The Astrophysical Journal, 624:L81–L84, 2005 May 10

IRAC MID-INFRARED IMAGING OF THE HUBBLE DEEP FIELD–SOUTH: STAR FORMATION HISTORIES AND STELLAR MASSES OF RED GALAXIES AT $z > 2$[1]

Ivo Labbé,[2,3] Jiasheng Huang,[4] Marijn Franx,[5] Gregory Rudnick,[6,7] Pauline Barmby,[4] Emanuele Daddi,[6,8] Pieter G. van Dokkum,[9] Giovanni G. Fazio,[4] Natascha M. Förster Schreiber,[10] Alan F. M. Moorwood,[11] Hans-Walter Rix,[12] Huub Röttgering,[5] Ignacio Trujillo,[12] and Paul van der Werf[5]
Received 2005 January 27; accepted 2005 April 4; published 2005 April 22

ABSTRACT

We present deep 3.6–8 μm imaging of the Hubble Deep Field–South with the Infrared Array Camera (IRAC) on the *Spitzer Space Telescope*. We study distant red galaxies (DRGs) at $z > 2$ selected by $J_s - K_s > 2.3$ and compare them with a sample of Lyman break galaxies (LBGs) at $z = 2$–3. The observed UV–to–8 μm spectral energy distributions are fitted with stellar population models to constrain star formation histories and derive stellar masses. We find that 70% of the DRGs are best described by dust-reddened star-forming models and 30% are very well fitted with old and "dead" models. Using only the $I - K_s$ and $K_s - 4.5$ μm colors, we can effectively separate the two groups. The dead systems are among the most massive at $z \sim 2.5$ (mean stellar mass $\langle M_* \rangle = 0.8 \times 10^{11}\ M_\odot$) and likely formed most of their stellar mass at $z > 5$. To a limit of $0.5 \times 10^{11}\ M_\odot$, their number density is ~10 times lower than that of local early-type galaxies. Furthermore, we use the IRAC photometry to derive rest-frame near-infrared J, H, and K fluxes. The DRGs and LBGs together show a large variation (a factor of 6) in the rest-frame K-band mass-to-light ratios (M/L_K), implying that even a *Spitzer* 8 μm–selected sample would be very different from a mass-selected sample. The average M/L_K of the DRGs is about 3 times higher than that of the LBGs, and DRGs dominate the high-mass end. The M/L_K values and ages of the two samples appear to correlate with derived stellar mass, with the most massive galaxies being the oldest and having the highest mass-to-light ratios, similar to what is found in the low-redshift universe.

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

1. INTRODUCTION

Galaxies at $z > 2$ exhibit very diverse properties: they range from the blue Lyman break galaxies (LBGs), which are bright in the rest-frame ultraviolet (Steidel et al 1996a, 1996b), to the distant red galaxies (DRGs), which are generally faint in the rest-frame UV and have fairly red rest-frame optical colors. The DRGs are selected in the observers' near-infrared (NIR) by the simple criterion $J_s - K_s > 2.3$ (Franx et al. 2003; van Dokkum et al. 2003). The variety in the galaxy population at $z > 2$ is comparable to that seen in the local universe, where colors range from very blue for young starbursting galaxies to very red for old elliptical galaxies. An urgent question is what causes the red colors of DRGs at $z > 2$. Are they "old and dead," or are they actively star-forming and more dusty than U-dropout galaxies?

First analyses of the optical-to-NIR spectral energy distributions (SEDs) and spectra have suggested that both effects play a role: they have higher ages, contain more dust, and have higher mass-to-light ratios (M/L) in the rest-frame optical than do LBGs (Franx et al. 2003; van Dokkum et al. 2004; Förster Schreiber et al. 2004, hereafter F04). Furthermore, many have high star formation rates, over 100 $M_\odot$ yr^{-1} (van Dokkum et al. 2004; F04). Unfortunately, the number of DRGs with rest-frame optical spectroscopy is very small. Inferences inevitably depend on SED fitting, which has large uncertainties (see, e.g., Papovich et al. 2001). The SED constraints on the stellar and dust content are expected to improve significantly when the photometry is extended to the rest-frame NIR.

