The spectral properties of Lyman break galaxies (LBGs) offer a means to isolate pure samples displaying either dominant Lyα in absorption or Lyα in emission using broadband information alone. We present criteria developed using a large z ∼ 3 LBG spectroscopic sample from the literature that enables large numbers of each spectral type to be gathered in photometric data, providing good statistics for multiple applications. In addition, we find that the truncated faint, blue-end tail of z ∼ 3 LBG population overlaps and leads directly into an expected Lyα emitter (LAE) population. As a result, we present simple criteria to cleanly select large numbers of z ∼ 3 LAEs in deep broadband surveys. We present the spectroscopic results of 32 r′ ≲ 25.5 LBGs and r′ ≲ 27.0 LAEs at z ∼ 3 preselected in the Canada–France–Hawaii Telescope Legacy Survey that confirm these criteria.

Key words: galaxies: evolution – galaxies: formation – galaxies: fundamental parameters – galaxies: photometry

1. INTRODUCTION

The Lyman break galaxies (LBGs; Steidel et al. 1996) are a high-redshift population of star-forming galaxies selected by their rest-frame ultraviolet colors. LBGs at z ∼ 3 are faint (L* corresponds to m_R ∼ 24.5 at z ∼ 3) and require long integrations using 8 m class telescopes to obtain spectral information having a signal-to-noise ratio (S/N) of a few. Nevertheless, >1500 z ∼ 3 spectra have been obtained and studied (e.g., Le Fèvre et al. 2005; Cooke et al. 2005; Steidel et al. 2003; hereafter CCS03). Lyα is the most prominent feature in LBG spectra and has been shown to be indicative of other spectral properties, such as interstellar medium (ISM) line strength and continuum profile (Shapley et al. 2003; hereafter AES03).

Inspection of the spectroscopic samples gathered to date shows that ~50% of LBGs exhibit dominant (or net) Lyα in absorption, with the remaining exhibiting dominant Lyα in emission (e.g., AES03; Cooke et al. 2005). A recent investigation of close and interacting pairs (Cooke et al. 2009a) finds evidence of an overabundance of z ∼ 3 LBGs exhibiting Lyα emission in close pairs; all LBGs with projected physical separations ≲15 kpc display Lyα emission. To properly explore this relationship, along with other spectroscopic trends with spatial distribution, large samples that reflect the spectral properties of LBGs are necessary. However, the conventional means of multi-object spectroscopy is inefficient in acquiring the spectra of galaxies closely spaced on the sky because of inherent mechanical constraints (see Cooke et al. 2009a). As a result, comprehensive LBG spectroscopic surveys are difficult and time intensive.

High-redshift star-forming systems with continua typically too faint for spectroscopic follow-up, but with prominent Lyα emission (Cowie & Hu 1998; Hu et al. 1998), have been detected via targeted spectroscopic and narrowband surveys (e.g., Dawson et al. 2004; Ouchi et al. 2005; Venemans et al. 2005; Gawiser et al. 2006). Because of the faint continua of these Lyα emitters (LAEs), and the difficulties involved in their acquisition, LAE surveys have been limited to narrow redshift ranges and typically clustered fields. As a result, fewer total spectra have been compiled relative to LBGs, with the details of the process, or processes, that contribute to the generation and escape of Lyα emission, less certain. In addition, the data-gathering techniques have provided only a few cases to date where the mass of LAEs from clustering can be reasonably inferred (Ouchi et al. 2003; Shimasaku et al. 2004; Kovač et al. 2007; Gawiser et al. 2007).

LBGs at z ∼ 3 that are dominated by Lyα in absorption separate sufficiently in color–magnitude space from those dominated by Lyα in emission to enable a simple means using broadband imaging to isolate subsets with desired Lyα features and associated spectral properties. Furthermore, we find that deep broadband criteria can select clean, EW unbiased samples of z ∼ 3 LAEs with photometrically detectable continua. As a result, very large numbers z ∼ 3 systems over defined volumes with desired spectral properties can be efficiently obtained for statistical study, complementing the limitations of narrowband surveys and extensive deep spectroscopic campaigns. We describe the observations used in this work in Section 2. In Section 3, we present the behavior of LBG spectroscopic subsets, the criteria for LBG spectral-type and LAE selection, and the results of our spectroscopic test of the criteria. Finally, we provide a conclusion in Section 4. Magnitudes presented here are in the AB (Fukugita et al. 1996) magnitude system.

