Draft version November 13, 2012
Preprint typeset using LaTeX style emulateapj v. 5/2/11

A CENSUS OF STAR-FORMING GALAXIES IN THE $Z\sim 9$-10 UNIVERSE BASED ON HST+SPITZER OBSERVATIONS OVER 19 CLASH CLUSTERS: THREE CANDIDATE $Z\sim 9$-10 GALAXIES AND IMPROVED CONSTRAINTS ON THE STAR FORMATION RATE DENSITY AT $Z\sim 9.2$[1]

R. J. Bouwens[2,3], L. Bradley[4], A. Zitrin[5], D. Coe[4], M. Franx[2], W. Zheng[6], R. Smit[2], O. Host[7], M. Postman[4], L. Moustakas[8], I. Labbé[2], M. Carrasco[5,9], A. Molino[10], M. Donahue[11], D.D. Kelson[12], M. Meneghetti[13], S. Jha[14], N. Benítez[10], D. Lemze[6], K. Umetsu[15], T. Broadhurst[16], J. Moustakas[17,18], P. Rosati[19], S. Jouvel[20], M. Bartelmann[5], H. Ford[6], G. Graves[21], C. Grillo[22], L. Infante[9], Y. Jimenez-Teja[10], O. Lahav[20], D. Maoz[23], E. Medezinski[6], P. Melchior[24], J. Merten[8], M. Nonino[25], S. Ogaz[4], S. Seitz[26]

Draft version November 13, 2012

ABSTRACT

We utilise a two-color Lyman-Break selection criterion to search for $z\sim 9$-10 galaxies over the first 19 clusters in the CLASH program. Key to this search are deep observations over our clusters in five near-IR passbands to 1.6μm, allowing us good constraints on the position of the Lyman break to $z\sim 10$. A systematic search yields three $z\sim 9$-10 candidates in total above a 6σ detection limit. While we have already reported the most robust of these candidates, MACS1149-JD, in a previous publication, two additional $z\sim 9$ candidates are also revealed in our expanded search. The new candidates have H_{160}-band AB magnitudes of $\sim$26.2-26.9 and are located behind MACSJ1115.9+0129 and MACSJ1720.3+3536. The observed $H_{160}-Spitzer/IRAC$ colors for the sources are sufficiently blue to strongly favor redshifts of $z\geq 9$ for these sources. A careful assessment of various sources of contamination suggests $\lesssim$1 contaminants for our $z\sim 9$-10 selection. To determine the implications of these search results for the LF and SFR density at $z\sim 9$, we introduce a new differential approach to deriving these quantities in lensing fields. Our procedure is to derive the evolution by comparing the number of $z\sim 9$-10 galaxy candidates found in CLASH with the number of galaxies in a slightly lower redshift sample (after correcting for the differences in selection volumes), here taken to be $z\sim 8$. This procedure takes advantage of the fact that the relative selection volumes available for the $z\sim 8$ and $z\sim 9$-10 selections behind lensing clusters are not greatly dependent on the details of the gravitational lensing models. We find that the normalization of the UV LF at $z\sim 9$ is just $0.22^{+0.30}_{-0.15}\times$ that at $z\sim 8$, $\sim 2^{+3}_{-1}\times$ lower than what we would infer extrapolating $z\sim 4$-8 LF results. These results therefore suggest a more rapid evolution in the UV LF at $z>8$ than seen at lower redshifts (although the current evidence here is weak). Compared to similar evolutionary findings from the HUDF, our result is much more insensitive to large-scale structure uncertainties, given our many independent sightlines on the high-redshift universe.

Subject headings: galaxies: evolution — galaxies: high-redshift

[1] Based on observations made with the NASA/ESA Hubble Space Telescope, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

[2] Leiden Observatory, Leiden University
[3] University of California, Santa Cruz
[4] Space Telescope Science Institute
[5] Universitat Heidelberg
[6] The Johns Hopkins University
[7] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen
[8] JPL, California Institute of Technology
[9] Universidad Catolica de Chile
[10] Instituto de Astrofísica de Andalucía
[11] Michigan State University
[12] The Carnegie Institute for Science; Carnegie Observatories
[13] INAF, Osservatorio Astronomico di Bologna
[14] Rutgers University
[15] Academia Sinica, Institute of Astronomy & Astrophysics
[16] University of the Basque Country
[17] University of California, San Diego
[18] Siena College
[19] European Southern Observatory
[20] University College London
[21] University of California, Berkeley
[22] Technische Universitat Munchen
[23] Tel Aviv University
[24] The Ohio State University
[25] INAF, Osservatorio Astronomico di Trieste
[26] Universitas Sternwarte, München

1. INTRODUCTION

Since the discovery of large numbers of $z \sim 3$ galaxies with the Lyman-break selection technique 17 years ago (Steidel et al. 1996), there has been a persistent effort to use the latest facilities to identify galaxies at higher and higher redshifts through photometric selections and follow-up spectroscopy. These efforts allow us to probe galaxies during the epoch of reionization to ascertain what role they may have in driving this process. Progressively, the high-redshift frontier has been extended to $z \sim$ 4-5 (e.g., Madau et al. 1996; Steidel et al. 1999), $z \sim 6$ (e.g., Stanway et al. 2003; Bouwens et al. 2003; Dickinson et al. 2004), $z \sim 7$ (e.g., Bouwens et al. 2004; Yan & Windhorst 2004; Bouwens & Illingworth 2006b; Iye et al. 2006; Fontana et al. 2009; Schenker et al. 2012), and $z \sim 8$ (e.g., Bouwens et al. 2010; McLure et al. 2010; Bunker et al. 2010; Yan et al. 2010).

The current frontier for identifying high-redshift galaxies now seems to lie firmly at $z \sim 10$, with three distinct $z \sim 10$ galaxy candidates having been reported. Bouwens et al. (2011a) presented the discovery of a plausible $z \sim 10.3$ galaxy in the full two-year HUDF09 observations over the HUDF (see also Oesch et al. 2012a). More recently, Zheng et al. (2012: hereinafter Z12) presented evidence for a highly-magnified $z \sim 9.6$ galaxy within the 524-orbit CLASH program (Postman et al. 2011), and Coe et al. (2012: hereinafter C12) reported the discovery of an even higher redshift triply-lensed $z \sim 10.8$ galaxy.

Despite the very interesting nature of earlier exploratory work, the total number of $z \sim$ 9-11 galaxies is small, and hence it is still somewhat challenging to obtain accurate constraints on how rapidly the luminosity function (LF) or star formation rate (SFR) density evolved in the very early universe, at $z > 8$. Earlier $z \sim 10$ searches using the very deep HUDF09 data (Bouwens et al. 2011a; Oesch et al. 2012a) found tentative evidence for a deficit of $z \sim 10$ galaxies relative to simple extrapolations from lower redshifts, pointing towards a very rapid evolution in the UV LF and SFR density at $z > 8$ (Oesch et al. 2012a). A rapid evolution of the UV LF at $z > 8$ is supported by several theoretical models (Trenti et al. 2010; Lacey et al. 2011), but may be in some tension with the discovery of a bright, multiply-lensed $z \sim 10.8$ galaxy in the CLASH program (C12), since one might have expected such sources to be quite rare assuming a rapid evolution of the UV LF.

Fortunately, there is an ever increasing quantity of observations now available (or soon to become available) to identify $z \sim$ 9-10 galaxies. One noteworthy near-term opportunity exists in the moderately deeper WFC3/IR observations acquired over the HUDF/XDF (GO 12498: PI Ellis). This program promises to extend current $z \sim$ 9-10 samples in that field deeper by $\sim$0.4 mag while increasing the number of sources 2-4×. However, another significant opportunity exists in ongoing observations over lensing clusters, as part of the 524-orbit CLASH program (Postman et al. 2012). The initial discovery papers only reported on the brightest and most robust $z \sim 10$ and $z \sim 11$ galaxy candidates from the CLASH program (Z12; C12), but it is possible to extend these searches somewhat fainter by $\sim$0.5-1.0 mag to the magnitude limit of the survey ($\sim$27 AB mag). At such magnitudes, we would expect to identify other plausible $z \sim$ 9-10 galaxies, potentially increasing the overall sample size to $\sim$3-5 sources in total.

The purpose of this paper is to capitalize on this opportunity and to extend $z \sim$ 9-10 searches to the magnitude limit of the CLASH program (Z12; see also C12). A deeper search for $z \sim$ 9-10 galaxies can be performed in a reasonably reliable manner taking full advantage of the substantial observations with Spitzer/IRAC instrument over the CLASH program – allowing us to distinguish potential star-forming galaxy candidates at $z \sim$ 9-10 from lower-redshift interlopers. We also incorporate HST observations over 2 more clusters from the CLASH program (utilizing a total of 19 clusters) to expand the total search area by 50% and 10% over what was considered in Z12 and C12, respectively.

The plan for this paper is as follows. In §2, we describe our observational data set. In §3, we discuss our procedure for catalog creation, the selection of $z \sim$ 9-10 galaxy candidates, quantifying their properties, and estimating the extent to which contamination may be a concern for our selection. In §4, we introduce a new differential approach to derive the evolution in the UV LF and SFR density at $z \gtrsim 9$ and then apply it to our search results at $z \sim 9$. Finally, in §5, we summarize the results from this paper and offer a prospective. Throughout this work, we quote results in terms of the luminosity $L^*_{z=3}$ Steidel et al. (1999) derived at $z \sim 3$: $M_{1700,AB} = -21.07$. We refer to the HST F390W, F435W, F475W, F606W, F625W, F775W, F814W, F850LP, F105W, F110W, F125W, F140W, and F160W bands as U_{390}, B_{435}, g_{475}, V_{606}, r_{625}, i_{775}, I_{814}, z_{850}, Y_{105}, J_{110}, J_{125}, JH_{140}, and H_{160}, respectively, for simplicity. Where necessary, we assume $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70\,\mathrm{km/s/Mpc}$. All magnitudes are in the AB system (Oke & Gunn 1983).

2. OBSERVATIONAL DATA

Our primary dataset for this study are the 20-orbit HST observations over the first 19 clusters with data from the 524-orbit CLASH multi-cycle treasury program (Postman et al. 2012: see Table 1). The HST observations over each of the CLASH clusters is typically distributed over 16 different bands using the WFC3/UVIS camera, the Advanced Camera for Surveys (ACS) wide field camera, and the WFC3/IR camera. These observations extend from 0.2μm (F225W filter) to 1.6μm (F160W) and reach to depths to 26.4-27.7 AB mag (5σ: 0.4″-diameter aperture) depending upon the passband.

Our reductions of these data were conducted using standard procedures, aligned, and then drizzled on the same frame (0.065″ pixel scale) with the multidrizzle software (Koekemoer et al. 2003). The FWHM for the PSF is $\sim$0.1″ in the WFC3/UVIS or ACS observations and $\sim$0.16-0.17″ for the WFC3/IR observations.

The typical area available over each cluster to search for $z \sim$ 9-10 galaxies is $\sim$4 arcmin^2 and is dictated by the area available within the WFC3/IR field-of-view. In total, we make use of $\sim$77 arcmin^2 over the first 19 CLASH clusters to search for $z \sim$ 9-10 galaxies. This corresponds to an approximate search volume of $\sim$7000 Mpc^3 (comoving) at $z \sim 9$ to probe faint, highly magnified $\mu > 5$ galaxies (assuming $\sim$25% of our WFC3/IR area is high magnification $\mu \gtrsim 5$ and a $\Delta z \sim 1$ width for our red-

TABLE 1
The 19 cluster fields from the CLASH program considered in the present $z \sim 9$ search.

Cluster	Redshift	High Magnification[a]
Abell 209	0.206	
Abell 383	0.187	
Abell 611	0.288	
Abell 2261	0.244	
MACSJ0329.70211	0.450	
MACSJ0416.1-2403	0.42	Y
MACSJ0647.8+7015	0.591	Y
MACSJ0717.5+3745	0.548	Y
MACSJ0744.9+3927	0.686	
MACSJ1115.9+0129	0.355	
MACSJ1149.6+2223	0.544	Y
MACSJ1206.20847	0.439	
MACSJ1720.3+3536	0.387	
MACSJ1931.82635	0.352	
MACSJ2129.40741	0.570	Y
MS2137.32353	0.313	
RXJ1347.51145	0.451	
RXJ1532.9+3021	0.363	
RXJ2129.7+0005	0.234	

[a] Clusters in the CLASH program were selected based on either their x-ray or magnification properties (Postman et al. 2012). Clusters marked here with a "Y" were included because of their magnification properties.

shift selection window: see Figure 1). To ensure that we have the maximum depth and filter coverage available for candidates uncovered in our search, we do not consider the small amount of data over each cluster with observations in only one of the two roll angles used for the CLASH program (see figure 11 of Postman et al. 2012 for an illustration of the two roll-angle strategy).

Each of the CLASH clusters also has a substantial amount of observations with the Spitzer/IRAC instrument (Fazio et al. 2004). The typical integration times range from ∼3.5 hours per IRAC band from the ICLASH program (GO #80168: Bouwens et al. 2011c) to ∼5 hours per IRAC band from the Spitzer IRAC Lensing Survey program (GO #60034: PI Egami). These observations reach to 1σ depths of ∼26.1-26.5 mag in both the 3.6μm and 4.5μm IRAC channels, allowing us to set useful constraints on the color of possible $z \sim$ 9-10 candidates redward of the break. The FWHM for the IRAC PSF at 3.6μm and 4.5μm is $\sim 1.8''$. We reduced the Spitzer/IRAC observations using the public MOPEX software available from the Spitzer Science Center (Makovoz et al. 2005). The reductions were drizzled onto a common output frame ($0.3''$-pixel scale).

3. RESULTS

3.1. *Catalog Construction*

Our procedure for constructing catalogs is similar to that previously utilized by Bouwens et al. (2007, 2011b, 2012b). These catalogs are distinct from those distributed as part of the CLASH program, but overall the results are in very good agreement.

We provide a brief outline of the procedure we use here. More details are provided in several of our previous publications (e.g., Bouwens et al. 2007, 2011b, 2012b). SExtractor (Bertin & Arnouts 1996) is run in dual-image mode, using the square root of the χ^2 image (Szalay et al. 1999) to detect sources and the PSF-matched images for photometry. The χ^2 image (similar to a coadded frame) is constructed from the imaging observations in the two passbands where we expect $z \sim 9$ candidates to show significant signal, i.e., the JH_{140} and H_{160} bands. For the photometry, PSF-matching is done to the WFC3/IR H_{160}-band. Fluxes and colors of sources are measured in apertures that scale with the size of sources, as recommended by Kron (1980) and using a Kron factor of 1.2. The small-aperture fluxes are then corrected to total magnitudes in two steps. First the excess flux around the source in a larger scalable aperture (Kron factor 2.5) is derived based on the square root of χ^2 image and this correction is applied to the measured fluxes in all HST bands. Second, a correction is made for the expected light outside the larger scaled aperture and on the wings of the PSF using the tabulated encircled energy distribution (e.g., from Sirianni et al. 2005).