Here we present the first results on rest-frame NIR photometry of DRGs and LBGs in the Hubble Deep Field–South (HDF-S) as observed with the Infrared Array Camera (IRAC; Fazio et al. 2004) on the *Spitzer Space Telescope*. Where necessary, we assume an $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, and a Salpeter initial mass function (IMF) between 0.1 and 100 $M_\odot$. Magnitudes are expressed in the AB photometric system unless explicitly stated otherwise.

2. OBSERVATIONS, PHOTOMETRY, AND SAMPLE SELECTION

We have observed the 5 arcmin^2 HDF-S WFPC2 field with IRAC, integrating 1 hr each in the mid-infrared (MIR) 3.6, 4.5, 5.8, and 8 μm bands. The observations, data reduction, and photometry will be described in I. Labbé et al. (2005, in preparation). Briefly, we reduced and calibrated the data using standard procedures (see, e.g., Barmby et al. 2004). The limiting

[1] Based on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 1407. Support for this work was provided by NASA through contract 125790 issued by JPL/Caltech. Based on service-mode observations collected at the European Southern Observatory, Paranal, Chile (program 164.O-0612). Based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555
[2] Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101; ivo@ociw.edu.
[3] Carnegie Fellow.
[4] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138.
[5] Sterrewacht Leiden, Postbus 9513, NL-2300 RA Leiden, Netherlands.
[6] National Optical Astronomy Observatory, 950 North Cherry Avenue, Tucson, AZ 85719.
[7] Goldberg Fellow.
[8] Spitzer Fellow.
[9] Department of Astronomy, Yale University, New Haven, CT 06520-8101.
[10] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, D-85748 Garching, Germany.
[11] European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany.
[12] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany.

depths at 3.6, 4.5, 5.8, and 8 μm are 25.0, 24.6, 22.6, and 22.5 mag, respectively (5 σ, 3″ diameter aperture). We PSF-matched the IRAC images and a very deep K_s image (Labbé et al. 2003) to the 8 μm band. Nearby sources were fitted and subtracted to avoid confusion (see Labbé 2004). We measured fluxes in 4″.4 diameter apertures. The K_s-IRAC colors were then combined with the HDF-S catalog of Labbé et al. (2003), resulting in 11-band photometry from 0.3 to 8 μm. In addition to the photometric errors, we add in quadrature an error of 10% to reflect absolute calibration uncertainties of IRAC.

We selected DRGs in the field of the HDF-S using the criteria of Franx et al. (2003), yielding 14 galaxies with photometric redshifts (Rudnick et al. 2003) in the range $1.9 < z < 3.8$. One blended source was excluded. In addition, we selected isolated LBGs in the same field to the same K_s-band limit and in the same redshift range, using the *U*-dropout criteria of Madau et al. (1996) on the WFPC2 imaging (Casertano et al. 2000). The two samples are complementary, with only one source in common. For both samples, the mean redshift is $z = 2.6$ with a standard deviation of 0.5.[13]

3. RED GALAXIES AT $z > 2$: OLD, DUSTY, OR BOTH?

In Figure 1, we analyze the $I-K_s$ versus the $K_s-4.5$ μm colors. The $K_s-4.5$ μm color has a sufficiently large baseline and much higher signal-to-noise ratio than the $K_s-5.8$ μm or K_s-8 μm color. The DRGs and LBGs lie in very different regions, with little overlap. The mean $K-4.5$ μm color of the DRGs is significantly redder than that of the LBGs, confirming that DRGs have higher M/L in the rest-frame optical. The red $I-K_s$ and $K-4.5$ μm colors of the DRGs imply that they must be prominent in IRAC-selected samples. Indeed, the majority of red $z-3.6$ μm sources selected by Yan et al. (2004) in the Great Observatories Origins Deep Survey satisfy the DRG selection criteria. We show color tracks for stellar population models (Bruzual & Charlot 2003), redshifted to a fixed $z = 2.6$: a single-age burst (SSP) model, and two models with constant star formation (CSF) and reddening typical for LBGs ($A_V = 0.6$; Shapley et al 2001) and for DRGs ($A_V = 2.0$; F04). The $K_s-4.5$ μm colors of LBGs are consistent with low-reddening models, whereas most of the DRGs lie in the area of models with substantial extinction.