2. OBSERVATIONS

We use the publicly available data set of CCS03 and related accessible files to quantify the spectroscopic properties of z ∼ 3 LBGs and establish the U_r GR spectral-type selection criteria. This survey consists of ~2500 U_r GR selected and ~800 spectroscopically confirmed LBGs from 17 separate fields that effectively sample the z ∼ 3 LBG population as a whole. As a key component to this work, we use the Lyα EW measurements of AES03 for 775 LB spectroscopic properties of the CCS03 sample (A. E. Shapley 2009, private communication).

In Section 3.3, we test the criteria presented below using the spectra of 32 u''g'i-selected z ∼ 3 LBGs in the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) Deep field “D4”.3 We r-band select z ∼ 3 LBGs from stacked images that combine four years of high-quality observations that reach
Figure 1. Left panel: color–magnitude diagram of $z \sim 3$ LBGs in the Steidel et al. (2003) sample. Photometrically selected galaxies are plotted as gray crosses, whereas spectroscopically confirmed $z \sim 3$ LBGs are the larger (colored) symbols coded by their Lyα emission and represent the four quartiles as defined in Shapley et al. (2003). Specifically, group 1—large (gray/red) squares, group 2—small (gray/pink) squares, group 3—small (black/light blue) triangles, and group 4—large (black/blue) triangles. The large square and triangle and associated error bars represent the mean and 1σ standard deviation for the distributions of LBGs with net Lyα in absorption (aLBGs) and net Lyα in emission (eLBGs), respectively. Right two panels: histograms of the aLBG (gray/red), back-hatched) and eLBG (black/blue, forward-hatched) $(G – R)$ color and $R$ mag distributions (see the text). The $(G – R)$ color-selection criteria omit the far red tail of the aLBG distribution in an effort to avoid low-redshift interlopers. The spectroscopic-justified magnitude truncation ($R \leq 25.5$) of eLBGs suggests that a continuation of this distribution probes a Lyα emitter population.

3. ANALYSIS

We construct a $(G – R)$ versus $R$ color–magnitude diagram (CMD) for the CCS03 LBG photometric sample, plotted in the left panel of Figure 1. Motivated by the results of Cooke et al. (2009a), we overlay the spectroscopically confirmed LBGs and code them by their AES03 defined Lyα EW quartiles. AES03 find that LBGs with strong Lyα have blue $\sim 1300$–$2000$ Å continua with a continuous reddening trend seen with decreasing Lyα emission EW, through to strong absorption. This is the main cause of the color separation on the CMD in Figure 1. The presence of the Lyα feature in the $G$ band also contributes to the color separation, but to a lesser extent, and is on order of the $\sim 0.2$ mag uncertainties.

We use the term “aLBGs” for LBGs with dominant Lyα in absorption (net Lyα EW $< 0$) and “eLBGs” for dominant Lyα in emission (net Lyα EW $> 0$). We find that the aLBG and eLBG distributions show significant overlap but are distinct limiting magnitudes of $u^* \sim 27.5$, $g \sim 27.5$, $r \sim 27.0$, and $i \sim 26.7$. The stacked $r$-band images probe $\geq 1.0$–$1.5$ mag deeper than the $R$-band images of CCS03 and thereby test a fainter regime. The $z \sim 3$ LBG criterion ($u^* – g \geq 1.25 (g – i)$) for the CFHT Megacam filters imposes a weaker restriction on the $u^*$-band depth as compared to the $(U_n – G) \geq (G – R)$ + 1.0 criteria of CCS03. This minimizes the introduction of a bias in $r > 25.5$ LBG selection based on their $(g – i)$ colors. For further information on image stacking and color-selection spectroscopic tests, see Cooke et al. (2009b, Supplementary Information).