The measurement of IRAC fluxes is important for a more secure identification of $z \sim 9$ candidates in our fields, since it allows us to quantify the approximate spectral slope of the sources redward of the spectral break observed at $\sim 1.2\mu$m and therefore distinguish potential star-forming galaxies at $z \sim$ 9-10 from interlopers at $z \sim$ 1-2. IRAC photometry can be challenging due to the significant overlap between nearby sources in existing data. Fortunately, there are well-established procedures to use the positions and spatial profiles of sources in available HST observations to model the IRAC image observations and extract fluxes (e.g., Shapley et al. 2005; Labbé et al. 2006; Grazian et al. 2006; Laidler et al. 2007).

Here we make use of the "Ivophot" software (Labbé et al. 2006, 2010a, 2010b, 2012) to do photometry on sources in our fields in the midst of this confusion. Since this software has been presented more extensively in other places, we only include a brief description here. The most important step for doing photometry on faint sources with this software is to remove confusion from neighboring sources. This is accomplished by using the deep WFC3/IR observations as a template to model the positions and isolated flux profiles of the foreground sources. These flux profiles are then convolved to match the IRAC PSF and then simultaneously fit to the IRAC imaging data leaving only the fluxes of the sources as unknowns. The best-fit model is then used to subtract the flux from neighboring sources and normal aperture photometry is performed on sources in a $2.5''$-diameter aperture. The measured 3.6μm and 4.6μm fluxes are then corrected to account for the light on the wings of the IRAC PSF (typically the correction is a factor of ∼2.2).

3.2. *Source Selection*

In this paper, we adopt a two-color Lyman-break selection to search for promising $z \sim$ 9-10 galaxy candidates in the CLASH program. This work takes advantage of the sharp break in the spectrum of star-forming galaxies due to absorption by neutral hydrogen. Many years of spectroscopic work has shown that the Lyman-break selection technique is very effective in selecting large numbers of very high-redshift galaxies, with minimal contamination (Steidel et al. 1996; Steidel et al. 2003; Bunker et al. 2003; Dow-Hygelund et al. 2007; Popesso et al. 2009; Vanzella et al. 2009; Stark et al. 2010).

In analogy with lower-redshift Lyman-break selections

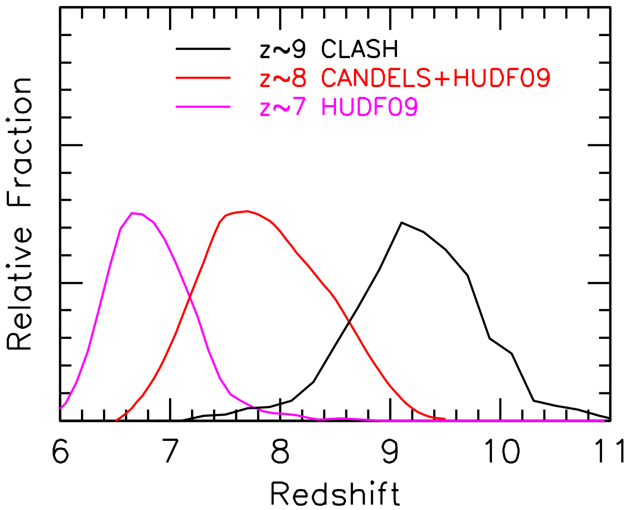

FIG. 1.— The redshift distribution we would expect for our present $z \sim 9$ selection based on the simulations we run in §4.3. These simulations allow us to assess the relative selection volume for our $z \sim 9$ selection and our comparison sample at $z \sim 8$. The mean redshift for our selection is 9.2. Our $z \sim 9$ selection cuts off at $z > 10$ due to our use of a $JH_{140} - H_{160} < 0.5$ criterion (§3.2: see also Figure 2). For context, we also show the expected redshift distributions for the $z \sim 7$ and $z \sim 8$ selections of Bouwens et al. (2011b) and Oesch et al. (2012b), respectively.

(e.g., Bouwens et al. 2007; Bouwens et al. 2011b), we devised the following two-color $z \sim$ 9-10 selection for the CLASH cluster fields:

$$((J_{110} + J_{125})/2 - H_{160} > 0.7) \wedge (JH_{140} - H_{160} < 0.5)$$

This criterion is very similar to the criteria previously presented in Z12, i.e., $(J_{110} - JH_{140} > 0.5) \wedge (JH_{140} - H_{160} < 0.5)$, but probe to slightly higher redshift sources on average, also folding in information from the redder J_{125}-band filter and requiring a sharper break in the spectrum. In general, it makes sense to combine the flux information from both the J_{110} and J_{125} bands to search for $z \gtrsim 9$ candidates because of their similar red-side cut-offs at 1.4μm. In applying the above criteria, the magnitudes of sources not detected at 1σ are set to their 1σ upper limits.

It is also important we detect sources at sufficient S/N that we can rely on the color information (and optical non-detections) to provide reliable redshift information on the sources and guarantee they are real. After some experimentation and extensive simulations (§3.5), we elected to require sources in our $z \sim 9$ selection be detected at $\geq 6\sigma$ in a combined JH_{140} and H_{160} bands (using a fixed $0.35''$-diameter aperture). For significance thresholds less than 6σ, our simulations (§3.5) suggest that our $z \sim 9$ selection would be subject to significant contamination from lower redshift interlopers.

To ensure that sources really have no flux in the spectrum blueward of the Lyman break, we also require sources be undetected ($<2.5\sigma$) in the Y_{105} band and any passband blueward of this.[27] Moreover, we combine the flux in all the bluer bands (U_{390}, B_{435}, g_{475}, V_{606}, r_{625}, i_{775}, I_{814}, z_{850}, and Y_{105}) to construct a χ^2 statistic for sources in our catalogs and exclude sources from our selection if the χ^2_{opt+Y} statistic is greater than 3.8. The particular threshold for χ^2_{opt+Y}, i.e., 3.8, was chosen to keep contamination in our $z \sim 9$ sample relatively low while not overly impacting the completeness of our samples (see figure 19 from Bouwens et al. 2011b for an illustration of how such a choice can be made). This criterion ensures that sources are not consistently detected at $> 1\sigma$ in more than three optical bands.

Here χ^2 is calculated as follows: $\chi^2_{opt+Y} = \Sigma_i \mathrm{SGN}(f_i)(f_i/\sigma_i)^2$ where f_i is the flux in band i in a consistent aperture, σ_i is the uncertainty in this flux, and $\mathrm{SGN}(f_i)$ is equal to 1 if $f_i > 0$ and -1 if $f_i < 0$ (Bouwens et al. 2011b). As in Bouwens et al. (2011b), we calculate this χ^2 statistic in three different apertures (scalable Kron apertures [Kron factor of 1.2], $0.35''$-diameter circular apertures, $0.18''$-diameter circular apertures) to ensure that there is absolutely no evidence for a significant excess of light blueward of the break, whether this light be tightly concentrated on the source itself or more diffuse. When computing the χ^2 statistic with $0.18''$-diameter apertures, we use the original unsmoothed ACS or WFC3/IR images (i.e., before PSF-matching to the WFC3/IR H_{160}-band data) to retain the maximum signal-to-noise for the purposes of rejecting low-redshift interlopers.

As one final step to ensure that our $z \sim 9$ candidates show no evidence for flux blueward of the break, we construct a second χ^2 statistic for each source, utilizing only the information in the three bands immediately blueward of the break, i.e., the I_{814}, z_{850}, and Y_{105} bands. We then exclude any source which has an χ^2_{I+z+Y} value greater than 3. This criterion ensures that sources are not detected at $\gtrsim 1\sigma$ on average in the I_{814}, z_{850}, and Y_{105} bands. Sources detected at $>2\sigma$ in the Y_{105}-band are also excluded to minimize the contribution of $z \sim 8$ galaxies to our selection.

In Figure 1, we show the approximate redshift selection window for our current selection. Details on how it is calculated will be presented in §4.3, but approximately involve adding artificial sources to the real data with realistic colors, sizes, and magnitudes, and then attempting to reselect them with the criteria given above. The mean redshift we derive for our $z \sim$ 9-10 selection from the simulations is 9.2. For context, we also present the redshift selection windows for samples at $z \sim 7$ and $z \sim 8$, as selected by Bouwens et al. (2011b) and Oesch et al. (2012b), respectively.

3.3. *Resulting $z \sim 9$ Sample*

We applied the selection criteria given in the previous section to the HST WFC3/UVIS+ACS+WFC3/IR observations from all 19 clusters in the current data set. We identified three sources which satisfy these selection criteria. The sources are found behind three different clusters MACSJ1149.6+2223, MACSJ1115.9+0129, and MACSJ1720.3+3536. The brightest of our three candidates, i.e., MACS1149-JD, was already presented in Z12. These sources were also flagged as the most interesting $z > 8$ sources using an independent, purely photometric redshift selection (L. Bradley et al. 2012, in prep).

We performed Spitzer/IRAC photometry on all three

[27] Since we combine the optical flux measurements into several χ^2 statistics that we use to test the plausibility of specific sources as $z \sim 9$ candidates, we only adopt a weaker 2.5σ threshold here to avoid unnecessarily excluding many plausible $z \sim 9$ candidates.

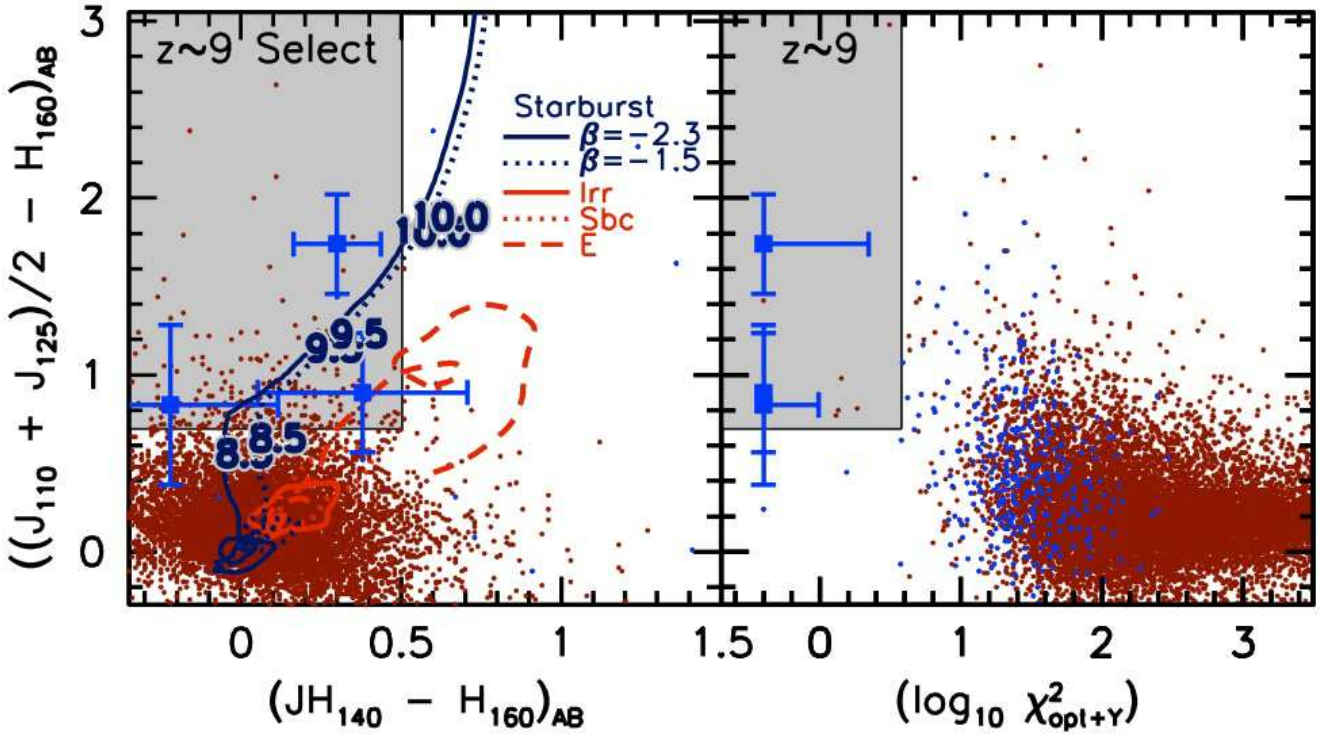

FIG. 2.— Selection criteria used here to identify $z \sim$ 9-10 galaxies over the CLASH program. (*left*) The $((J_{110}+J_{125})/2 - H_{160})_{AB}$ vs. $(JH_{140} - H_{160})_{AB}$ diagram shows the first of our two primary criteria we use to identify $z \sim$ 9-10 galaxies from the CLASH program. Selected sources must fall in the gray region defined by two LBG-like color criteria, with a $(J_{110}+J_{125})/2 - H_{160} > 0.7$ criterion defining the Lyman break and a $JH_{140} - H_{160} < 0.5$ criterion providing a constraint on the spectral slope redward of the break. The large blue squares show the sources that made it into our $z \sim$ 9-10 sample. The error bars on these points are the 1σ uncertainties. The blue lines show the expected colors for star-forming galaxies with varying UV-continuum slopes as a function of redshift while the red lines show the expected colors for different SED templates at lower redshift (Coleman et al. 1980). The small dark red points show the colors of sources in our photometric sample where the χ^2_{opt+Y} statistic is > 3.8. The blue points show these colors for sources where the χ^2_{opt+Y} statistic is < 3.8. See §3.2 (and Bouwens et al. 2011b) for a definition of the χ^2_{opt+Y} statistic, but it roughly includes a stack of all the flux information in the Y_{105} band and bluer bands. (*right*) The $((J_{110}+J_{125})/2 - H_{160})_{AB}$ vs. χ^2_{opt+Y} diagram shows the second of our two primary criteria we use to identify $z \sim$ 9-10 galaxies from the CLASH program. The selected sources must fall in the gray region and therefore must show no flux in the optical or Y_{105} bands (i.e., $\chi^2_{opt+Y} < 3.8$). The three selected $z \sim 9$ candidates are the blue squares. The dark red points indicate sources in our photometric sample which are either detected in the Y_{105} band ($>2\sigma$) or where the $JH_{140} - H_{160}$ colors are greater than 0.5. The blue points are those sources where neither condition is satisfied. This figure is similar to Figure 2 of Oesch et al. (2012b). Using both the two-color criteria and our χ^2_{opt+Y} criteria, we observe a clear separation between our $z \sim$ 9-10 candidates and the bulk of our photometric sample.

candidates using the software described in §3.1. None of the three sources is nearby a bright foreground source and so all of our IRAC flux measurements should be reliable. None of the sources are detected at $>3\sigma$ significance in the Spitzer observations, and an average of the 3.6μm and 4.5μm flux measurements show a range of measured magnitudes, from $\sim$25.3 mag for MACS1149-JD to an upper limit on the IRAC $3.6\mu + 4.5\mu$ flux of MACSJ1115-JD1.

The coordinates and photometry of these candidates are provided in Table 2, while postage stamp images of the candidates are shown in Figure 3. In Table 2, we also present a mean spectral energy distribution for galaxies at $z \sim 9$, which we computed on the basis of our HST+Spitzer photometry for the three $z \sim 9$ candidates. In computing this mean SED, the fluxes of each source are rescaled such that its average JH_{140}+H_{160} flux matches the average $JH_{140} + H_{160}$ flux for the sample (prior to rescaling).