3.1. *Dead Galaxies at* $z = 2$–3

Three of 13 DRGs (Fig. 1, *stars*) lie well outside the area of CSF models. Their $I-K_s$ colors are too red for their $K-4.5$ μm color, and they lie close to the line of an old SSP model with ages of 2–3 Gyr. The candidate old, passive galaxies are shown in Figure 2. Their SEDs are very well fitted by an old single-burst population. The gray model curves are predictions based on fits to the exceptionally deep optical/NIR data only, demonstrating that the *Spitzer* fluxes lie very close to the expected values. Dead galaxies have bluer $K-4.5$ μm colors than do dusty star-forming DRGs, and with *Spitzer* we can now effectively separate them.

Models with ongoing star formation and dust reddening fit the SEDs badly. Star formation histories (SFHs) with exponentially declining star formation rates (SFRs) and timescales greater than 300 Myr can be ruled out at the 99.9% confidence level. For the marginally acceptable models, the ratio of ongoing to past average SFR is $\mathrm{SFR}(t)/\langle\mathrm{SFR}\rangle = 0.001$, indicating that these three galaxies are truly "red and dead." The best-fit ages and implied formation redshifts depend on the assumed metallicity Z, as expected. Supersolar, $Z = 0.05$ models give a mean age of $\langle t\rangle = 1.4$ Gyr and mean formation redshift $\langle z_f\rangle = 5$; solar, $Z = 0.02$ models yield $\langle t\rangle = 2.9$ Gyr and $\langle z_f\rangle \gg 5$, while subsolar-metallicity models fail to provide good fits to the IRAC fluxes, as they are too blue. Hence we infer from the models that the "dead" galaxies had formed most of their stellar mass by $z \sim 5$ ($Z = 0.05$) or much earlier, $z \gg 5$ ($Z = 0.02$).

The number density of "dead" galaxies is 1.9×10^{-4} h_{70}^3 Mpc^{-3}, assuming a top-hat redshift distribution between $z = 2$ and $z = 3$. For stellar masses above 0.5×10^{11} $M_\odot$ and the same IMF, that number density is 10 times lower than that of early-type galaxies in the nearby universe (Bell et al. 2003). Obviously, the sampled volume is too small to allow firm conclusions, but the result is indicative of strong evolution of passive galaxies from $z = 2.5$ to $z = 0$.

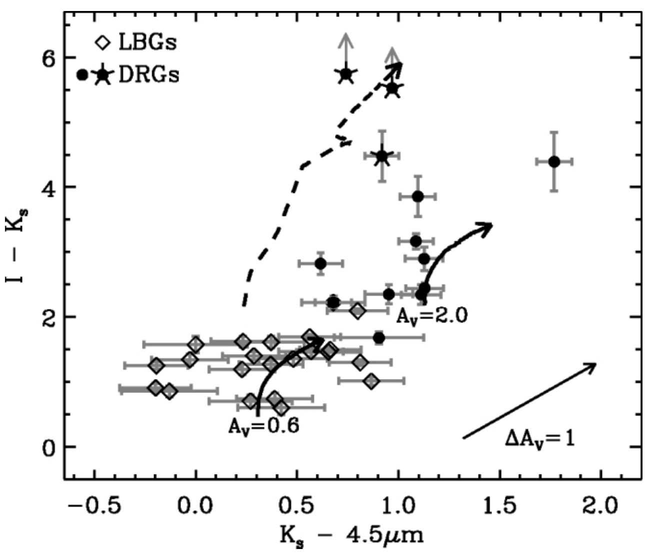

FIG. 1.—Observed $I-K_s$ vs. $K_s-4.5$ μm color-color diagram of two samples of $z > 2$ galaxies. Distant red galaxies (DRGs; *filled circles and stars*) occupy a different color region than $z \sim 2.5$ Lyman break galaxies (LBGs; *diamonds*). DRGs that are best described by an old single-age burst (SSP) model (*stars*) have colors distinct from those better described by constant star formation (CSF) models and dust, indicating that IRAC fluxes may help in separating these populations. The curves show the color evolution tracks of Bruzual & Charlot (2003) models at a fixed $z = 2.6$: an SSP model with ages ranging from 0.3 to 3 Gyr (*dashed line*) and two CSF models with ages ranging from 0.1 to 3 Gyr and reddenings of $A_V = 0.6$ and $A_V = 2.0$, respectively (*solid lines*). The vector indicates a reddening of $A_V = 1$ mag for a Calzetti et al. (2000) law.