We acquired the spectroscopic data using the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995; McCarthy et al. 1998) mounted on the Keck I telescope on 2009 July 20 using the 400/3400 grism on the new blue arm and 400/8500 grating on the new red arm yielding a blue/red resolution of $\sim 300$ km s$^{-1}$. Time constraints resulted in a total integration time of 3600 s with $\sim 0.9$ seeing FWHM. This is a $\sim 0.5 \times$ the total integration time of typical $z \sim 3$ LBG spectroscopy, but is found to be sufficient to identify most $R \geq 25.5$ candidates and the Lyα emission feature of $R \geq 27.0$ LAEs.

Figure 2. Similar to Figure 1, but for LBGs that exhibit the opposite extremes in Lyα features and associated spectral profiles. Squares denote the subset of 150 LBGs that display the strongest Lyα emission and triangles denote the subset of 150 LBGs that display the strongest Lyα absorption. The solid (green) line denotes our primary statistical cut of the two distributions and is the basis for subsequent selection criteria. The dashed (red) and dot-dashed (blue) lines indicate a $(G – R)$ distribution 1.5σ separation from the primary cut for the Lyα absorber and Lyα emitter subsets, respectively. This corresponds to $\geq 2.5\sigma$ from each of the respective $(G – R)$ and $R$ magnitude mean values. Objects beyond these cuts produce $\geq 95\%$ pure samples of either spectral type. Objects fainter than $R = 25.5$ that lay $\geq 3\sigma$ from the aLBG color distribution mean value comprise an expected population of LAEs. This region is shown bounded by dotted lines.

Next, we study two LBG spectral subsets that exhibit the opposite extremes in Lyα EW behavior. Subset 1 contains the 150 strongest aLBGs (net Lyα EW $\leq 12.0$) and subset 2 consists of the 150 strongest eLBGs (net EW $\geq 26.5$) of the full spectroscopic sample. The $(G – R)$ versus $R$ CMD of the two subsets is plotted in Figure 2 and shows two cohesive distributions with a more distinct separation when compared to populations. A two-sided K–S test of the aLBG and eLBG $(G – R)$ distributions produces a probability of $p = 1.5 \times 10^{-13}$ that the two distributions are pulled from the same population. Similarly, a K–S test produces $p = 3.0 \times 10^{-13}$ for the $R$ mag distributions. This is illustrated in the right panels of Figure 1.

Time constraints resulted in a total integration time of 3600 s with $\sim 0.9$ seeing FWHM. This is a $\sim 0.5 \times$ the total integration time of typical $z \sim 3$ LBG spectroscopy, but is found to be sufficient to identify most $R \geq 25.5$ candidates and the Lyα emission feature of $R \geq 27.0$ LAEs.
the aLBG/eLBG distribution values shown in Figure 1. Using the statistical mean values of the two subsets, we sever the populations in both color and magnitude as shown by the solid line in Figure 2. This line falls $\sim 1\sigma$ from both the color and magnitude mean values of both distributions and defines our primary cut. We show that more restrictive selections based on the slope of this cut (defined below), are very effective in isolating LBG spectral types with differing $\text{Ly}_\alpha$ and continuum profiles.

3.1. Spectral-type Selection Criteria

We use the parameters of the subsets 1 and 2 distributions to refine the primary cut in an effort to generate pure samples of each spectral type. Choosing CMD cuts that are separated from the primary cut by $1.5 \sigma$ of the $(G - R)$ distributions of each subset and along the same slope as the primary cut, selects LBG spectral types $\gtrsim 2.5 \sigma$ from the mean values of the opposite distribution. We find that the following criteria select nearly pure samples of aLBGs:

\[
(G - R) \gtrsim 0.4047 R - 9.376 + 1.5 \sigma_A, \tag{1}
\]

and eLBGs

\[
(G - R) \lesssim 0.4047 R - 9.376 - 1.5 \sigma_E, \tag{2}
\]

where $\sigma_A = 0.3095$ and $\sigma_E = 0.2392$ and are the $1\sigma$ standard deviations of the $(G - R)$ distributions for subset 1 and subset 2, respectively. These selection cuts are indicated in Figure 2.

