pannella et al. 2009) of the same BzK sample. In Figure 7(b) we show the redshift evolution of all four models and best-fit model parameters together with the observations at $z \sim 3.7$ and $z \sim 2$ (shaded regions). A 1σ scatter of 0.25 dex is added for the B_W dropouts to account for the redshift distribution and measurement errors, while for BzK s we use the published value of 0.23 dex (Daddi et al. 2007).¹⁹ As for BX/BM galaxies, the scatter is observed to be ≈ 0.46 dex, about 70% larger than that for BzK s.

It is evident from Figure 7(b) that the CSF and DLY models predict ≈ 0.5 and 0.7 dex decline in the amplitude of the Ψ - M_* scaling relation, clearly too steep to be supported by the observations at $z \sim 2$ and 3.7 . In particular, the CSF model requires that the formation redshift be $z_f \approx 4.8$ (≈ 500 Myr prior to the observations) to reproduce the observed relation at $z \sim 3.7$.²⁰ The implication of the relatively low formation redshift ($z_f \approx 4.8$ for the B_W -band dropouts and $z_f \approx 2.3$ for $z \sim 2$ galaxies) is that the galaxy samples observed at different redshift bins cannot be connected to one another through cosmic time. Such a scenario would not only contradict long duty cycles suggested by the tight Ψ - M_* scaling relation, but also require an unrealistic picture in which the majority of the galaxies observed

¹⁹ The quoted value of the semi-interquartile range of 0.16 dex (Daddi et al. 2007) corresponds to 0.23 dex in the standard deviation assuming a normal distribution.

²⁰ Note that exponentially declining τ models will generally behave in a similar way to CSF models. However, for a given $[\Psi, M_*]$ value, they require the formation redshift to be even closer to the time of the observations than CSF models because their SFR was higher in the past.

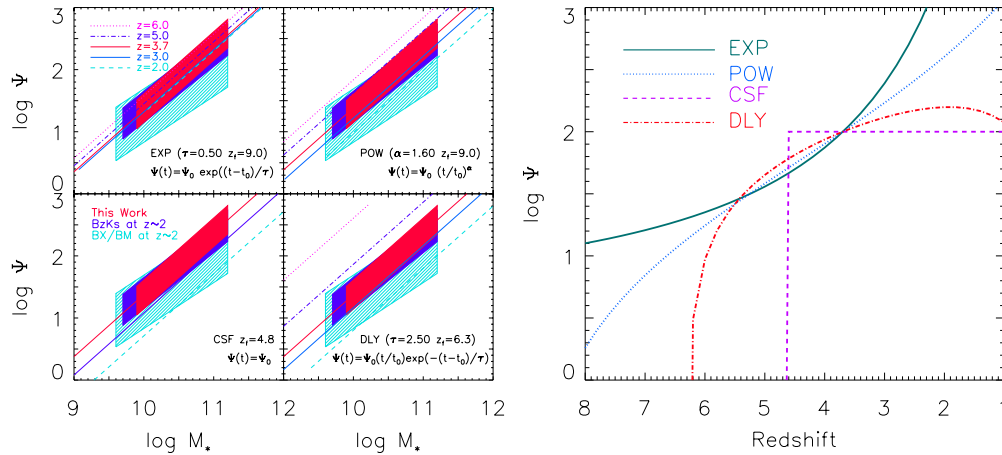


Figure 7. Left: the redshift evolution of the Ψ - M_* relation is explored for four SFH models from $z = 6$ to 2 in comparison with the observations at $z \sim 3.7$ (this work) and $z \sim 2$ (Daddi et al. 2007; Reddy et al. 2006). The parameters α or τ and the formation redshift z_f are chosen for each model to best reproduce the observed relation at $z \sim 3.7$. The EXP and POW models with suitable choices for the model parameter may provide a reasonable description of the existing data, although the CSF and DLY models do not since they predict a steep decline in the scaling relation with cosmic time (see the text). Right: four SFH models that best fit the observed Ψ - M_* scaling relation at $z \sim 3.7$. All models are normalized to have $\Psi = 100 M_\odot \text{yr}^{-1}$ at $z = 3.7$. The time evolution of the SFH (before and after the observation times) and, as a result, the Ψ - M_* scaling law varies dramatically for different models. Hence, more robust determinations of the observed Ψ - M_* scaling relations at different redshifts has the potential to elucidate the SFHs of high-redshift galaxy populations and their descendants at later epochs.

at $z \sim 3.7$ would suddenly start forming stars at a very short time interval around $z = 4.8$.