3.2. *Dusty Star-forming Galaxies*

In Figure 2, we show the SEDs for three of the eight DRGs whose optical/NIR fluxes were better fitted with CSF models and reddening by a Calzetti et al. (2000) dust law. The observed MIR flux points are often different from predictions based on shorter wavelengths, especially where the flux density F_λ was still rising at the observed K_s band. Hence, for the very dusty DRGs ($A_V > 1.5$) the MIR fluxes help to better constrain the age and dust in the models. The average ages and extinction

[13] The photometric redshift accuracy of DRGs from other studies (F04; S. Wuyts et al. 2005, in preparation) is $|z_{\rm spec} - z_{\rm phot}|/(1 + z_{\rm spec}) \approx 0.1$. The estimate is based on direct comparison of 16 galaxies with both photometric and spectroscopic redshifts. Uncertainties in rest-frame magnitudes and model parameters are based on Monte Carlo simulations that take into account uncertainties in the observed fluxes and redshift estimates.

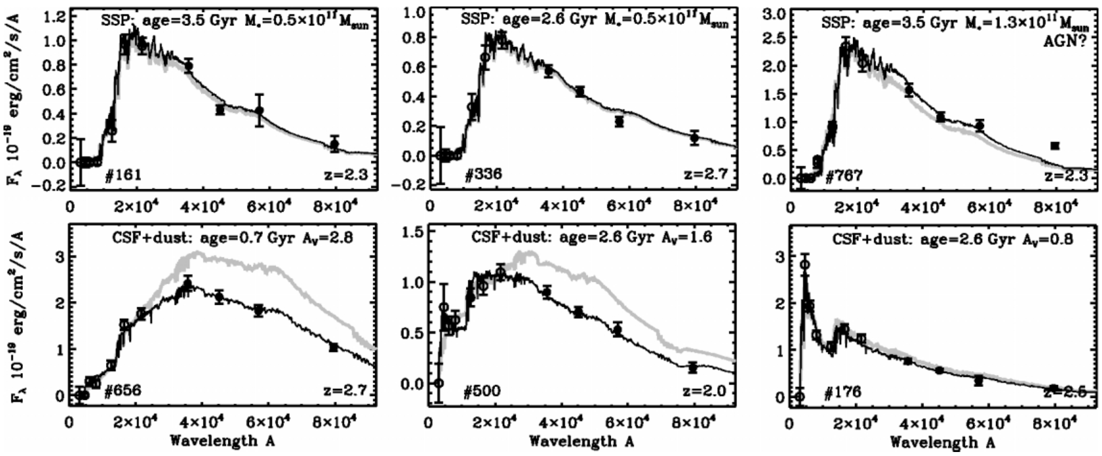

FIG. 2.—The UV–to–8 μm spectral energy distributions of DRGs. The best-fit Bruzual & Charlot (2003) models to the full SEDs (*black solid lines*) are shown, as well as predictions based on the ultradeep optical and near-infrared fluxes only (*gray solid lines*). *Top*: The three DRGs best fitted with an SSP model. The mid-infrared IRAC fluxes (*filled circles*) directly confirm the predictions. Old and dead DRGs have bluer $K_s - 4.5$ μm colors than very dusty star-forming DRGs. Galaxy 767 shows a flux excess at 8 μm, possibly related to AGN activity. *Bottom*: Mid-infrared predictions for galaxies better fitted with CSF models and Calzetti et al. (2000) dust reddening. For highly reddened galaxies, the MIR observations can be very different from the predictions. Here IRAC fluxes help to better constrain the total age, dust content, and stellar mass in the models.

of the fits changed mildly after inclusion of the IRAC data: $\langle t \rangle$ from 1.1 to 1.3 Gyr, and $\langle A_V \rangle$ from 1.9 to 1.5 mag.

Finally, two remaining galaxies could not be fitted well by any model, most likely because of emission-line contamination. One of them is spectroscopically confirmed as a very strong Lyα emitter (S. Wuyts et al. 2005, in preparation). However, the generally good fits to the UV–to–8 μm SEDs of the DRGs indicate that the red $J_s - K_s$ colors were caused by old age and dust-reddened starlight, and not other anomalies. We note that one of the dead galaxies (see Fig. 1) has an apparent flux excess at 8 μm, suggesting the presence of an obscured active galactic nucleus (AGN) that starts to dominate the flux in the rest-frame K band (see, e.g., Stern et al. 2004).