As shown in Figure 3, MACS1149-JD is clearly resolved (see the Supplementary Information to Z12). MACS1149-JD also shows distinct elongation along the shear axis (Figure 1 from Z12) predicted from our gravitational lensing model for MACSJ1149.6+2223 (Z12). The other two plausible $z \sim 9$ candidates in our selection are quite small and show no clear evidence for gravitational shearing in the expected directions. However, since we would expect faint $z \geq 9$ galaxies to be small and the predicted magnification to be only modest (magnifications of $\sim$5-9$\times$ in total), it is not clear that the structural properties of the sources teach us anything definitive.

In Figure 4, we indicate the position of these candidates within the field of view of our MACSJ1149.6+2223, MACSJ1115.9+0129, and MACSJ1720.3+3536 observations (*magenta circles*). On Figure 4, we have also overplotted the approximate critical lines for these clusters based on the lens models we have for these clusters (*white contours*: Z12; Zitrin et al. 2012, in preparation; Carrasco et al. 2012, in preparation). We caution that the lens models we have for MACSJ1115.9+0129 and MACSJ1720.3+3536 are still somewhat preliminary and are not totally finalized yet. The models are constructed based on the assumption that mass traces light, with typ-

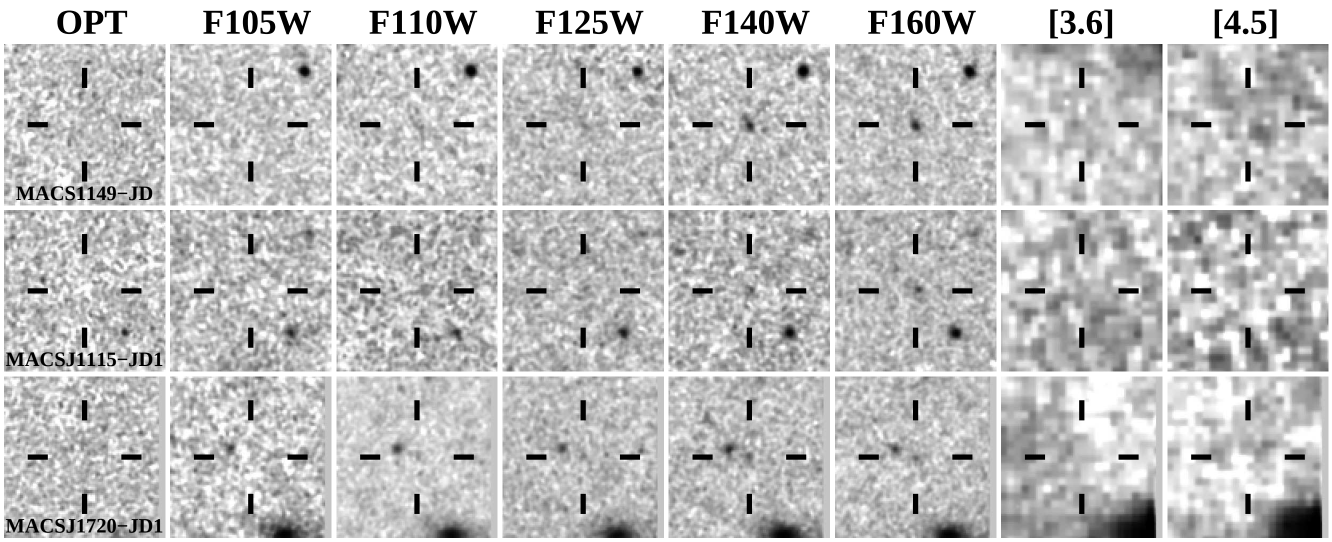

FIG. 3.— Postage stamp images (6.6″×6.6″) of the three $z \sim 9$ galaxy candidates we identify in the current 19-cluster CLASH observations. The source in uppermost row is the same $z \sim 9.6$ candidate as we reported in Z12 (though our redshift estimate for this source is a very consistent $z \sim 9.7$: see §3.4). The leftmost postage stamp shows a stack of the deep ACS $B_{435} + g_{475} + V_{606} + r_{625} + i_{775} + I_{814} + z_{850}$ optical observations, while the other stamps show the observations in specific HST WFC3/IR and Spitzer/IRAC bands. All three of our $z \sim 9$ candidates are detected at $> 6.8\sigma$ in a coadded $JH_{140} + H_{160}$ image (0.35″-diameter aperture: see Table 2). The Spitzer fluxes we measure for the sources are sufficiently faint, as to substantially prefer a $z > 6$ solution for the sources rather than a low redshift solution. None of the sources show any significant detections in the optical ACS observations.

ically only one lower-redshift system for normalization.

We can use these magnification models to estimate the approximate magnification factors for our candidate $z \sim 9$ galaxies. The approximate magnification factors are 14.5, 9.3, and 5.0 and suggest intrinsic delensed $H_{160,AB}$ magnitudes for the sources of 28.5, 28.6, and 28.6 mag, respectively, for MACS1149-JD, MACSJ1115-JD1, and MACSJ1720-JD1. The intrinsic magnitudes inferred for the first three $z \sim 9$ galaxy candidates in the CLASH sample are only slightly brighter than was found for the Bouwens et al. (2011a) $z \sim 10$ candidate, i.e., $H_{160,AB} \sim 28.7$ mag, and seem consistent with expectations.

The predicted positions of any possible counterimages to our $z \sim 9$ candidates are also shown on Figure 4 (*dashed yellow circles*). The only case where the counterimages are expected to be bright enough to detect is for MACSJ1720-JD1. Unfortunately, we were unable to locate the counterimages to MACSJ1720-JD1 at the predicted positions − which could mean that our lensing model may require further refinements, the counterimages are blended with foreground sources, or that the redshift identification is incorrect. For MACSJ1115-JD1, the counterimage is expected to be too faint to detect.

3.4. *Best-fit Photometric Redshifts*

The three candidate $z \sim 9$ galaxies presented in the previous section were selected using a two-color Lyman-Break selection, and therefore their photometry is likely a reasonable fit to a model star-forming galaxy SED at $z \sim 9$. However, since one can often fit the same photometry with SED templates at different redshifts, it is worthwhile for us to examine these candidates using standard photometric redshift procedures to look for possible degeneracies. Our use of photometric redshift procedures also allow us to naturally fold in the IRAC flux information we have for our $z \sim 9$ candidates.

To this end, we used the EAZY photometric redshift software (Brammer et al. 2008) to estimate photometric redshifts for the sources based on the observed photometry and to calculate the relative probability that sources in sample are more likely star-forming galaxies at $z \sim 9$ or galaxies at lower redshift (i.e., $z < 3$). The photometric redshift fitting is conducted using the EAZY_v1.0 template set. This template set consists of five SED templates from PEGASE library (Fioc & Rocca-Volmerange 1997) derived based on the Blanton & Roweis (2007) algorithm and one young, dusty template (50 Myr, $A_V = 2.75$).

We consider three different priors in looking at the redshift likelihood distribution of our three $z \sim 9$ candidates: (1) a flat prior, (2) a prior calibrated to published LFs or LF trends, and (3) a prior tuned to reproduce the contamination rate estimated in the next section (§3.5). Our second prior is based on the LF results of Giallongo et al. (2005) and R. Quadri et al. (2012, private communication) for red $z \sim 1.3$-2 galaxies while at $z > 7$ we utilize the LF-fitting formula of Bouwens et al. (2011b). The third prior accounts for the effect of noise on the photometry of lower-redshift galaxies in our search fields and the fact that in some rare events, noise could cause ∼1-2 sources from our fields to seem like highly probable $z \sim 9$ galaxies (§3.5). Our third prior is calibrated to reproduce the results from our photometric scattering experiments. A more detailed description of these priors is provided in Appendix A.

The results are shown in Figure 5. The left panels show a comparison of the observed photometry with the best-fit $z \sim 9$-10 galaxy (*blue line*) and best-fit $z < 3$ galaxy (*red line*), while the right panels show the probability that a given source in our sample has a particular redshift. The best-fit redshifts for MACS1149-JD, MACSJ1115-JD1, and MACSJ1720-JD1 using the flat priors were 9.7, 9.2, and 9.0, respectively. The 68% confidence intervals on the derived redshifts based on these same priors are [9.42,9.91], [8.64,9.55], and [8.16, 9.34], respectively.

TABLE 2
COORDINATES, ESTIMATED REDSHIFTS AND MAGNIFICATION FACTORS, AND PHOTOMETRY FOR PRESENT $z \sim 9$ SAMPLE.[a]

	MACS1149-JD[b]	MACSJ1115-JD1	MACSJ1720-JD1	Stack[c]
R.A.	11:49:33.58	11:15:54.50	17:20:12.76	—
Decl	22:24:45.7	01:29:47.9	35:36:17.5	—
z_{photo}[d]	$9.7^{+0.2}_{-0.3}$[e]	$9.2^{+0.4}_{-0.6}$	$9.0^{+0.3}_{-0.8}$	—
Magnification	14.5	9.3	5.0	—
S/N ($JH_{140} + H_{160}$)[f]	15.4	7.8	6.9	—
U_{390}	−8±25	−14±37	16±32	1±18
B_{435}	−1±26	−117±39	4±32	−35±19
g_{475}	−3±19	−23±25	−10±20	−12±12
V_{606}	−9±14	−0±35	−11±28	−7±16
r_{625}	−27±22	10±24	−9±17	−6±11
i_{775}	0±27	49±47	−35±38	0±22
I_{814}	−3±11	−13±20	−27±17	−16±10
z_{850}	−38±34	−32±55	4±39	−15±25
Y_{105}	−3±17	−39±23	−20±20	−21±12
J_{110}	27±13	37±19	22±13	26±8
J_{125}	56±16	63±21	44±16	49±10
JH_{140}	146±15	80±22	80±15	86±10
	(=26.0±0.1)	(=26.6±0.3)	(=26.6±0.2)	(=26.6±0.1)
H_{160}	193±15	115±19	66±16	100±10
	(=25.7±0.1)	(=26.2±0.2)	(=26.9±0.3)	(=26.4±0.1)
[3.6]	219±115	46±147	234±186	155±94
[4.5]	353±133	−103±167	209±122	122±78
$\frac{1}{2}$([3.6] + [4.5])	286±88	−57±111	221±111	139±61
	(=25.3±0.3)	(>26.3)	(=25.5±0.5)	(=26.0±0.5)

[a] The fluxes in this table are in units of nJy.
[b] The same candidate as is presented in Z12. The fluxes presented in this table were derived independently from those presented in Z12, but are very similar in general.
[c] This column gives the average fluxes in all HST+IRAC bands blueward of 0.4μm for the three $z \sim 9$ candidates in our selection. The fluxes of each source are rescaled such that its average $JH_{140}+H_{160}$ flux matches the average $JH_{140} + H_{160}$ flux of the sample (prior to rescaling).
[d] These photometric redshift estimates are based on the EAZY photometric redshift software (Brammer et al. 2008: see §3.4). In §3.4, we also provide photometric redshift estimates for sources using BPZ and Le PHARE.
[e] Z12 prefer a slightly lower redshift of 9.6 for this source based on the photometry, but within the uncertainties, the present estimate is fully consistent with that given in Z12.
[f] S/N of our $z \sim 9$ candidates in the JH_{140} and H_{160} bands added in quadrature (0.35″-diameter circular aperture). The S/N limit for our $z \sim 9$ selection was 6.0. Our highest S/N candidates are much less likely to correspond to lower-redshift contaminants (see §3.5, Figure 5, and Figure 6).

No substantial changes in these results are seen using our other two priors, except for the integrated probability within the $z \sim$ 1-2 peak. For our second LF-calibrated prior (*red line*), the lower-redshift peak is actually smaller than in the case of the flat prior. This simply reflects the extreme rarity of faint red (old and/or dusty) galaxies at $z \sim$ 1.3-2 as found in the Giallongo et al. (2005) and R. Quadri et al. (2012, private communication) probes (see also Stutz et al. 2008 and Figure 11 from Oesch et al. 2012a). For our third prior (*dotted blue lines*), the lower-redshift peak is larger, particularly in the case of MACSJ1720-JD1. While this prior was tuned so as to reproduce the expected contamination level for our $z \sim 9$ selection over the first 19 CLASH clusters (suggesting some possible contamination of our selection by lower redshift interlopers), we should emphasize that we have no particular evidence that MACSJ1720-JD1 corresponds to such a source.

We also derived redshift likelihood confidence intervals using the Le PHARE photometric redshift package (Arnouts et al. 1999; Ilbert et al. 2006, 2009) for our three candidates. The SED templates we used with Le PHARE were the same ones as optimized for the COSMOS survey (Scoville et al. 2007) and included three elliptical and six spiral SEDs as generated by Polletta et al. (2007) using the GRASIL code (Silva et al. 1998) as well as 12 starburst galaxies ranging in age from 30 Myr to 3 Gyr using the Bruzual & Charlot (2003) GALAXEV library. We supplemented these with four additional elliptical emplates for a total of seven elliptical templates. Dust extinction was added in ten steps up to $E(B-V) = 0.6$. With these templates, we used Le PHARE to derive the following 68% confidence intervals for the candidates: [9.41,9.84] for MACS1149-JD, [8.86,9.66] for MACSJ1115-JD1, and [8.55,9.25] for MACSJ1720-JD1. The best-fit redshifts for these three candidates were 9.6, 9.3, and 8.9, respectively. The above results are for a flat prior and are quite similar to what we derived using EAZY. Use of the two other priors resulted in similar changes to the redshift likelihood distributions as shown in Figure 5.