Applying Equations (1) and (2) to the full spectroscopic data set of CCS03 produces spectral-type samples consisting of 40 aLBGs and 60 eLBGs with contamination fractions of 0.100 and 0.067, respectively (Figure 3; top panel). In addition, Figure 3 presents the net $\text{Ly}_\alpha$ EW and redshift distributions for the two isolated samples (center and bottom panels, respectively). The samples exhibit $66 \pm 36$ Å EW for the eLBGs and $-20 \pm 10$ Å EW for the aLBGs and their redshift distributions are found to be representative of the full data set ($z_{\text{FULL}} = 2.97 \pm 0.27$, $z_{\text{aLBG}} = 3.05 \pm 0.25$, and $z_{\text{eLBG}} = 2.88 \pm 0.24$). The above criteria, of course, can be made more strict by using a larger coefficient of $\sigma_A$ and $\sigma_E$, thereby producing samples that are more pure, at the cost of sample size. In this manner, large photometric data sets can obtain very clean spectral-type samples while still maintaining a large number of objects for good statistics.

3.2. $\text{Ly}_\alpha$ Emitters

LAEs at $z \sim 3$ have colors similar to eLBGs with the bulk expected to have $(G - R) \lesssim 0.0$ (see Reddy & Steidel 2009, Appendix A). The colors and mass of LAEs (e.g., Gawiser et al. 2007; Lai et al. 2008) appear to provide a natural extension of the LBG population and would help to complete the $R \sim 25.5$ eLBG magnitude truncation.

From the aLBG/eLBG color and magnitude distributions and those of their spectral subsets, and the results of the spectral-type selection criteria above, we find that a pure sample of LAEs...
(defined here as having $R > 25.5$) can be obtained using

$$(G - R) \leq 0.4047R - 9.376 - 2.0\sigma_E$$ and $R \geq 25.5$, (3)

which modifies the eLBG cut to a larger separation from the primary cut. This effectively avoids systems displaying Ly$\alpha$ in absorption by selecting objects $>3\sigma$ from the mean of the aLBG distribution. The region of the CMD defined by Equation (3) is indicated in Figure 2 and shows that the few spectroscopically identified objects in the CCFHTLS sample that meet these criteria exhibit Ly$\alpha$ in emission. Similar to above, the purity of the LAE sample can be determined by the $\sigma_E$ coefficient.

### 3.3. Spectroscopic Tests of the Spectral-type Predictions

We use 34316 $u'g'i$-selected $z \sim 3$ LBGs in the CCFHTLS Deep field “D4” to test the predictions of the spectral-type criteria and our LAE assumptions. Figure 4 plots the $(g - i)$ versus $i$ mag CMD for the “D4” field and shows our tentative spectral-type cuts based on the $r < 25.5 z \sim 3$ LBG densities and the results of the above analysis. The primary cut shown here produces aLBG/eLBG samples with a very similar $r \leq 25.5$ ratio (1460/5200) as compared to the R ratio of CCFHTLS when maintaining conventional color-selection criteria, and specifically the constraint $(g - i) \geq -0.2$.

For more complete color-space detection, we relax the criterion to $(g - i) \geq -1.0$ to include the small number of bluer objects to probe the full eLBG and expected LAE $(g - i)$ distribution. The relaxation of the color in this manner does not increase the fraction of interlopers (e.g., CCFHTLS; Cooke et al. 2005). The “D4” field is reflective of the remaining three CFHTLS Deep fields in depth and generates spectral-type samples of $\sim 1600$ aLBGs and $\sim 14000$ eLBGs, $\sim 8000$ of which fall in the LAE selection region (LAE magnitude definition shown in Figure 4 is $i > 25.5$).