4. MASS-TO-LIGHT RATIO VARIATIONS AT $z = 2$–3

We next study the M/L values inferred from the SED fits described above, adopting the best-fitting SFH (SSP or CSF). Figure 3*a* shows the modeled M/L_K (rest-frame K) versus rest-frame $U - V$. The curves show the dust-free Bruzual & Charlot (2003) model M/L_K values. Galaxies generally lie to the red of the model curves, as a result of dust attenuation. The fits imply a large range in M/L_K for DRGs and LBGs together (a factor of 6). Furthermore, M/L_K correlates with rest-frame $U - V$ color, such that galaxies that are red in the rest-frame optical have higher M/L_K values than blue galaxies. The average M/L_K of the DRGs is 0.33 ± 0.04, about 3 times higher than that of the LBGs. Figure 3*b* shows the derived M/L_K against the derived age. The M/L_K correlates well with the age, as expected from the models. The role of extinction is greatly reduced in the rest-frame K band, implying that the differences in the M/L_K values for DRGs and LBGs are driven by age differences.

Figure 3*c* shows M/L_K against the stellar mass M_*. DRGs dominate the high-mass end. The highest-mass galaxies ($>0.5 \times 10^{11}$ $M_\odot$) in this sample all have high values of M/L_K, and here the ratio does not depend strongly on M_*. At lower masses, galaxies have much lower values of M/L_K. This may be a selection effect caused by our magnitude cutoff, as we would miss low-mass galaxies with high M/L_K. However, high-mass galaxies with *low* M/L_K would be detected if they existed, and hence the lower envelope of the M/L_K versus M_* distribution is real. Two intriguing possibilities are that the mean M/L_K decreases toward lower stellar mass, or that the intrinsic scatter in M/L_K increases.

Our DRGs and LBGs were selected in the observed K_s band (rest-frame V band). We find a large variation in M/L_V (a factor of 25), and hence selection effects play a major role in NIR studies at high redshift. Selection using MIR observations with *Spitzer* would improve the situation, but the wide range in M/L_K values and ages found here indicates that even an IRAC 8 μm–selected sample would still be very different from a mass-selected sample.

5. DISCUSSION AND CONCLUSIONS

We have presented rest-frame UV-to-NIR photometry of a sample of DRGs and LBGs at $z = 2$–3 in the HDF-S. These galaxies span a wide range in properties, similar to low-redshift galaxies. The rest-frame NIR photometry from IRAC helps significantly: first, by allowing us to separate "old and dead" from dusty star-forming DRGs using only the observed I, K_s, and 4.5 μm band, and second, by improving model constraints on the heavily obscured DRGs.

The wide range in galaxy properties at $z = 2.5$ raises several important issues. First, it demonstrates that it is impossible to obtain mass-selected samples photometrically. Even in the rest-frame K band, the mass-to-light ratio varies by a factor of 6 for the DRGs and LBGs in the HDF-S. Second, we need to understand what produces these variations. If a relation between total stellar mass and M/L_K exists (see Fig. 3*c*), then stellar mass might be driving the variations. Deeper IRAC data are needed in order to establish this well, as incompleteness effects may play a role. It is tantalizing that Kauffmann et al. (2003) find a similar correlation at $z = 0.1$, with the most massive galaxies being the oldest and having the highest mass-to-light ratios. Kauffman et al. also found that above a stellar mass $M_* = 6 \times 10^{10}$ $M_\odot$, galaxy properties correlate only weakly with M_*, similar to what we find at $z = 2$–3. The simplest explanation is that we observe the same galaxies at $z = 0.1$ and $z = 2.5$, although we note that hierarchical formation scenarios predict significant merging for galaxies at $z > 2$. Alternatively, perhaps we observe merging processes occurring