Finally, we also estimated the photometric redshifts of our three candidates with BPZ (Bayesian Photometric Redshift Code: Benítez 2000; Coe et al. 2006). Similar to the analyses in C12 and Z12, we modelled the photometry using SEDs from PEGASE (Fioc & Rocca-Volmerange 1997) adjusted and recalibrated to match the observed photometry of galaxies with known spectroscopic redshifts in the FIREWORKS catalog (Wuyts et al. 2008). This FIREWORKS catalog includes pho-

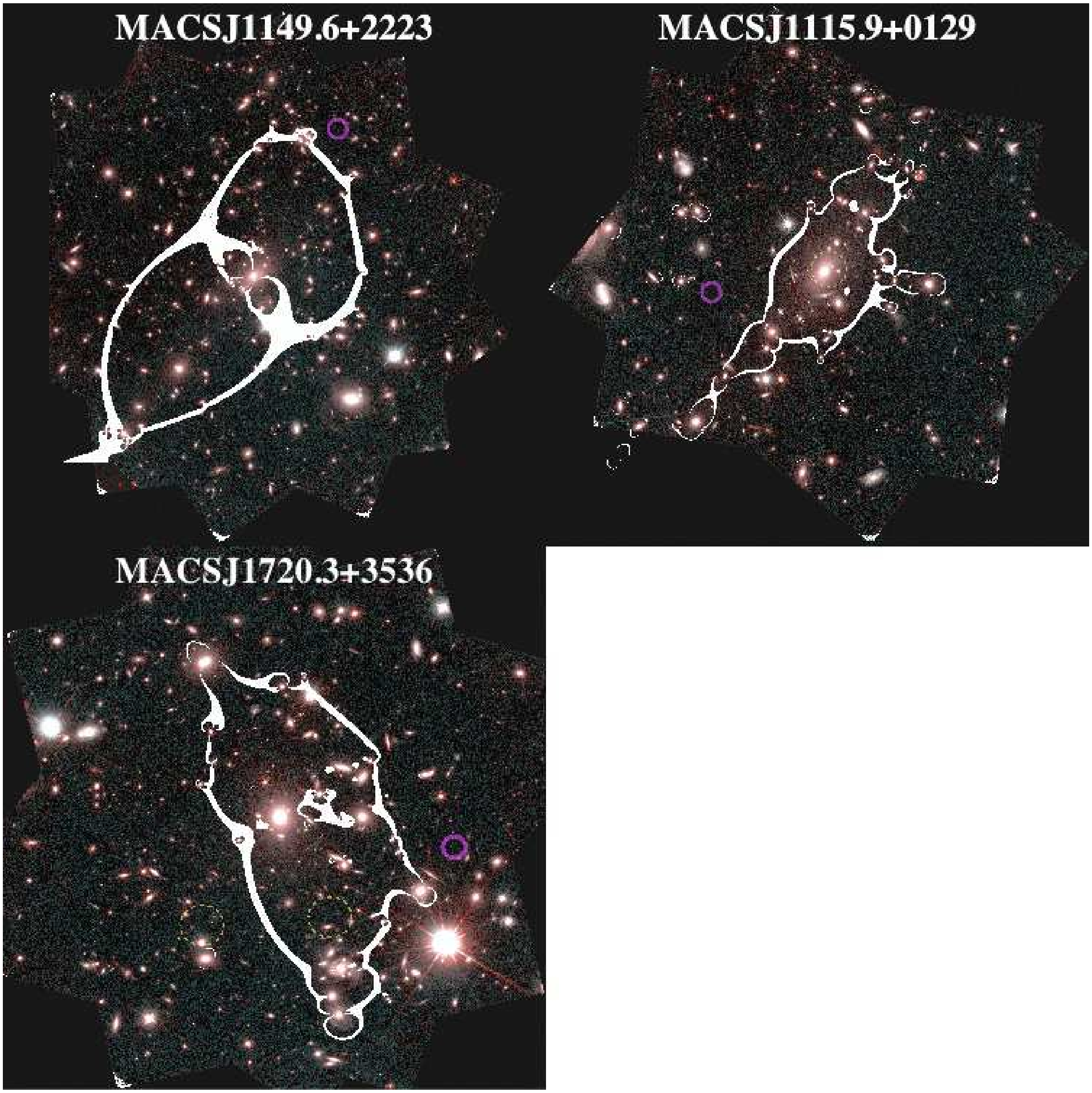

FIG. 4.— Position of the three $z \sim 9$ galaxy candidates we identify over MACSJ1149.6+2223, MACSJ1115.9+0129, and MACSJ1720.3+3536. The color images shown are based on the HST I_{814} + H_{160} observations of these clusters with CLASH and are shown over those regions with deep WFC3/IR observations. Overlaid on these images are the expected ultra high-magnification regions ($\mu > 100$) for a source at $z = 9.2$ based on the gravitational lensing models we have for the three clusters (Z12; A. Zitrin et al. 2012, in prep; M. Carrasco et al. 2012, in prep). Our lensing models for MACSJ1115.9+0129 and MACSJ1720.3+3536 are still preliminary and have not yet been finalized, constructed merely with the assumption that mass traces light, with typically only one lower-redshift system for normalization. The position of our three candidates is indicated by the large magenta circles. The dashed yellow circles indicate the position of possible counterimages as predicted by our preliminary lensing models.

tometry to 24.3 AB mag (5σ) in K_s, for galaxies with $z \sim 3.7$. The best-fit photometric redshifts we derive with BPZ are 9.7, 9.3, and 8.9 for MACS1149-JD, MACSJ1115-JD1, and MACSJ1720-JD1, respectively. For MACS1149-JD, the redshift likelihood distribution is predominantly uni-modal though in the other two cases the distribution is more bimodal, with modest peaks at lower redshift. 17% and 30% of the total probability for MACS1115-JD1 and MACS1720-JD1, respectively, is at $z \sim$ 1.5-1.8. Focusing on the dominant $z \sim 9$ peaks (excluding all $z < 5$ solutions), the 68% confidence intervals on the redshifts for our three candidates are [9.53, 9.87], [8.63, 9.58], and [8.30,9.26], respectively. These results are for a flat prior and are somewhat similar to what we derived using the other two photometric redshift codes, although the low-redshift peaks are slightly more significant with BPZ. We opted not to make use of the BPZ prior in computing the redshift distribution

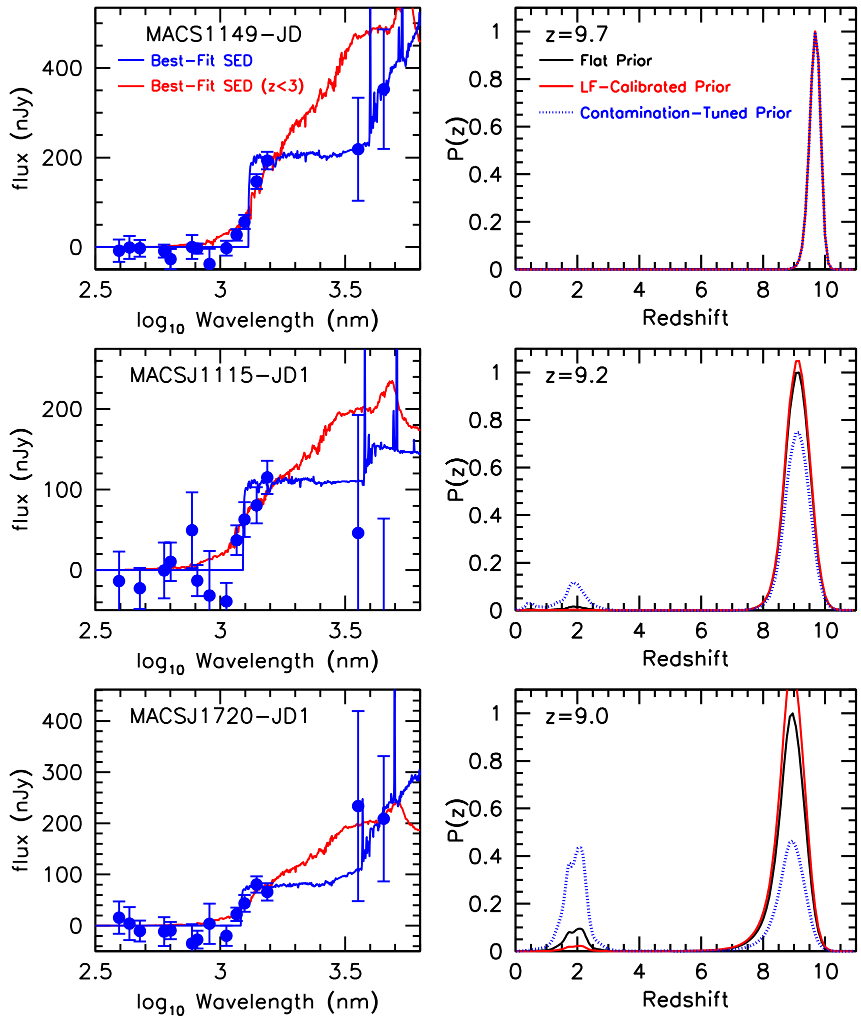

FIG. 5.— (*left*) Observed spectral energy distributions (*solid blue circles*) for three $z \sim 9$ galaxy candidates in our selection. The blue line shows the SED template which best fits our observed photometry (using the EAZY photometric redshift code), while the red line shows the best-fit $z \sim$ 0-3 SED template. The candidate in the uppermost row was previously presented in Z12. (*right*) The redshift likelihood distribution computed for our three $z \sim 9$ candidates using the EAZY photometric redshift software (see §3.4). We consider three different priors in computing the redshift likelihood distributions: (1) a flat prior (*black line*), (2) a prior calibrated to reproduce published LFs or LF trends (*red line*: Giallongo et al. 2005; Bouwens et al. 2011b; R. Quadri et al. 2012, private communication), and (3) a prior tuned to reproduce the results from our photometric scattering simulations (*dotted blue line*: §3.5). Appendix A provides a more detailed description of these priors. Results from our third prior account for the fact that $\sim$1 faint source from our selection might be expected to resemble plausible $z \sim 9$ galaxies, due to the effects of noise (see §3.5). However, even though we might expect a source to possibly scatter into our $z \sim 9$ selection, we have no evidence that any particular source in our sample actually corresponds to such a low-redshift interloper.

for our sources, due to the relative weight it assigns to faint red galaxies at $z \sim$ 1.3-2 and blue galaxies at $z \sim 9$ (which differs by more than a factor of 30 from what we compute based on published LFs or LF trends: see Appendix A). We find that a flat prior comes much closer to accurately representing the relative surface densities of these two populations.

3.5. *Possible Contamination*

While the sources in our current selection are consistent with being $z \sim 9$ galaxies, these sources are faint enough that they could easily have a very different nature. Important sources of contamination for $z > 8$ selections include low-mass stars, supernovae, emission line galaxies (van der Wel et al. 2011), and the photometric scatter of various low-redshift galaxies. Readers are referred to Bouwens et al. (2011a), Z12, and C12 for rather extended discussions of these issues.

In general, the most important source of contamination for high-redshift samples results from faint lower-redshift galaxies entering these samples (and hence satisfying their selection criteria) due to the effects of noise. Noise can cause intrinsically red, optically-detected galaxies to look bluer and disappear entirely at optical wavelengths.

Here we test for contamination from faint lower-redshift sources scattering into our high-redshift selection through the effects of noise, by using all intermediate magnitude sources in the CLASH cluster fields that are detected at $> 2\sigma$ in the I_{814} band and Y_{105} bands (and therefore likely at redshifts $z < 6$) to implicitly define the

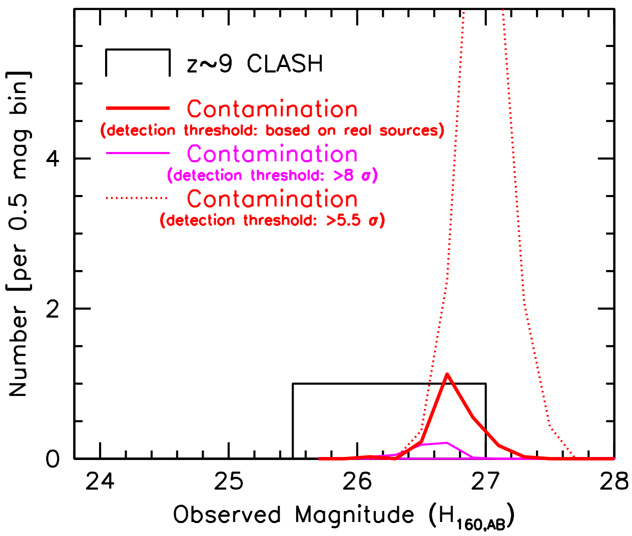

FIG. 6.— The number of $z \sim 9$ galaxy candidates we find in our CLASH cluster search as a function of the $H_{160,AB}$-band magnitude. Also plotted (*red line*) is the number of contaminants we would expect to select in our search fields for sources, due to the effects of noise on the photometry of other lower redshift sources in search fields (see §3.5 for details). The total number that we estimate for our search fields is 0.7 (versus the 3 $z \sim 9$ candidates in our selection). In modeling possible contamination of our selection, we only allow for three contaminants at maximum and the nth contaminant must have a higher signal to noise than the nth lowest signal-to-noise source. For context, we also show the contamination expected for a $>8\sigma$ selection and for a $>5.5\sigma$ selection. Clearly, contamination from lower redshift sources (due to photometric scatter) is only especially significant for sources with $H_{160,AB}$-band magnitudes faintward of 26.5 AB mag. For sources detected at just 5.5σ in the $JH_{140} + H_{160}$ bands (with magnitudes $\sim$27 AB mag), contamination from lower redshift becomes very important.

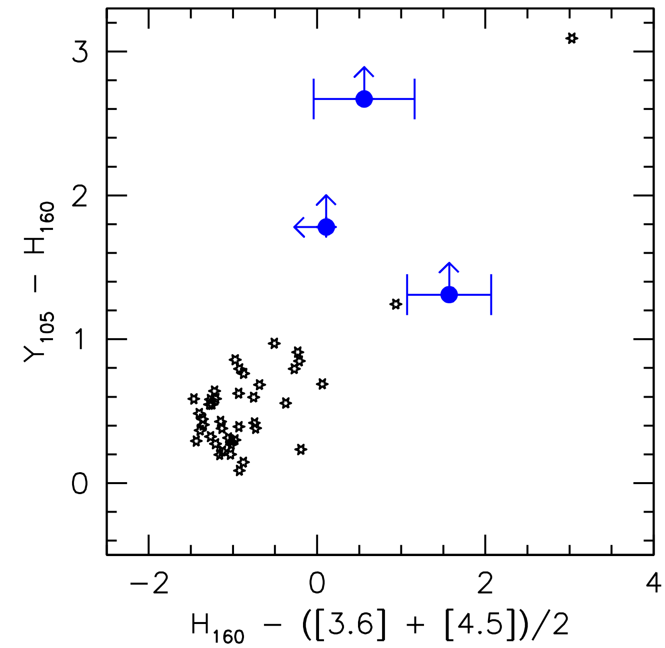

FIG. 7.— A comparison of the $Y_{105} - H_{160}$ vs. $H_{160} - ([3.6] + [4.5])/2$ colors of the three $z \sim 9$ candidates in our sample (*solid blue circles*, 1σ *error bars*, and *arrows indicating* 1σ *limits*) with the observed colors of various stars. The black starlike symbols are the colors derived from the substantial library of stellar spectra observed with IRTF (Cushing et al. 2005; Rayner et al. 2009), with sources ranging from very low-mass stars to higher mass Mira-type variable stars (*the black starlike symbol in the upper right of this figure*). Two of our $z \sim 9$ candidate galaxies have $Y_{105} - H_{160}$ colors which are clearly too red to match those colors observed by the broad set of stars encompassed by this library.

color distribution for potential interlopers to our high-redshift samples. Then, we take all the faint sources in all the CLASH cluster fields (with their $H_{160,AB}$ magnitudes and errors), randomly match them up with a source from the sample which defines our color distribution, give this faint source the same colors as the intermediate-magnitude source, add noise to the photometry of the sources in its bluer bands (assuming a normal distribution), and then see if this source satisfies our $z \sim 9$ selection criteria. Our procedure here is essentially identical to what we performed in many previous analyses (e.g., Bouwens et al. 2011a; Bouwens et al. 2011b). In modeling possible contamination of our selection, we only allow for three contaminants at maximum and the nth contaminant per CLASH data set must have a higher signal to noise than the nth lowest signal-to-noise source.

Applying this procedure to all the sources in the CLASH fields 100×, we find that only 0.7 lower-redshift ($z \lesssim 6$) sources enter our $z \sim 9$ selection by chance (per Monte-Carlo simulation for the entire CLASH program). The magnitude distribution of these contaminants is shown in Figure 6. The small number of contaminants we find from these simulations demonstrate that the overall level of contamination for the present probe is likely only modest ($\sim$23%), becoming important faintward of 26.5 mag.