The EXP model provides a good description of the observations at both redshifts when $\tau = 0.50$ Gyr is assumed (consistent with the age constraints). A slight change in the formation redshift ($z_f = 7$ or 11 instead of 8) does not change the timescale more than 50 Myr. The POW model also provides a good description at $z \sim 3.7$ when $\alpha = 1.60$ (1.70) and $z_f = 9$ (11) is assumed. It is interesting that the exponent α (found simply by requiring it to reproduce the scaling relation at $z \sim 3.7$) is very similar to that found by Papovich et al. (2011), who obtained the exponent α by fitting the redshift evolution of galaxy samples at a fixed number density at $z = 3$ –8 (see their Figure 2). Unlike the EXP model, the POW model predicts a moderate evolution in the amplitude of the scaling relation (≈ 0.3 dex from $z = 3.7$ to 2). While the predicted decline is formally not consistent with the BzK measurements, the real uncertainties (random and systematic) are likely under-represented in Figure 7(b) (shaded area). We note that Reddy et al. (2006) reported much larger scatter (≈ 0.4 – 0.5 dex) for BX/BM star-forming galaxies at a similar redshift range. Furthermore, there are other studies that have found similarly large or even larger scatter at $z \sim 3$ and a higher normalization than the B_W dropouts in the scaling law (Mannucci et al. 2009; Magdis et al. 2010). Such discrepancies in the literature highlight the existing uncertainties and difficulties in measuring such quantities at high redshift.

In summary, our analyses suggest that both the EXP and the POW models do better than the constant SFH model. A more definitive answer will have to await better data on samples at multiple redshifts. The existing data clearly suffer from different selection biases, random (and likely significant) photometric errors in the rest-optical photometry of high-redshift galaxies, and uncertainties associated with SED modeling assumptions. Nevertheless, it is worth noting that comparing the observed Ψ - M_* relations at different redshifts to simple models, such as those discussed here, has the potential to elucidate the SFHs of high-redshift galaxy populations. The key to discriminating between, say, SFHs that rise exponentially or as a power law lies in unambiguously determining the overall amplitude of the Ψ - M_* scaling relation at multiple redshifts. While interesting,

the prescriptions for SFHs described above are very simplistic, and one can easily imagine that SFHs of galaxies can be mass-dependent. However, considering more detailed models is beyond the scope of this present work and unwarranted by the data. A future determination of the scatter in the relationship (through measurements of individual galaxies, as opposed to stacked samples) will bear directly on the distribution of SFH parameters (e.g., formation redshifts, SFR amplitudes, or SFH modes) for the galaxy population.

6.3. What is the Duty Cycle of Star Formation?

The UV luminosity of a star-forming galaxy is dominated by hot stars and therefore measures the instantaneous SFR of the system. However, the existence of the tight Ψ - M_* relation observed within the $z \approx 3.7 \pm 0.4$ star-forming population suggests that the duty cycle of star formation must be significantly longer than the time range sampled (i.e., $\gtrsim 400$ Myr). Furthermore, the positive correlation between the two quantities— $dM_*/dt \propto M_*$ —yields the solution of $M_* \propto t^\alpha$ (or $\propto \exp[t/\tau]$) where α (or τ) is positive, unless the galaxies are observed in a special time of their formation histories. If the onset of star formation in these galaxies is not coordinated, the direct implication would be that the SFR is increasing with time. The increase in the SFR would be better determined by observing similar relations at different cosmic epochs as discussed in Section 6.2.2. In such a scenario, most galaxies are expected to have a fairly long duty cycle until the process is terminated or slowed down by a feedback mechanism.