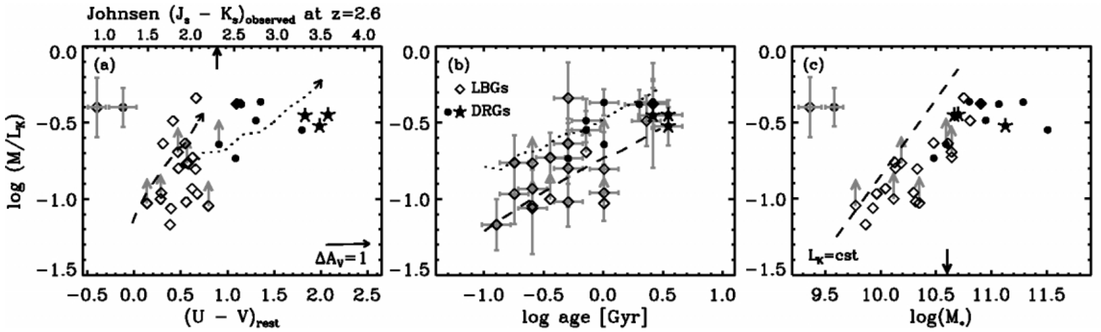

FIG. 3.—Mass-to-light ratios in the rest-frame K band (M/L_K) of DRGs and $z = 2$–3 LBGs. Symbols are as in Fig. 1. (*a*) M/L_K vs. rest-frame $U-V$ color, showing a clear correlation. The top axis indicates the corresponding observed J_s-K_s color and the $J_s-K_s > 2.3$ limit (*upward-pointing arrow*). Color tracks of an SSP model (*dotted line*) and a CSF model (*dashed line*) are shown. (*b*) Relation between M/L_K and best-fit age. The M/L_K is more sensitive to the age of the stellar population than to dust extinction; hence, the DRGs are on average older than, not just more dusty versions of, LBGs. (*c*) M/L_K vs. stellar mass. The highest-mass galaxies all have high M/L_K. The mass completeness limit (*downward-pointing arrow*) corresponds to our observed $K_{s,tot} = 22.5$ mag cutoff and a maximal M/L. A selection by rest-frame K light corresponds to taking everything to the right of the constant-L_K line (*dashed line*). Obviously this is still far from a vertical cutoff, which would represent a selection by stellar mass.

at a critical mass of about $6 \times 10^{10}\ M_\odot$. A partial explanation may be a simple relation between mass and metallicity: higher mass galaxies might have higher metallicity and thus be redder.

It is very noteworthy that there are “red and dead” galaxies at $z = 2$–3. Previous authors found such galaxies at $z = 1$–2 (e.g., Cimatti et al. 2004; McCarthy et al. 2004). Whereas the number density of “dead” galaxies at $z = 2.5$ is probably much lower than at $z = 0$, the mere existence of these systems at such high redshift raises the question what caused such a rapid and early decline in star formation. Our model fits imply that their star formation stopped by $z = 5$ or higher, close to or during the epoch of reionization. It is possible that a strong feedback mechanism caused this, such as an AGN or galactic-scale outflow. We note that these galaxies are among the most massive galaxies at $z = 2.5$ and hence were probably at the extreme tail of the mass function at $z = 5$.

Obviously, many questions remain unanswered by this study. Only a very small field has been studied, and similar studies on wider fields are necessary. Finally, the mass-to-light ratios derived here remain model dependent and vary with the assumed SFH, metallicity, and IMF (see, e.g., Bell et al. 2003). Direct mass determinations are presently very challenging but are becoming more important to understand the $z = 2$–3 galaxy population.

This research was supported by the Carnegie Institution of Washington, the Netherlands Foundation for Research (NWO), the Leids Kerkhoven-Bosscha Fonds, the Lorentz Center, and the Smithsonian Institution.

REFERENCES

Barmby, P., et al. 2004, ApJS, 154, 97
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Casertano, S., et al. 2000, AJ, 120, 2747
Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, ApJ, 533, 682
Cimatti, A., et al. 2004, Nature, 430, 184
Fazio, G. G., et al. 2004, ApJS, 154, 10
Förster Schreiber, N. M., et al. 2004, ApJ, 616, 40 (F04)
Franx, M., et al. 2003, ApJ, 587, L79
Kauffmann, G., et al. 2003, MNRAS, 341, 54
Labbé, I. 2004, Ph.D. thesis, Univ. Leiden
Labbé, I., et al. 2003, AJ, 125, 1107
Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, MNRAS, 283, 1388
McCarthy, P. J., et al. 2004, ApJ, 614, L9
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Rudnick, G., et al. 2003, ApJ, 599, 847
Shapley, A. E., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, ApJ, 562, 95
Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996a, AJ, 112, 352
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996b, ApJ, 462, L17
Stern, D., et al. 2004, ApJL, submitted (astro-ph/0410523)
van Dokkum, P. G. et al. 2003, ApJ, 587, L83
———. 2004, ApJ, 611, 703
Yan, H., et al. 2004, ApJ, 616, 63