We also considered the implications for contamination if we had restricted our selection to sources with a $JH_{140} + H_{160}$ detection significance of $>8\sigma$ and $>5.5\sigma$ (weaker than our $z \sim 9$ selection criteria). The results are shown in Figure 6 with the magenta and dotted red lines, respectively. Only $\sim$0.2 contaminants are expected for a 8σ detection threshold while for a $\sim 5.5\sigma$ threshold the expected number of contaminants is $\sim$7.5 sources – and hence might be a significant concern if we had considered a lower detection threshold for our selection.

This being said, it is worth noting that our estimate of the total contamination here may be a little high (perhaps by a factor of $\sim$2-3), due to our use of an intermediate magnitude ($\sim$24.0-25.5 mag) population of galaxies to model the colors of somewhat fainter galaxies (i.e., 26.0-27.0 mag). Since the intermediate magnitude population are somewhat redder in general than $\sim$26.0-27.0 mag population (e.g., see Figure 11 of Oesch et al. 2012a), they are more likely to scatter into $z \sim 9$ selections via noise than is the actual situation for $\sim$26.0-27.0 mag galaxies.

In any case, the results of these simulations strongly suggest that the two most significantly detected sources in our sample, i.e., MACSJ1115-JD1 and especially MACS1149-JD, are unlikely to be contaminants. For sources with lower S/N than this, we must remain concerned about contamination – even though we cannot establish the exact rate. The issue will contribute to the overall errors in our SFR density estimates at $z \sim 9$ (§4).

Other possible sources of contamination are from stars or SNe. Both possibilities would require that sources in our selection are unresolved. Comparing the coadded $JH_{140} + H_{160}$ profile of our candidates with the WFC3/IR PSF, it is clear that 2 out of our 3 candidates are resolved (see also discussion in Z12 which demonstrate clearly that MACS1149-JD is resolved). Only MACSJ1115-JD1 does not show any spatial extension. In any case, as Figure 7 demonstrates, the colors of the candidates do not clearly support a stellar origin. The redder Mira-variable stars would appear to give the best match, but their intrinsic luminosities are such that we would need to observe them well outside our own Milky Way galaxy (Whitelock et al. 1995; Dickinson et al. 2000). We can also safely exclude the possibility of a SNe, given that deep optical observations of our cluster fields were obtained over the same two month time window as our deep near-IR observations (Postman et al. 2012; see also Z12 and C12).

One final possibility is that some candidates may correspond to more local solar system or Oort cloud objects. To be consistent with the constraints we can set on the proper motion of our candidates based on the ~2 month observational baseline we have (see Figure 6 of C12 for an illustration of the constraints we can set), such a source would need to be at 50,000 AU. However, at such distances, Oort cloud objects would be extremely faint (e.g., faintward of 40 mag), even if as these sources were as large as the moon (see also discussion in Z12 and C12).

4. A NEW DIFFERENTIAL DETERMINATION OF THE UV LUMINOSITY FUNCTION AT $Z \sim 9$

The present $z \sim 9$ sample is the largest such sample available to date and should allow us to substantially improve our constraints on the $z \sim 9$ luminosity function. However, before providing a detailed discussion of the specific constraints we are able to set, we must first include a few words on the procedure we adopt.

4.1. *UV LF Evolution from Lensing Cluster Searches: Rationale for Using a Differential Approach*

Normally, we would derive the luminosity function for $z \sim 9$ galaxies using the same approach that has been followed in the field, i.e., (1) distribute the sources in one's samples into different magnitude intervals, (2) count the total number of sources in a magnitude interval (after correcting for contamination) and (3) divide these numbers by the effective volume where such sources could be found. Such a procedure has been followed in a number of previous works (e.g., Santos et al. 2004; Richard et al. 2008).

However, even the simple process of placing sources into different intrinsic magnitude bins can be quite uncertain due to its dependence on a particular magnification model. Calculations of the selection volumes are just as equally model dependent. While in many cases these model dependencies may not result in large overall uncertainties in one's results, the uncertainties clearly do become large (~0.3-0.4 mag or larger) near the critical curves of the lensing models where the magnification factors become nominally infinite (e.g., see Figure 2 of Maizy et al. 2010). These issues can potentially have a huge effect on luminosity functions derived in the context of lensing clusters.[28]

[28] Of course, for realistic LFs, these uncertainties may not be especially problematic. Indeed, for LFs with an effective faint-end slope close to −2, uncertainties in the magnification factor trade off almost perfectly with uncertainties in the search volume so as to have no large effect on the inferred LFs. Because of this fact, one potentially very effective approach for minimizing the impact these uncertainties on the derived LFs is by marginalizing over the magnification factor in performing the comparisons with the observed numbers (C12). The excellent agreement between the present estimate of the SFR density at $z \sim 9$ and that obtained by

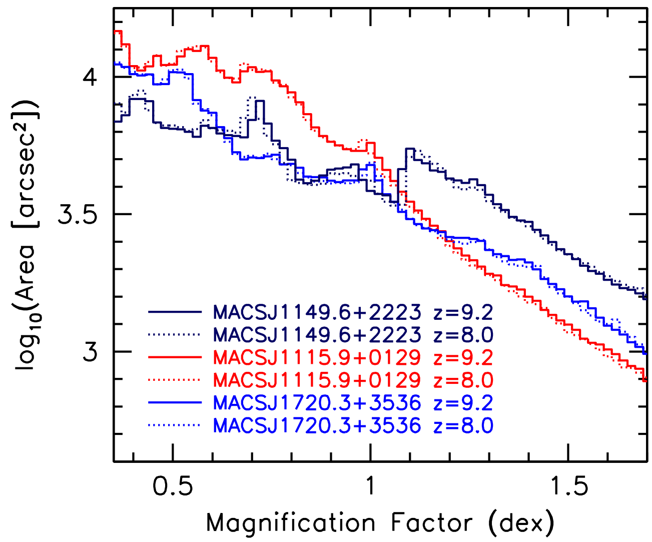

FIG. 8.— Search area (per unit dex) behind select galaxy clusters subject to varying levels of magnification by gravitational lensing. Results are shown for sources at $z = 8$ and $z = 9$ based on the lens models for MACSJ1149.6+2223, MACSJ1115.3+0129, and MACSJ1720.3+3536 (Z12; Zitrin et al. 2012, in prep; Carrasco et al. 2012, in prep). It is obvious from these results that the total search volume behind a cluster (given the area magnified to various levels) can show a huge variation from one cluster to another. However, if one utilises the same cluster to search for sources at similar but slightly different redshifts (*compare the dotted and solid lines representing $z \sim 8$ and $z \sim 9$ selections*), almost exactly the same selection area is available for selecting sources at a given magnification factor (although we remark that the selection area is slightly larger (~1-3%) at $z \sim 9$ than at $z \sim 8$). As a result, we would expect the relative selection volumes for a $z \sim 9$ search behind lensing clusters and a $z \sim 8$ search behind lensing clusters to be very well defined, *if the same clusters are utilized for the two searches.*

TABLE 3
ESTIMATED SCHECHTER PARAMETERS FOR THE UV LF AT $z \sim 9$ AND A COMPARISON WITH UV LF DETERMINATIONS AT OTHER REDSHIFTS $z \sim$ 4-10 (SEE §4.4).

Dropout Sample	Redshift	M^*_{UV} [a]	ϕ^* (10^{-3} Mpc^{-3})	α
$J_{110} + J_{125}$	9.2	−20.04 (fixed)	$0.11^{+0.16}_{-0.08}$	−2.06 (fixed)
B	3.8	-20.98 ± 0.10	1.3 ± 0.2	-1.73 ± 0.05
V	5.0	-20.64 ± 0.13	1.0 ± 0.3	-1.66 ± 0.09
i	5.9	-20.24 ± 0.19	$1.4^{+0.6}_{-0.4}$	-1.74 ± 0.16
z	6.8	-20.14 ± 0.26	$0.86^{+0.70}_{-0.39}$	-2.01 ± 0.21
Y	8.0	$-20.04^{+0.44}_{-0.48}$	$0.50^{+0.70}_{-0.33}$	$-2.06^{+0.35}_{-0.28}$

[a] Values of M^*_{UV} are at 1600 Å for the Bouwens et al. (2007, 2011b) $z \sim 4$, $z \sim 5$, and $z \sim 7$ LFs, at ~ 1350 Å for the Bouwens et al. (2007) $z \sim 6$ LF, and at ~ 1750 Å for the Oesch et al. (2012b) constraints on the $z \sim 8$ LF. Since z~6-8 galaxies are blue ($\beta \sim -2$: Stanway et al. 2005; Bouwens et al. 2006a, 2009, 2012b), we expect the value of M^*_{UV} to be very similar ($\lesssim 0.1$ mag) at 1600 Å to its value quoted here.

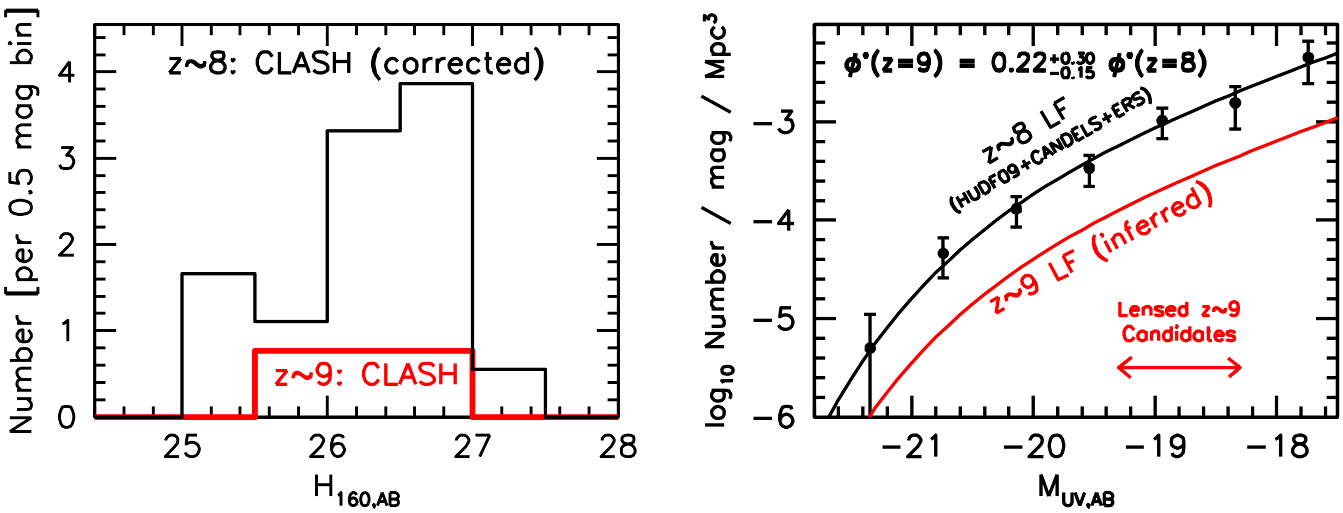

FIG. 9.— Illustration of our differential approach to deriving the UV LF at $z \sim 9$. (*left*) The contamination-corrected number of $z \sim 9$ galaxy candidates we find within CLASH (*red histogram*) vs. the number of $z \sim 8$ galaxy candidates (*black histogram*) behind the same CLASH clusters, corrected to have the same selection volume as at $z \sim 9$. For simplicity, the contamination rate correction is applied in a magnitude-independent manner (although the contamination rate will clearly be higher near the faint ends of our two samples). The number of $z \sim 9$ galaxy candidates in CLASH, after contamination correction, is just $0.22^{+0.30}_{-0.15}\times$ that at $z \sim 8$. A simple comparison of these surface densities should give us a fairly model independent measure of the relative normalization of the UV LF at $z \sim 8$ and the UV LF at $z \sim 9$ – assuming that the shape of the LF (i.e., M^* and α) does not change very dramatically from $z \sim 9$ to $z \sim 8$. (*right*) The observed UV LF at $z \sim 8$ as derived by Oesch et al. (2012b: *black points, error bars, and line*) based on the HUDF09+CANDELS+ERS data set and our newly inferred UV LF at $z \sim 9$ (*red line*) based on our differential comparison of our $z \sim 8$ and $z \sim 9$ selections. We infer that the UV LF at $z \sim 9$ has an effective ϕ^* that is just $0.22^{+0.30}_{-0.15}\times$ that at $z \sim 8$. The red horizontal arrow towards the bottom of this panel indicates the approximate luminosities inferred for our 3 $z \sim 9$ candidates (after correction for lensing magnification: see §3.2).

Ideally, we would like to determine the UV LF at $z \sim 9$ in a way that avoids these uncertainties. One possible way for us to do this is (1) to leverage existing well-determined LFs that already exist at $z \sim$ 7-8 from blank field studies (e.g., Bouwens et al. 2011b; Oesch et al. 2012b; Bradley et al. 2012) not subject to potentially large selection volume uncertainties and (2) then to use our searches for $z \sim$ 7-10 galaxies behind lensing clusters to derive the differential evolution in the LF from $z \sim 9$ to $z \sim$ 7-8. This provides us with a somewhat indirect approach to deriving the LF at $z \sim 9$ and takes advantage of the very similar effect gravitational lensing from low-redshift clusters has on light from the high-redshift universe, regardless of the exact redshift of the source. Fundamentally, this is due to the fact that the D_{LS}/D_S factor is very insensitive to redshift when the lensed source is at $z > 5$ (i.e., very distant) and the lensing cluster is relatively close (i.e., $z \sim$ 0.1-0.5). For example, for a $z \sim 0.4$ lensing cluster, the computed D_{LS}/D_S factor for $z \sim 9$ background sources is only ∼1% higher than the D_{LS}/D_S factor for $z \sim 8$ sources. D_{LS} and D_S are the angular-diameter distances from the cluster lens to source and from observer to source, respectively (e.g., Narayan & Bartelmann 1996).

As a result, for sources seen behind a given lensing cluster, the $z \sim 8$ universe is magnified in almost exactly the same way as the $z \sim 9$ universe. This can be illustated using the lensing models we have available for three of the CLASH clusters (Figure 8). The total area available behind a given cluster to magnify the background light by more than a factor of 3 is almost exactly the same for the $z \sim 8$ universe as for the $z \sim 9$ universe. Note that this is true, even if the precise position of the critical curves at $z \sim 9$ lies in a slightly different position from the critical curves at e.g. $z \sim 8$.

Because of the very similar way a given set of clusters magnifies galaxies at $z \sim 9$ and at other similar redshifts (e.g., $z \sim 8$), one might expect it to have the same effect on the total surface densities of these galaxies one finds on the sky. Therefore, if one starts with the same luminosity function of galaxies at both $z \sim 8$ and $z \sim 9$, one would expect to find roughly the same surface density of these galaxies on the sky, modulo two slight differences. The $z \sim 9$ galaxy distribution would be shifted to slightly fainter magnitudes (e.g., by ∼0.3 mag versus $z \sim 8$) to reflect their slightly larger luminosity distances and would be present at slightly lower surface densities (by ∼10% versus $z \sim 8$) reflecting the smaller cosmic volume available at $z \sim 9$.

Even multiple imaging of the same high-redshift sources would not appreciably affect the ratio of sources seen at different redshifts, since one would expect galaxies at $z \sim 9$ and similar redshifts to give rise to lensed multiplets to approximately the same degree, and therefore the ratio of surface densities should be preserved. However, since multiple images of a single background source are not independent events, not accounting for this effect could have a slight effect on the uncertainties we estimate for the relative surface densities of galaxies at different redshifts.