A complementary constraint on the duty cycle of star formation can be obtained from clustering measurements: the measured clustering strength (combined with a theoretical model) can provide an estimate of the dark halo mass hosting the typical sample galaxy, and a comparison between the expected number density of these halos and the observed number density of the sample galaxies yields an estimate of the duty cycle (e.g., Martini & Weinberg 2001; Lee et al. 2006, 2009; McLure et al. 2009). Indeed, the long duty cycle suggested by our results is supported by the clustering studies of luminous high-redshift galaxies (e.g., Adelberger et al. 2005; Ouchi et al. 2005) and the

interpretations in relation to the theoretical predictions of dark matter halo clustering at the same epoch (Conroy et al. 2006).

In contrast, for less luminous star-forming galaxies at similar redshifts, the evidence seems to point to the contrary. Lee et al. (2009) found that small-scale clustering of faint UV-selected galaxies at $z \sim 4$ and 5 (the deepest data with which the clustering of high-redshift galaxies has been measured thus far) is much weaker than that for dark matter halos, and argued that it may be a result of short duty cycles ($\approx 10\%$ – 30% , or ≤ 0.4 Gyr at the 1σ level) governing their star formation. Furthermore, Stark et al. (2009), who studied the SED properties of similar samples of UV-faint star-forming galaxies, found little evolution in ages for galaxies at $z \sim 4, 5$, and 6. They argued that their star formation may not be continuous, at least at the constant level, as it would lead to older ages, on average, for galaxies at $z \sim 4$ compared to $z \sim 6$. Alternatively, they argued, their young ages could also be produced by rising SFHs in the same way as in more UV-luminous galaxies. However, such a scenario would be inconsistent with the Lee et al. (2009) clustering study, as the small-scale clustering of the galaxies would closely reflect that of dark halos.

Taken at face value, this apparent difference between the implied duty cycles for star formation in UV-luminous and UV-faint galaxies suggests that star formation is a luminosity-dependent process such that luminous galaxies rapidly grow in mass by continuously forming stars (at an increasing rate), while faint galaxies flicker with cosmic time as their UV light turns on and off at short intervals (≈ 200 – 300 Myr). However, it is also true that our understanding of UV-faint high-redshift galaxies is more limited than that of their more luminous counterparts because it is more challenging to obtain spectroscopic observations to identify even their redshifts, not to mention more detailed physical properties. Further investigations are needed to determine whether such differences are due to an observational limitation or an intrinsic one.

6.4. The Descendants of the Most UV-luminous Galaxies at $z \approx 3.7$

The extinction-corrected SFRs of $L \gtrsim L^*$ galaxies span 50 – $250 M_\odot \text{ yr}^{-1}$ at the epoch of observation (i.e., $z \approx 3.7$). The observed Ψ – M_* relation suggests that the SFRs will increase with time (as $\propto e^{t/\tau}$ or $\propto t^{1.6-1.7}$; see Figure 7(a)). Assuming an exponential rise, we estimate that these galaxies will double their stellar mass every ≈ 300 Myr until (or unless) the star formation is quenched or slowed by a feedback mechanism. In the absence of quenching, by $z = 2$ (i.e., ≈ 1.5 Gyr after $z = 3.7$), $L \approx L^*$ galaxies with $\langle M_* \rangle \approx 10^{10} M_\odot$ will grow to $M_* \approx 10^{11.4} M_\odot$, while $2.5L^*$ galaxies will be in excess of $10^{12} M_\odot$. Hence, the descendants of these UV-luminous star-forming systems at $z \approx 3.7$ will likely be galaxies at the top end of the mass function by $z \approx 2$.

It is not clear if such growth in the SFR is sustainable either by the fuel supply within a galaxy or by the physical mechanisms of star formation. More detailed studies of individual star-forming galaxies can shed light on the gas content and fueling of star formation and the role of feedback processes (e.g., AGNs). In addition, comparative statistical studies of the UV-luminosity function, mass function, and large-scale clustering of galaxies at different redshifts (e.g., $z \approx 3.7$ and 2) can constrain the duty cycles of star formation and thereby determine what fraction of the UV-brightest galaxies at high redshift will have evolved into the most massive galaxies at lower redshift (K. S. Lee et al. 2011, in preparation). We intend to investigate this UV-bright sample

using both detailed and statistical approaches, with the goal of elucidating their evolution and determining how and when the feedback processes (such as AGN feedback) will shape the fate of these rare galaxies in the future (e.g., Marchesini et al. 2010).