The analysis of our 32 spectra (Section 2) indicates that our tentative spectral-type cuts are very effective. Of the 11 LBGs targeted in the aLBG spectral-type cut, eight have confirmed Ly$\alpha$ in absorption, two show complex absorption and emission profiles with net Ly$\alpha$ emission, and one is unidentified as a result of its low S/N. Nine of the 17 LBGs targeted in the eLBG region were unidentified (although six show weak evidence of emission), with eight exhibiting Ly$\alpha$ emission. Seven of the eight galaxies that fall in the LAE region show Ly$\alpha$ emission, with the remaining galaxy being unidentified. This demonstrates a high efficiency in identifying $i \lesssim 27.0$ LAEs from their broadband colors and helps to confirm the LAE extension to the LBG population.

### 4. CONCLUSION

We present an analysis of the photometric properties of the $z \sim 3$ LBG spectroscopic sample of Steidel et al. (2003). The relationships between Ly$\alpha$ EW and $(G - R)$ and Ly$\alpha$ EW and $R$ mag enable a spectral separation of the LBG population using broadband data. We define statistical photometric cuts to reliably generate $\gtrsim 90\%$ pure spectral-type samples of LBGs displaying dominant Ly$\alpha$ in absorption and dominant Ly$\alpha$ in emission. In addition, the spectral-type broadband criteria are extended to isolate clean samples of LAEs.

Our spectroscopic sample of 32 $z \sim 3$ LBGs from the CCFHTLS demonstrates the efficiency of the broadband criteria presented here in identifying galaxies based on their Ly$\alpha$ feature, including LAEs. Use of this method will allow the statistical study of $z \sim 3$ galaxies populations from the large numbers easily acquired in deep broadband surveys. This circumvents the expense and constraints of investigations using multi-object spectroscopy and the limitations of narrowband searches. Further tests of the criteria, and the extension of the results to other redshifts and filter sets, are presented in a future paper.

J.C. thanks A. E. Shapley for helpful discussions and access to the spectroscopic values. J.C. gratefully acknowledges generous support by Gary McCue.

**REFERENCES**

Cooke, J., Berrier, J. C., Barton, E. J., Bullock, J. S., & Wolfe, A. M. 2009a, MNRAS, submitted

Cooke, J., Sullivan, M., Barton, E. J., Bullock, J. S., Carlberg, R. G., Gal-Yam, A., & Tollerud, E. 2009b, Nature, 460, 237

Cooke, J., Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ApJ, 621, 596

Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319

Dawson, S., et al., 2004, ApJ, 617, 707

Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748

Gawiser, E., et al., 2006, ApJ, 642, L13

Gawiser, E., et al., 2007, ApJ, 671, 278

Hu, E. M., Cowie, L. L., & McMahon, R. G. 1998, ApJ, 502, L99

Kovač, K., Somerville, R. S., Rhoads, J. E., Malhotra, S., & Wang, J. 2007, ApJ, 668, 15

Lai, K., et al., 2008, ApJ, 674, 70

Le Fèvre, O., et al., 2005, A&A, 439, 845

McCarthy, J. K., et al., 1998, Proc SPIE, 3355, 81

Oke, J. B., et al., 1995, PASP, 107, 375

Ouchi, M., et al., 2003, ApJ, 582, 60

Ouchi, M., et al., 2005, ApJ, 620, L1

Reddy, N. A., & Steidel, C. C. 2009, ApJ, 692, 778

Shapley, A. E., Steidel, C. C., Adelberger, K. L., & Pettini, M. 2003, ApJ, 588, 65

Shimasaku, K., et al., 2004, ApJ, 605, L93

Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, 17

Venemans, B. P., et al., 2005, A&A, 431, 793

---

**Figure 4.** $(g - i)$ vs. $i$-band CMD for $u'g'i'$-selected $z \sim 3$ objects in the CCFHTLS Deep field “D4” (see the text). Objects are plotted similar to Figure 2 but use light gray (black) triangles to indicate those meeting the adopted LAE criteria. A sample of 32 systems were targeted for spectroscopic follow-up. Black squares mark confirmed aLBGs, black (cyan) triangles mark confirmed eLBGs, and black (yellow) crosses indicate low S/N unidentified systems. Large squares indicate strong absorption and large triangles indicate strong emission. Similar to above, the purity of the LAE population accessible by this method.