Given this situation, it seems quite clear we should be able to use the relative surface densities of galaxies in different redshift samples to make reasonably reliable inferences about the relative volume densities of the galaxy

C12 based on the Z12 search results would seem to support this conclusion.

population at different epochs (after making small adjustments to the numbers to account for the factors discussed above).

4.2. $z \sim 8$ *Comparison Sample*

Redshift $z \sim 8$ selections serve as the perfect comparison sample for our $z \sim 9$ studies. Not only is the $z \sim 8$ universe close enough to $z \sim 9$ to make differences in the lensing effects quite small overall, but the ∼70-80 $z \sim 8$ galaxies available in current WFC3/IR surveys allow the LF there to be robustly established from field studies (e.g., Bouwens et al. 2011b; Lorenzoni et al. 2011; Oesch et al. 2012b; Bradley et al. 2012). This allows us to put together the new information we have on the differential evolution of the LF from $z \sim 9$ to $z \sim 8$ with previous $z \sim 8$ LF determinations to estimate the approximate UV LF at $z \sim 9$.

Finally, given the observed rate of evolution in M^* and α (e.g., Bouwens et al. 2012a), we would expect the shape of the LF at $z \sim 8$ to be similar to the shape of the LF at $z \sim 9$, i.e., $\Delta(M^*(z = 8) - M^*(z = 9)) \lesssim 0.4$ and $\Delta\alpha(z = 8) - \alpha(z = 9) \lesssim 0.06$, so we can model any evolution in the LF very simply assuming a change in the normalization ϕ^* (though modeling the evolution in terms of the characteristic luminosity M^* is only slightly more involved).

For our $z \sim 8$ comparison sample, we use the same selection criteria as previously utilized in Bouwens et al. (2011b) and Oesch et al. (2012b), i.e.,

$$(Y_{105} - J_{125} > 0.45) \wedge (J_{125} - H_{160} < 0.5)$$

As in these two previous works (and as performed for our $z \sim 9$-10 selection), we also require sources to be undetected in the I_{814} band and blueward both in individual bands at $<2\sigma$ and using the χ^2_{opt} statistic discussed earlier (§3.2). We also demand that sources be detected at $> 6\sigma$ in a combined JH_{140} and H_{160} image (0.35″-diameter aperture), as performed for our primary $z \sim 9$-10 selection. Since these color criteria and selection criteria are very similar to that used by Bouwens et al. (2011b) and Oesch et al. (2012b) in identifying $z \sim 8$ galaxies, the redshift distribution for the present $z \sim 8$ selection should be approximately the same as shown in Figure 1 (*red line*).

Applying this selection criteria to the 19-cluster CLASH dataset, we find a total of 19 sources which satisfy our $z \sim 8$ criteria. These sources have $H_{160,AB}$ magnitudes ranging from 25.0 to 27.3 mag. We allow for a potential contamination of ∼1 source in our $z \sim 8$ sample, consistent with the contamination level found by Bouwens et al. (2011b) for their $z \sim 8$ sample.[29]

4.3. *Relative selection volumes at* $z \sim 8$ *and* $z \sim 9$*: Expected sample sizes assuming no evolution*

In order to utilize the relative surface density of $z \sim 8$ and $z \sim 9$ galaxy candidates we observe to make inferences about the evolution of the luminosity function, we must have an estimate for how many galaxies we would expect in the two samples if the UV LF did not evolve at all between the two epochs. Then, based on the relative number of sources expected in the two samples assuming no evolution, we can determine the approximate evolution in the LF from $z \sim 9$ to $z \sim 8$. With this step, we effectively account for the approximate difference in selection volume for our $z \sim 8$ and $z \sim 9$ samples.

The simplest way for us to account for any evolution in the UV LF is through the normalization ϕ^*[30] – since it simply requires that we compare the number of sources we find in our $z \sim 8$ and $z \sim 9$ samples with that found in our simulations (see below) to derive the approximate evolution, i.e.,

$$\phi^*(z = 9) = \phi^*(z = 8) \frac{n_{obs,z=9}}{n_{obs,z=8}} \frac{n_{no-evol-sim,z=8}}{n_{no-evol-sim,z=9}} \quad (1)$$

where $n_{obs,z=9}$ is the number of sources in our $z \sim 9$ selection after correction for contamination (i.e., ∼2.3), $n_{obs,z=8}$ is the number of sources in our $z \sim 8$ selection after correction for contamination (i.e., ∼18), $n_{no-evol-sim,z=8}$ is the number of $z \sim 8$ candidates we find in our simulations for our $z \sim 8$ selection based on a fiducial lensed LF, and $n_{no-evol-sim,z=9}$ is the number of $z \sim 9$ candidates we find in our simulations for our $z \sim 9$ selection based on this same LF.

As in our previous papers, we estimate these selection volumes directly from simulations. To perform these simulations, we insert artificial sources with a variety of redshifts and luminosities into the real observations and then attempt to select these objects using our $z \sim 8$ and our $z \sim 9$-10 selection criteria. We generate artificial images for each source in these simulations using our well-tested cloning software (Bouwens et al. 1998; Bouwens et al. 2003; Bouwens et al. 2007) which we use to artificially redshift similar luminosity $z \sim 4$ galaxies from the HUDF to higher redshift. We scale the size of galaxies at fixed luminosity as $(1 + z)^{-1}$ to match the observed size-redshift scaling at $z > 3$ (e.g., Bouwens et al. 2004; Oesch et al. 2010; Mosleh et al. 2012). We take the UV-continuum slope β of galaxies in our simulations to have a mean value and 1σ scatter of -2.3 and 0.45, respectively, to match the observed trends extrapolated to $z \sim 8$-9 (Bouwens et al. 2012b; Finkelstein et al. 2012).

For simplicity, we run the above simulations without incorporating the deflection maps derived for all 19 CLASH clusters we are considering in the present search. In general, we would not expect any sizeable changes in their surface brightnesses as a result of the lensing magnification, and therefore the results of our simulations should be roughly valid (though the surface brightnesses will change somewhat for unresolved sources in our fields and there will be differences in the shapes and sizes of sources in our fields). Fortunately, any inaccuracies in this approximation should not have a large effect on our final result since it would affect the selection volume results at both $z \sim 8$ and $z \sim 9$ and we are interested in

[29] We emphasize that the current $z \sim 8$ selection is derived independently from the $z \sim 8$ sample that will be reported in the upcoming $z \sim 6$-9 sample paper from CLASH (L. Bradley et al. 2012, in preparation) though obviously there is substantial overlap in the sources which compose the two selections.

[30] While we have previously advocated parameterizing the evolution of the UV LF in terms of the characteristic luminosity M^* (e.g., Bouwens et al. 2007; Bouwens et al. 2008), the shape of the $z \sim 7$-9 UV LFs appears to be such that evolution in ϕ^* and M^* is almost entirely degenerate, particularly over a limited baseline in redshift (e.g., see Figure 8 from Oesch et al. 2012b). For convenience, we parameterize the evolution of the LF in terms of ϕ^* here.

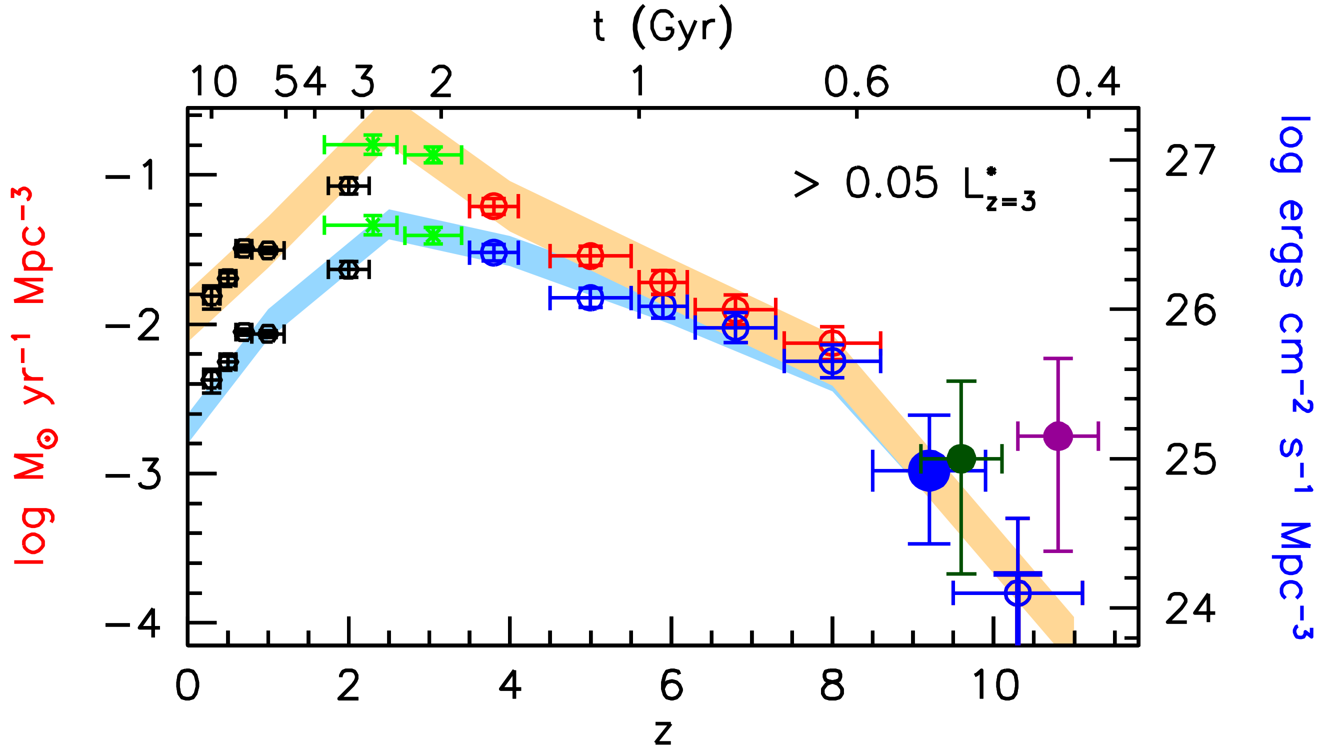

FIG. 10.— The UV luminosity density (*right axis*) and star formation rate density (*left axis*) versus redshift. The UV luminosity and SFR density shown at $z \sim 9$ (*large blue solid circle*) are from the present work and inferred based on the relative number of $z \sim 8$ and $z \sim 9$ galaxies found within the CLASH cluster program (see §4.5). These luminosity densities and SFR densities are only considered down to a limiting luminosity of −17.7 AB mag – which is the approximate limit of both the HUDF09 probe (Bouwens et al. 2011b) and the present search assuming a maximum typical magnification factor of ∼9 and limiting magnitude of ∼27.0 mag. The UV luminosity is converted into a star formation rate using the canonical UV-to-SFR conversion factors (Madau et al. 1998; Kennicutt 1998). The upper set of points at every given redshift and orange contour show the dust-corrected SFR densities, while the lower set of points and blue contours show the inferred SFR densities before dust correction. Dust corrections at $z > 3$ are estimated based on the observed UV-continuum slope distribution and are taken from Bouwens et al. (2012b). At $z \leq 3$, the dust corrections are from Schiminovich et al. (2005) and Reddy & Steidel (2009). UV luminosity density and SFR density determinations from the literature are from Schiminovich et al. (2005) at $z < 2$ (*black hexagons*), Reddy & Steidel (2009) at $z \sim$ 2-3 (*green crosses*, Bouwens et al. (2007) at $z \sim$ 4-6 (*open red and blue circles*), Bouwens et al. (2011b) at $z \sim 7$ (*open red and blue circles*), Oesch et al. (2012b) at $z \sim 8$ (*open red and blue circles*), and Oesch et al. (2012a) at $z \sim 10$ (*open blue circle and upper limit*). Estimates of the SFR density at $z \sim 9.6$ and $z \sim 10.8$ as derived in C12 based on the $z \sim 9.6$ Z12 and $z \sim 10.8$ C12 candidates are also shown (*dark green and magenta solid circles, respectively*). Conversion to a Chabrier (2003) IMF would result in a factor of ∼1.8 (0.25 dex) decrease in the SFR density estimates given here. The present $z \sim 9$ determination is in good agreement with the trend in the SFR density and UV luminosity, as defined by the Oesch et al. (2012a) and Z12 estimates.

the ratio of the selection volumes at the two redshifts.

The UV LFs we input into the simulations have the following parameters: $M^*_{UV} = -22.4$, $\alpha = -2.0$, and $\phi^* = 5.5 \times 10^{-5}$ Mpc^{-3}. These luminosity parameters were chosen to implicitly include a factor of ∼9 magnification from gravitational lensing – which is the median magnification estimated for sources in our selection – so the effective M^* at $z \sim 8$ is chosen to be ∼2.4 mag brighter than seen in blank field studies (e.g., Oesch et al. 2012b). The faint-end slope we assume approximately matches what we would expect based on the UV LF results at $z \sim$ 7-8 (Bouwens et al. 2011b; Oesch et al. 2012b; Bradley et al. 2012) which point to faint-end slopes α of −2. No change is required in the faint-end slope α of the LF, due to the perfect trade-off between magnification and source dilution effects for slopes of −2 (e.g., Broadhurst et al. 1995). The normalization ϕ^* we choose has no effect on our final results (due to the differential nature of this calculation). While the LFs we adopt for these simulations could, in principle, affect our evolutionary results, the overall size of such effects will be small due to the differential nature of the comparison we are making. We also verified that the surface density of $z \sim 8$ sources predicted by this model LF showed a very similar magnitude dependence as seen for our $z \sim 8$ sample.

Using the above simulation procedure and aforementioned LF, we repeatedly added artificial sources to the real CLASH observations for all 19 CLASH clusters, created catalogs, and repeated our $z \sim 8$ and $z \sim 9$ selections. In total, we repeated the described simulations 20 times for each cluster field to obtain an accurate estimate of the total number of sources (selection volume and redshift distribution) we would expect to find in each sample, given the described luminosity function.