We thank the referee for her/his careful and thorough report, which helped improve our paper. K.S.L. gratefully acknowledges the generous support of Gilbert and Jaylee Mead for their namesake fellowship in the Yale Center for Astronomy and Astrophysics. K.S.L. also thanks Adam Muzzin, C. Meg Urry, and Casey Papovich for useful discussions and suggestions. A.H.G. acknowledges support for this work from the National Science Foundation under grant AST-00708490. The analysis in this paper is made partly with the Keck observations awarded through Yale University. This work is based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory at the California Institute of Technology. We are grateful for the expert assistance of the staff of Kitt Peak National Observatory where the optical and near-infrared observations of the NDWFS Boötes field were obtained. The authors thank NOAO for supporting the NDWFS. The research activities of A.D. and B.T.J. are supported by NOAO, which is operated by the Association of Universities for Research in Astronomy under a cooperative agreement with the National Science Foundation. The LBT is an international collaboration among institutions in the United States, Italy, and Germany. LBT Corporation partners are the University of Arizona on behalf of the Arizona university system; Istituto Nazionale di Astrofisica, Italy; LBT Beteiligungsgesellschaft, Germany, representing the Max Planck Society, the Astrophysical Institute Potsdam, and Heidelberg University; The Ohio State University; and The Research Corporation, on behalf of the University of Notre Dame, University of Minnesota, and University of Virginia. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership between the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. We thank the staff of the W. M. Keck Observatory, particularly Gregory Wirth. The analysis pipeline used to reduce the DEIMOS data was developed at UC Berkeley with support from NSF grant AST-0071048. The authors also recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community.

REFERENCES

- Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N. A., & Erb, D. K. 2005, *ApJ*, **619**, 697
- Ashby, M. L. N., et al. 2009, *ApJ*, **701**, 428
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J., & Franx, M. 2004, *ApJ*, **611**, L1
- Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, *ApJ*, **670**, 928
- Bouwens, R. J., et al. 2009, *ApJ*, **705**, 936
- Bouwens, R. J., et al. 2010, *ApJ*, **708**, L69
- Brown, M. J. I., Dey, A., Jannuzi, B. T., Brand, K., Benson, A. J., Brodwin, M., Croton, D. J., & Eisenhardt, P. R. 2007, *ApJ*, **654**, 858
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, *ApJ*, **533**, 682
- Chabrier, G. 2003, *PASP*, **115**, 763
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, *ApJ*, **647**, 201
- Daddi, E., Cimatti, A., Renzini, A., Fontana, A., Mignoli, M., Pozzetti, L., Tozzi, P., & Zamorani, G. 2004, *ApJ*, **617**, 746