In total, we find 657 sources that satisfy our $z \sim 8$ selection criteria and 383 sources that satisfy our $z \sim 9$ selection criteria, based on the same luminosity and simulation area (so $n_{no-evol-sim,z=8} = 657$ and $n_{no-evol-sim,z=9} = 383$ in Eq. 1 above). This suggests that the effective volume for our $z \sim 9$ selection is just 58% as large as it is for our $z \sim 8$ selection, and therefore to make a fair comparison between our $z \sim 8$ and $z \sim 9$ samples we need to multiply the surface densities in our $z \sim 8$ sample by 0.58 (Eq. 1 above). In Figure 9, we show the comparison of the surface densities of $z \sim 8$ galaxies

TABLE 4
UV Luminosity Densities and Star Formation Rate Densities to -17.7 AB mag ($0.05\ L^*_{z=3}$: see §4.5).[a]

Dropout Sample	$<z>$	$\log_{10}\mathcal{L}$ (ergs s^{-1} Hz^{-1} Mpc^{-3})	$\log_{10}$ SFR density ($M_\odot$ Mpc^{-3} yr^{-1}) Uncorrected	Corrected[b]
J	9.2	$24.92^{+0.37}_{-0.49}$	$-2.98^{+0.37}_{-0.49}$	$-2.98^{+0.37}_{-0.49}$
B	3.8	26.38±0.05	−1.52±0.05	−1.21 ± 0.05
V	5.0	26.08±0.06	−1.82±0.06	−1.54 ± 0.06
i	5.9	26.02±0.08	−1.88±0.08	−1.72 ± 0.08
z	6.8	25.88±0.10	−2.02±0.10	−1.90 ± 0.10
Y	8.0	25.58±0.11	−2.32±0.11	−2.20 ± 0.11
J[d]	10.3	$24.1^{+0.5}_{-0.7}$	$-3.8^{+0.5}_{-0.7}$	$-3.8^{+0.5}_{-0.7}$
J[d]	10.3	<24.2[c]	< -3.7[c]	< -3.7[c]

[a] Integrated down to $0.05\ L^*_{z=3}$. Based upon the $z\sim9$ inferred here (Table 3: §4.4) and the LF parameters in Oesch et al. (2012a,b) and Table 2 of Bouwens et al. (2011b) (see §4.5). The SFR density estimates assume $\gtrsim 100$ Myr constant SFR and a Salpeter IMF (e.g., Madau et al. 1998). Conversion to a Chabrier (2003) IMF would result in a factor of ∼1.8 (0.25 dex) decrease in the SFR density estimates given here.
[b] Dust corrections are from Bouwens et al. (2012b) and are based on the observed UV-continuum slopes. No dust correction is assumed at $z \gtrsim 9$.
[c] Upper limits here are 1σ (68% confidence).
[d] $z\sim10$ determinations and limits are from Oesch et al. (2012a: see also Bouwens et al. 2011a) and assume 0.8 $z\sim10$ candidates in the first case and no $z\sim10$ candidates (i.e., an upper limit) in the second case.

found over the first 19 CLASH clusters (corrected for the difference in selection volume) with the surface densities of $z\sim9$ galaxies found over these clusters.

4.4. *Inferred UV LF at $z\sim9$*

After correction for possible contamination of our selection by possible low redshift contaminants (see §3.5), the total number of $z\sim9$ candidates in our $z\sim9$ selection is 2.3. This number is just $0.22^{+0.30}_{-0.15}\times$ the total number of $z\sim8$ sources to a similar luminosity limit (corrected for differences in the selection volume: see Figure 9). In calculating the uncertainties on the fraction $0.22^{+0.30}_{-0.15}$, we have accounted for the Poissonian errors on the total number of galaxies in the $z\sim8$ and $z\sim9$ samples, as well as the Poissonian uncertainties in the contamination rates.

Assuming that we can approximate the differences between the $z\sim8$ and $z\sim9$ LFs as occurring simply through density evolution (i.e., by changing ϕ^*), we infer that the value of ϕ^* at $z\sim9$ is just $0.22^{+0.30}_{-0.15}\times\ \phi^*$ at $z\sim8$ (or $0.22^{+0.65}_{-0.20}\times$ if 2σ errors are used). The present search is inconsistent with no evolution at >98.4% confidence.[31]

Using the recent Oesch et al. (2012b) determination from HUDF09+CANDELS+ERS field studies (Bouwens et al. 2011b; Grogin et al. 2011; Koekemoer et al. 2011; Windhorst et al. 2011) that ϕ^* at $z\sim8$ is $5.0^{+7.0}_{-3.3}\times10^{-4}$ Mpc^{-3}, we estimate that ϕ^* at $z\sim9$ is $1.1^{+1.6}_{-0.8}\times10^{-4}$ Mpc^{-3}. For the purpose of parametrizing a $z\sim9$ LF, we will assume that M^* and α at $z\sim9$ match that derived by Oesch et al. (2012b) at $z\sim8$. The resultant $z\sim9$ LF is illustrated in the right panel of Figure 9 and compared with the Oesch et al. (2012b) $z\sim8$ LF. Use of the Bradley et al. (2012) $z\sim8$ LF results (where the bright end of the LF is based on the wide-area [∼274 arcmin^2] BORG results: Trenti et al. 2011) instead of the Oesch et al. (2012) results would yield an effective ϕ^* that is just 7% higher.

What effect will field-to-field variations (i.e., "cosmic variance") have on the overall uncertainties here? To estimate the size of these uncertainties, we first considered the case of a single cluster field. We used the Trenti & Stiavelli (2008) cosmic variance calculator, assumed a mean redshift of 8.0 and 9.2 for our two samples (as estimated from our simulations: see Figure 1), took the Δz width for these redshift distributions to be 0.8, and assumed that the relevant area in the source plane was $0.4'\times0.4'$. The latter area in the source plane assumes a factor of ∼10 dilution of the total search area (consistent with the mean magnification factors found here) and further that only 30% of the total area on our WFC3/IR images is effective for finding $z\sim9$ galaxies. The fractional uncertainty we estimated in our volume density estimates from field-to-field variations is 0.55 and 0.58 for our $z\sim8$ and $z\sim9$ selections, respectively, over a single CLASH cluster field.

Since each of our cluster fields provides an independent sightline on the high redshift universe, we need to reduce the derived variance by $\sim19^{0.5}\sim4.4$ (though we remark that the actual reduction will be slightly less than this since all our clusters will not receive equal weight in the total volume calculation and hence the gains from independent sightlines will be less). This results in fractional uncertainties of ∼0.13 in the total number of sources in the current $z\sim8$ and $z\sim9$ samples. Since our $z\sim9$

[31] Given the approximate degeneracy between evolution in M^* and ϕ^* for LFs at $z\sim$ 7-9 where a $\Delta M^*=1$ mag change trades off for a $\Delta\phi^*$ change (e.g., Figure 8 of Oesch et al. 2012b), we can reframe the inferred evolution in ϕ^* from $z\sim9$ to $z\sim8$ in terms of an equivalent evolution in M^* (as we have parameterized the LF evolution in the past, e.g., Bouwens et al. 2007; Bouwens et al. 2008; C12). We estimate that the effective M^* at $z\sim9$ is $0.7^{+0.5}_{-0.4}$ mag fainter than at $z\sim8$ (keeping ϕ^* fixed).

LF estimate is based on a differential comparison of the present 19-cluster $z \sim 8$ and $z \sim 9$ samples, we must add both of these uncertainties in quadrature to derive the approximate fractional uncertainty. The result is 0.19. By comparison, $z \sim 9$ searches using a single 4.4 arcmin^2 deep field would yield a fractional uncertainty of $\sim$0.5 in the volume density of $z \sim 9$ galaxies from large-scale structure ("cosmic variance"). This is much higher than the present uncertainties arising from large-scale structure.

Overall, the uncertainties from large-scale structure only have a fairly marginal impact on our total uncertainty in ϕ^* for the $z \sim 9$ LF, increasing it by just 3% over what one would estimate based on the small numbers in the current $z \sim 9$ selection.

4.5. *UV Luminosity and Star Formation Rate Density at $z \sim 9$*

We can utilize our newly estimated $z \sim 9$ LF to determine the approximate UV luminosity density and SFR density at $z \sim$ 9-10. We compute these luminosity densities to a limiting luminosity 0.05 $L^*_{z=3}$, which is the effective limit of the Oesch et al. (2012b) $z \sim 8$ LF we used as a reference point for inferring the $z \sim 9$ LF. This limiting luminosity is also what one would expect for $z \sim$ 8-9 searches in the CLASH program to $\sim$27 AB mag assuming a $\sim 9\times$ magnification factor – which is equivalent to the average magnification factor for $z \sim 9$ galaxy candidates uncovered in the present search. We can convert the UV luminosities we estimate to SFR densities using the canonical UV luminosity-to-SFR conversion factor (Madau et al. 1998: see also Kennicutt 1998).

The $z \leq 8$ SFR density determinations are corrected for dust extinction based on the values Bouwens et al. (2012b) estimate based on the observed UV-continuum slopes β. Given the observed trends towards bluer UV-continuum slopes β at very high redshifts (e.g., Bouwens et al. 2012b; Stanway et al. 2005; Finkelstein et al. 2012; Wilkins et al. 2011), we would expect the dust extinction at $z \sim$ 9-10 to be zero, and therefore apply no dust correction to the SFR density determinations there.

We present the UV luminosity and SFR densities we estimate at $z \sim 9$ in Figure 10 and also in Table 4. For context, we also provide the SFR and UV luminosity densities of several noteworthy determinations in the literature over the redshift range $z \sim 0$ to $z \sim 10$ (Bouwens et al. 2007, 2010; Oesch et al. 2012a). We also show the SFR density estimates at $z \sim 9.6$ and $z \sim 10.8$ from the $z \sim 9.6$ Z12 and $z \sim 10.8$ C12 candidates, as estimated by C12.

4.6. *Implications of the present $z \sim$ 9-10 search for the evolution of the LF at $z > 6$*

One of our primary motivations for obtaining constraints on the UV LF at $z \sim 9$ was to characterize the evolution of the UV LF at $z > 8$ and to test whether the UV LF at $z > 8$ really evolves more rapidly as a function of redshift – as recently found by Oesch et al. (2012a: see also Bouwens et al. 2012a) – or the evolution is more consistent with a simple extrapolation of the UV LF trends found by Bouwens et al. (2011b: see also Bouwens et al. 2008) over the redshift range $z \sim$ 4-8. Several theoretical models (Trenti et al. 2010; Lacey et al. 2011) support the idea that the UV LF should evolve faster at $z > 8$ as a function of redshift than at $z \sim$ 4-8 (Figure 8 of Oesch et al. 2012a), and we want to test this hypothesis using our current results.

To determine which of these two scenarios the present observations favor, we first compute the change in ϕ^* each would predict. Using the Bouwens et al. (2011b) fitting formula for the evolution of the UV LF, we estimate an expected change of $\Delta M^*_{UV} \sim 0.4$ in the UV LF from $z \sim 8$ to $z \sim 9.2$ (the mean redshift of our sample). Taking advantage of the approximate degeneracy between M^* and ϕ^* at $z \sim 7$ and $z \sim 8$ ($\Delta M^* = 1$ is nearly degenerate with $\Delta \log_{10} \phi^* = 1$: see Figure 8 of Oesch et al. 2012b), we can convert this to a change in ϕ^* over the redshift interval $z \sim 8$ to $z \sim 9.2$, i.e., $\Delta \log_{10} \phi^* \sim 0.4$ dex so that $\phi^*(z = 9.2) = 0.4\phi^*(z = 8)$. Oesch et al. (2012a) also estimate the rate of evolution from $z \sim 10$ to $z \sim 8$ based on their $z \sim 10$ HUDF09+ERS+CANDELS search results, which is more rapid than implied by the Bouwens et al. (2011b) fitting formula. The best-fit evolution in ϕ^* that Oesch et al. (2012a) find is a $0.54^{+0.36}_{-0.19}$ dex change per unit redshift, so that $\phi^*(z = 9.2) = 0.23^{+0.15}_{-0.15}\phi^*(z = 8)$.

The evolution we measure from $z \sim 9.2$ to $z \sim 8$ (§4.4) is such that $\phi^*(z = 9.2) = 0.22^{+0.30}_{-0.15}\phi^*(z = 8)$ (fixed M^* and α). As compared with the two different evolutionary scenarios, we can see that the observed evolution shows a preference for a more rapid evolution in the UV LF at $z > 8$, as found by Oesch et al. (2012a), but is still nevertheless quite consistent at $\sim 0.6\sigma$ with the rate of evolution seen in the $z \sim$ 4-8 LFs (versus one would expect utilizing the Bouwens et al. 2011b LF fitting formula where $dM^*/dz \sim 0.33$). More information on the prevalence of galaxies at $z \sim 9$ – and the evolution of the LF at $z > 8$ should soon be available soon using the new ultra-deep WFC3/IR observations over the HUDF/XDF field and also future deep observations over clusters (from CLASH and other programs).

In any case, it appears that searches for galaxies at even higher redshifts (i.e., $z \gtrsim 10$) may be required to provide the necessary leverage to determine the rate at which the LF really evolves at very early times. Again, while the recent HUDF09 $z \sim 10$ search results (Bouwens et al. 2012a; Oesch et al. 2012a) argue for a more rapid evolution in the UV LF at $z > 8$ (based on the small number of plausible $z \sim 10$ candidates found in the HUDF09 program), the recent discovery of a bright, multiply-lensed $z \sim 10.8$ galaxy found in the CLASH program (C12) seems more consistent with less rapid evolution.

5. SUMMARY

In this paper, we have explored the use of a two-color Lyman-Break selection to search for $z \sim$ 9-10 galaxies in the 19-cluster CLASH data set. Building on the important exploratory studies of Z12 and C12, we extend the CLASH $z \sim$ 9-10 selections even deeper to the approximate magnitude limit of the CLASH program ($\sim$27 mag). Such a search is possible making full use of the noteworthy Spitzer/IRAC observations over the CLASH clusters (Egami et al. 2008; Bouwens et al. 2011c), allowing us to determine which $z \sim$ 9-10 galaxy candidates have a blue spectral slope redward of the break (and therefore strongly favor a $z \sim 9$-10 solution) and which candidates do not.

In total, we find three plausible $z \sim 9$-10 galaxy candidates over the 19-cluster CLASH data set that satisfy a two-color Lyman-Break-like selection criteria (i.e., $(J_{110} + J_{125})/2 - H_{160} > 0.7$ and $JH_{140} - H_{160} < 0.5$) and have a combined $JH_{140} + H_{160}$ S/N of $\geq$6.0. The $H_{160,AB}$ magnitudes for sources in our selection range from $\sim$25.7 AB mag to 26.9 AB mag. The candidates are found behind the galaxy clusters MACSJ1149.6+2223, MACSJ1115.9+0129, and MACSJ1720.3+3536. The highest S/N source in our selection is the $z \sim 9.6$ Z12 candidate (here $z_{ph} \sim 9.7$). All three of our candidates have reasonably blue $H_{160,AB} - IRAC$ colors strongly favoring the $z \sim 9$-10 solution for all three sources in our selection.

As with our previous $z \sim 9$ and $z \sim 10$ selections (Bouwens et al. 2011a; Z12; C12), we have carefully considered the possibility of contamination. We find that the only significant source of contamination is from the "photometric scatter" of lower redshift galaxies into our selection and that this likely contributes only $\sim$0.7 source to our $z \sim 9$-10 sample (§3.5). However, we emphasize that we cannot completely exclude the possibility that the contamination rate may be somewhat higher.

To determine the implications of the present search results for the UV LF, UV luminosity density, and SFR density at $z \sim 9$-10, we introduce a novel differential approach for deriving these quantities. Our procedure is to simply compare the number of candidate $z \sim 9$ galaxies found in the CLASH fields with the number of $z \sim 8$ galaxies found in the CLASH fields and then correct this ratio for the relative selection volume at $z \sim 8$ and $z \sim 9$. This procedure takes advantage of the fact that the ratio of selection volumes at $z \sim 8$ and $z \sim 9$ for a given cluster is not greatly dependent on details of the gravitational lensing model one is utilizing (e.g., see Figure 8). This procedure therefore provides us with a very robust technique for measuring the evolution of the UV LF to $z > 9$ using searches over lensing cluster fields. The $z \sim 8$ and $z \sim 9$ selection volumes we derive are from detailed simulations where artificial sources are added to the real imaging data and then reselected using the same criteria as applied to the real data (§4.3).

Comparing our sample of three candidate $z \sim 9$ galaxies with a sample of 19 $z \sim 8$ galaxies found to similar 6σ detection significance over the same CLASH cluster fields (and correcting for the expected 23% contamination in our $z \sim 9$ selection), we derive the approximate evolution in the UV LF to $z \sim 9$. One strength of the present evolutionary estimate is that we are particularly insensitive to large-scale structure uncertainties due to our many independent lines of sight on the high redshift universe (§4.4).

We find that ϕ^* for the $z \sim 9$ LF is just $0.22^{+0.30}_{-0.15}\times$ the equivalent ϕ^* at $z \sim 8$ (§4.4: keeping M^* and α fixed).[32] We would have expected the normalization ϕ^* of the LF at $z \sim 9$ to be just $0.4\times$ that at $z \sim 8$, if the evolution in the UV LF proceeded at the same rate as seen at $z \sim 4$-8. The present result is therefore consistent with the idea that the UV LF may evolve more rapidly at $z > 8$, similar to what Oesch et al. (2012a) found based on a search for $z \sim 10$ galaxies within the HUDF09+ERS+CANDELS observations. Using the Oesch et al. (2012a) interpolation between the $z \sim 8$ and $z \sim 10$ results, we would have predicted the normalization of our $z \sim 9$ LF to be $0.23^{+0.15}_{-0.15}\times$ that at $z \sim 8$. Several theoretical models (Trenti et al. 2010; Lacey et al. 2011) support the idea that the UV LF should evolve faster at $z > 8$ as a function of redshift than at $z \sim 4$-8 (Figure 8 of Oesch et al. 2012a). Despite the excellent agreement between the present evolutionary result and that expected from the Oesch et al. (2012a) study, the uncertainties on this result are still somewhat large.

In the future, we expect further advances in our constraints on the UV LF at $z \geq 9$ based on continuing observations from the CLASH program (observations over 6 CLASH clusters still remain), from deeper observations over the HUDF/XDF (GO 12498: PI Ellis), and from the STScI Deep Field Initiative. Substantially deeper Spitzer observations over the CLASH clusters, as part of the Surf's Up program (Bradac et al. 2012) and other programs, should allow us both to obtain better constraints on the nature of current $z \sim 9$ candidates and to provide an initial estimate of the stellar mass density at $z \sim 9$.

We are grateful for extensive feedback on our manuscript from Pascal Oesch and Garth Illingworth. We thank Anton Koekemoer for providing us with high quality reductions of the available HST observations of our CLASH cluster fields. We acknowledge Dan Magee for assisting with the reductions of the IRAC data for our clusters using MOPEX. We are greatly appreciative to Ryan Quadri for sending us his results for the rest-frame V band LF of red $z \sim 2$ galaxies based on a UKIDSS UDS search. We acknowledge support from ERC grant HIGHZ #227749, an NWO vrij competitie grant, and the NASA grant for the CLASH MCT program. The dark cosmology center is funded by the DNRF.

[32] We can also express the observed evolution in the UV LF from $z \sim 9$ to $z \sim 8$ in terms of M^* (keeping ϕ^* fixed). For a fixed ϕ^*, the approximate value of M^* at $z \sim 9$ is $0.7^{+0.5}_{-0.4}$ mag fainter than M^* at $z \sim 8$: see §4.4).

REFERENCES

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Benítez, N. 2000, ApJ, 536, 571
Bertin, E. and Arnouts, S. 1996, A&AS, 117, 39
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Bouwens, R., Broadhurst, T. and Silk, J. 1998, ApJ, 506, 557
Bouwens, R., Broadhurst, T., & Illingworth, G. 2003, ApJ, 593, 640
Bouwens, R. J., Illingworth, G. D., Rosati, P., et al. 2003, ApJ, 595, 589
Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J., & Franx, M. 2004, ApJ, 611, L1
Bouwens, R. J., Thompson, R. I., Illingworth, G. D., et al. 2004, ApJ, 616, L79
Bouwens, R.J., Illingworth, G.D., Blakeslee, J.P., & Franx, M. 2006a, ApJ, 653, 53
Bouwens, R. J., & Illingworth, G. D. 2006b, Nature, 443, 189

Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2008, ApJ, 686, 230
Bouwens, R. J., et al. 2009, ApJ, 705, 936
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, ApJ, 709, L133
Bouwens, R. J., Illingworth, G. D., Labbé, I., et al. 2011a, Nature, 469, 504
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011b, ApJ, 737, 90
Bouwens, R., Zheng, W., Moustakas, L., et al. 2011c, Spitzer Proposal, 80168
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, ApJ, 752, L5
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, ApJ, 754, 83
Bradac, M., Gonzalez, A., Schrabback, T., et al. 2012, Spitzer Proposal, 90009
Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bunker, A. J., Stanway, E. R., Ellis, R. S., McMahon, R. G., & McCarthy, P. J. 2003, MNRAS, 342, L47
Bunker, A. J., Wilkins, S., Ellis, R. S., et al. 2010, MNRAS, 409, 855
Casertano, S., de Mello, D., Dickinson, M., et al. 2000, AJ, 120, 2747
Chabrier, G. 2003, PASP, 115, 763
Cimatti, A., Daddi, E., Mignoli, M., et al. 2002, A&A, 381, L68
Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926
Coe, D., Zitrin, A., Carrasco, M., et al. 2012, ApJ, in press (C12)
Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Cushing, M.C., Rayner, J.T., & Vacca, W.D. 2005, ApJ, 623, 1115
Dickinson, M., Hanley, C., Elston, R., et al. 2000, ApJ, 531, 624
Dickinson, M., Stern, D., Giavalisco, M., et al. 2004, ApJ, 600, L99
Dow-Hygelund, C. C., et al. 2007, ApJ, 660, 478
Egami, E., Ellis, R., Fazio, G., et al. 2008, Spitzer Proposal, 60034
Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10
Ferguson, H. C. et al. 2004, ApJ, 600, L107
Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, ApJ, 756, 164
Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950
Fontana, A., Vanzella, E., Pentericci, L., et al. 2010, ApJ, 725, L205
Giallongo, E., Salimbeni, S., Menci, N., et al. 2005, ApJ, 622, 116
Grazian, A., Fontana, A., de Santis, C., et al. 2006, A&A, 449, 951
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Iye, M., Ota, K., Kashikawa, N., et al. 2006, Nature, 443, 186
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kirkpatrick, J. D., Gelino, C. R., Cushing, M. C., et al. 2012, ApJ, 753, 156
Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2003, HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System, 337
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kron, R. G. 1980, ApJS, 43, 305
Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, ApJ, 649, L67
Labbé, I., et al. 2010a, ApJ, 708, L26
Labbé, I., et al. 2010b, ApJ, 716, L103
Labbé, I., Oesch, P. A., Bouwens, R. J., et al. 2012, ApJ, submitted, arXiv:1209.3037
Lacey, C. G., Baugh, C. M., Frenk, C. S., & Benson, A. J. 2011, MNRAS, 412, 1828
Laidler, V. G., Papovich, C., Grogin, N. A., et al. 2007, PASP, 119, 1325
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Lorenzoni, S., Bunker, A. J., Wilkins, S. M., et al. 2011, MNRAS, 414, 1455
Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388
Madau, P., Pozzetti, L. & Dickinson, M. 1998, ApJ, 498, 106
Maizy, A., Richard, J., de Leo, M. A., Pelló, R., & Kneib, J. P. 2010, A&A, 509, A105
Makovoz, D., & Khan, I. 2005, Astronomical Data Analysis Software and Systems XIV, 347, 81
McLure, R. J., Dunlop, J. S., Cirasuolo, M., et al. 2010, MNRAS, 403, 960
Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64
Mosleh, M., Williams, R. J., Franx, M., et al. 2012, ApJ, 756, L12
Narayan, R., & Bartelmann, M. 1996, arXiv:astro-ph/9606001
Oesch, P. A., Bouwens, R. J., Carollo, C. M., et al. 2010, ApJ, 709, L21
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012a, ApJ, 745, 110
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012b, ApJ, 759, 135
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Popesso, P., et al. 2009, A&A, 494, 443
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Rayner, J.T., Cushing, M.C., & Vacca, W.D. 2009, ApJS, 185, 289
Reddy, N. A., & Steidel, C. C. 2009, ApJ, 692, 778
Richard, J., Stark, D. P., Ellis, R. S., et al. 2008, ApJ, 685, 705
Santos, M. R., Ellis, R. S., Kneib, J.-P., Richard, J., & Kuijken, K. 2004, ApJ, 606, 683
Schenker, M. A., Stark, D. P., Ellis, R. S., et al. 2012, ApJ, 744, 179
Schiminovich, D., et al. 2005, ApJ, 619, L47
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, ApJ, 626, 698
Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103
Sirianni, M., et al. 2005, PASP, 117, 1049
Stanway, E. R., Bunker, A. J., & McMahon, R. G. 2003, MNRAS, 342, 439
Stanway, E. R., McMahon, R. G., & Bunker, A. J. 2005, MNRAS, 359, 1184
Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408, 1628
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M. and Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728
Stutz, A. M., Papovich, C., & Eisenstein, D. J. 2008, ApJ, 677, 828
Szalay, A. S., Connolly, A. J., & Szokoly, G. P. 1999, AJ, 117, 68
Trenti, M., & Stiavelli, M. 2008, ApJ, 676, 767
Trenti, M., Stiavelli, M., Bouwens, R. J., et al. 2010, ApJ, 714, L202
Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2011, ApJ, 727, L39
van der Wel, A., Straughn, A. N., Rix, H.-W., et al. 2011, ApJ, 742, 111
Vanzella, E., et al. 2009, ApJ, 695, 1163
Whitelock, P., Menzies, J., Feast, M., et al. 1995, MNRAS, 276, 219
Wilkins, S. M., Bunker, A. J., Stanway, E., Lorenzoni, S., & Caruana, J. 2011, MNRAS, 417, 717
Williams, R. E., Blacker, B., Dickinson, M., et al. 1996, AJ, 112, 1335
Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, ApJS, 193, 27
Wuyts, S., Labbé, I., Schreiber, N. M. F., et al. 2008, ApJ, 682, 985
Yan, H., & Windhorst, R. A. 2004, ApJ, 612, L93
Yan, H.-J., Windhorst, R. A., Hathi, N. P., et al. 2010, Research in Astronomy and Astrophysics, 10, 867

Zheng, W., Postman, M., Zitrin, A., et al. 2012, Nature, 489, 406 (Z12)

APPENDIX

A. DESCRIPTION OF THE REDSHIFT PRIORS

In §3.4, we present redshift likelihood distributions for the three candidate $z \sim 9$ galaxies in our selection. This allows us to estimate the relative probability that sources in this sample correspond to higher or lower redshift galaxies (Figure 5). However, in doing so, we must utilize a prior. We consider three different redshift priors: (1) a flat redshift-independent prior, (2) a prior calibrated to published LF or LF trends, and (3) a prior tuned to reproduce the results from our photometric scattering experiments (§3.5). This section discusses the latter two priors in detail.

LF-calibrated Prior: The second prior we consider is calibrated according to published LFs or LF trends. For this prior, we give special attention to two galaxy populations: star-forming galaxies at $z \sim 9$ and faint red galaxies at $z \sim 1.3$-2. These are the only two galaxy populations which can at least provide approximate fits to the sources in our selection and therefore have nominal χ^2's that are not especially large. For the faint red $z \sim 1.3$-2 galaxy case, we calibrate our priors based on the LF results of Giallongo et al. (2005) for red galaxies using deep near-IR observations available over the HDF-North and HDF-South fields (Williams et al. 1996; Casertano et al. 2000) and the K20 spectroscopic sample (Cimatti et al. 2002). At $z \sim 2$, their $< m^*/m(bimodal)$ LF results correspond to $M^*_{B,0} = -21.90$ mag, $\phi^* = 2 \times 10^{-4}$ Mpc^{-3} mag^{-1}, and $\alpha = -0.53$. The basic validity of these LF results has been verified with much improved statistics based on new results for red galaxies over the UKIDSS Ultra Deep Survey field (Lawrence et al. 2007) where fits yield $M_V = -21.9$, $\phi^* = 2 \times 10^{-4}$ Mpc^{-3}, and $\alpha = 0.07$ (R. Quadri et al. 2012, private communication). Meanwhile, at $z \sim 1.3$, the Giallongo et al. (2005) $< m^*/m(bimodal)$ LF results correspond to $M^*_{B,0} = -21.49$ mag, $\phi^* = 5 \times 10^{-4}$ Mpc^{-3} mag^{-1}, and $\alpha = -0.53$. Finally, for the $z \sim 9$ star-forming galaxy case, our priors use the Bouwens et al. (2011b) LF fitting formula as a guide (which is a parameterization of the evolution of the UV LF from $z \sim 8$ to $z \sim 4$: see §7.5 of that paper).

Assuming a deep blank search at $\sim$28.5 mag (the approximate intrinsic magnitude of our candidates after correction for magnification) with a $\Delta z \sim 1$, Δmag ~ 1 selection window, we find that these LFs predict $\sim$1.2 $z \sim 9.2$ galaxies per arcmin^2, but 0.14 faint red galaxies per arcmin^2 over the redshift range $z \sim 1.3$-2.5. Surprisingly enough, these results suggest that we would be much more likely (i.e., by $\sim 9\times$) to find a blue galaxy at $z \sim 9$ with our selection than a faint red galaxy at $z \sim 1.3$-2. Even correcting these predictions based on the present search results for $z \sim 9$ galaxies (where we find just $\sim 55^{+75}_{-38}$% as many galaxies as expected from the Bouwens et al. 2011b fitting formula), $z \sim 9$ galaxies would still be $5\times$ more abundant on the sky at $\sim$ 28.5 mag than red (old and/or dusty) galaxies at $z \sim 1.3$-2.5. For the purposes of our "LF calibrated" prior, we will assume that $z \sim 9$ galaxies have a $5\times$ higher surface density on the sky than $z \sim 1.3$-2.5 red (old and/or dusty) galaxies.

Contamination-Tuned Prior: Of course, it is not simply the faint red (old and/or dusty) galaxies at $z \sim 1.3$-2 that can contaminate $z > 8$ selections. Other galaxies can scatter into $z \sim 9$ selections through noise. This makes the low-redshift solution more likely than what we would calculate based on observationally-based LFs. Considered by themselves, each photometrically-scattered source would be unlikely to look very much like a probable $z \sim 9$ candidate, but one must account for the fact that there are some $\sim 4 \times 10^4$ sources in our fields which noise could conspire to make look like such a $z \sim 9$-10 candidate. We account for this possibility with our third "contamination tuned" prior. With this prior, we adjust the relative likelihood of the high and low redshift peaks for our entire three source $z \sim 9$-10 galaxy sample so that it matches the 23% contamination rate estimated in our photometric scattering experiments described in §3.5.[33] However, it is worth keeping in mind that results based on the third prior likely overweight the probability that sources are low-redshift contaminants. This is due to our photometric scatter simulations not accounting for the fact that red (old and/or dusty) galaxies are even rarer at $\sim$27-28 mag than the $\sim$24-25.5 magnitude sources we use as inputs to our photometric scattering simulations (e.g., Figure 11 from Oesch et al. 2012a).

[33] Admittedly, a more accurate approach would be to determine the actual redshift distribution of the intermediate-magnitude sources scattering into our selection and present it in Figure 5.