- Daddi, E., et al. 2007, *ApJ*, **670**, 156
- Davé, R. 2008, *MNRAS*, **385**, 147
- Dekel, A., Sari, R., & Ceverino, D. 2009, *ApJ*, **703**, 785
- Desai, V., et al. 2008, *ApJ*, **679**, 1204
- Dunne, L., et al. 2009, *MNRAS*, **394**, 3
- Elbaz, D., et al. 2007, *A&A*, **468**, 33
- Elsner, F., Feulner, G., & Hopp, U. 2008, *A&A*, **477**, 503
- Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., & Adelberger, K. L. 2006, *ApJ*, **646**, 107
- Faber, S. M., et al. 2003, *Proc. SPIE*, **4841**, 1657
- Ferguson, H. C., et al. 2004, *ApJ*, **600**, L107
- Finlator, K., Oppenheimer, B. D., & Davé, R. 2011, *MNRAS*, **410**, 1703
- Genzel, R., et al. 2006, *Nature*, **442**, 786
- Giavalisco, M., et al. 2004, *ApJ*, **600**, L103
- Jannuzi, B. T., & Dey, A. 1999, in ASP Conf. Ser. 191, Photometric Redshifts and the Detection of High Redshift Galaxies, ed. R. Weymann et al. (San Francisco, CA: ASP), 111
- Kajisawa, M., Ichikawa, T., Yamada, T., Uchimoto, Y. K., Yoshikawa, T., Akiyama, M., & Onodera, M. 2010, *ApJ*, **723**, 129
- Kennicutt, R. C., Jr. 1998, *ARA&A*, **36**, 189
- Keres, D., Katz, N., Fardal, M., Dave, R., & Weinberg, D. H. 2009, *MNRAS*, **395**, 160
- Kolatt, T. S., et al. 1999, *ApJ*, **523**, L109
- Kron, R. G. 1980, *ApJS*, **43**, 305
- Lee, K., Giavalisco, M., Conroy, C., Wechsler, R. H., Ferguson, H. C., Somerville, R. S., Dickinson, M. E., & Urry, C. M. 2009, *ApJ*, **695**, 368
- Lee, K.-S., Giavalisco, M., Gnedin, O. Y., Somerville, R. S., Ferguson, H. C., Dickinson, M., & Ouchi, M. 2006, *ApJ*, **642**, 63
- Leitherer, C., et al. 1999, *ApJS*, **123**, 3
- Madau, P. 1995, *ApJ*, **441**, 18
- Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, *MNRAS*, **283**, 1388
- Magdis, G. E., Rigopoulou, D., Huang, J., & Fazio, G. G. 2010, *MNRAS*, **401**, 1521
- Mannucci, F., et al. 2009, *MNRAS*, **398**, 1915
- Maraston, C., Pforr, J., Renzini, A., Daddi, E., Dickinson, M., Cimatti, A., & Tonini, C. 2010, *MNRAS*, **407**, 830
- Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., Franx, M., Labbé, I., & Wuyts, S. 2009, *ApJ*, **701**, 1765
- Marchesini, D., et al. 2010, *ApJ*, **725**, 1277
- Martini, P., & Weinberg, D. H. 2001, *ApJ*, **547**, 12
- McLure, R. J., Cirasuolo, M., Dunlop, J. S., Foucaud, S., & Almaini, O. 2009, *MNRAS*, **395**, 2196
- Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, **521**, 64
- Mihos, J. C., & Hernquist, L. 1996, *ApJ*, **464**, 641
- Noeske, K. G., et al. 2007, *ApJ*, **660**, L43
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Ouchi, M., et al. 2005, *ApJ*, **635**, L117
- Pannella, M., et al. 2009, *ApJ*, **698**, L116
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, **559**, 620
- Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, *MNRAS*, **412**, 1123
- Pentericci, L., Grazian, A., Fontana, A., Salimbeni, S., Santini, P., de Santis, C., Gallozzi, S., & Giallongo, E. 2007, *A&A*, **471**, 433
- Quadri, R., et al. 2007, *ApJ*, **654**, 138
- Reddy, N. A., Erb, D. K., Pettini, M., Steidel, C. C., & Shapley, A. E. 2010, *ApJ*, **712**, 1070
- Reddy, N. A., & Steidel, C. C. 2009, *ApJ*, **692**, 778
- Reddy, N. A., Steidel, C. C., Fadda, D., Yan, L., Pettini, M., Shapley, A. E., Erb, D. K., & Adelberger, K. L. 2006, *ApJ*, **644**, 792
- Renzini, A. 2009, *MNRAS*, **398**, L58
- Sawicki, M., & Thompson, D. 2006, *ApJ*, **642**, 653
- Shapley, A. E., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, *ApJ*, **562**, 95
- Shapley, A. E., Steidel, C. C., Erb, D. K., Reddy, N. A., Adelberger, K. L., Pettini, M., Barmby, P., & Huang, J. 2005, *ApJ*, **626**, 698
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *ApJ*, **588**, 65
- Somerville, R. S., Primack, J. R., & Faber, S. M. 2001, *MNRAS*, **320**, 504
- Stark, D. P., Ellis, R. S., Bunker, A., Bundy, K., Targett, T., Benson, A., & Lacy, M. 2009, *ApJ*, **697**, 1493
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, *ApJ*, **519**, 1
- Vanzella, E., et al. 2009, *ApJ*, **695**, 1163