STRONG FIELD-TO-FIELD VARIATION OF Lyα NEBULAE POPULATIONS AT z ≃ 2.3

YUJIN YANG1,2, ANN ZABLUDOFF2, DANIEL EISENSTEIN2, and ROMEELE DAVE2
1 Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany; yyang@mpia.de
2 Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA; azabludoff@as.arizona.edu, eisenste@as.arizona.edu, rad@as.arizona.edu

Received 2010 May 17; accepted 2010 June 30; published 2010 August 2

ABSTRACT

Understanding the nature of distant Lyα nebulae, aka “blobs,” and connecting them to their present-day descendants requires constraining their number density, clustering, and large-scale environment. To measure these basic quantities, we conduct a deep narrowband imaging survey in four different fields, Chandra Deep Field South (CDFS), Chandra Deep Field North (CDFN), and two COSMOS subfields, for a total survey area of 1.2 deg2. We discover 25 blobs at z = 2.3 with Lyα luminosities of $L_{\text{Ly}\alpha} \gtrsim (0.7–8) \times 10^{43}$ erg s$^{-1}$ and isophotal areas of $A_{\text{iso}} = 10–60$′′. The transition from compact Lyα emitters (LAEs; $A_{\text{iso}} \sim$ a few ′′) to extended Lyα blobs ($A_{\text{iso}} > 10$′′) is continuous, suggesting a single family perhaps governed by similar emission mechanisms. Surprisingly, most blobs (16/25) are in one survey field, the CDFS. The six brightest, largest blobs with $L_{\text{Ly}\alpha} \gtrsim 1.5 \times 10^{43}$ erg s$^{-1}$ and $A_{\text{iso}} > 16$′′ lie only in the CDFS. These large, bright blobs have a field-to-field variance of $\sigma_n \sim 1.5$ (150%) about their number density $n \sim 10^{1.8} \times 10^{-5}$ Mpc$^{-3}$. This variance is large, significantly higher than that of unresolved LAEs ($\sigma_n \sim 0.3$ or 30%), and can adversely affect comparisons of blob number densities and luminosity functions (LFs) among different surveys. Our deep, blind survey allows us to construct a reliable blob LF. We compare the statistics of our blobs with dark matter halos in a 1 N-body simulation. At $z = 2.3$, $n$ implies that each bright, large blob could occupy a halo of $M_{\text{halo}} \gtrsim 10^{13} M_\odot$ if most halos have detectable blobs. The predicted variance in $n$ is consistent with that observed and corresponds to a bias of ≃ 7. Blob halos lie at the high end of the halo mass distribution at $z = 2.3$ and are likely to evolve into the $\sim 10^{14} M_\odot$ halos typical of galaxy clusters today. On larger scales of $\sim 10$ comoving Mpc, blobs cluster where compact LAEs cluster, indicating that blobs lie in coherent, highly overdense structures.

Key words: galaxies: formation – galaxies: high-redshift – intergalactic medium

Online-only material: color figures

1. INTRODUCTION

Lyα nebulae, or “blobs,” are extended sources at $z \sim 2$–6 with typical Lyα sizes of $\gtrsim 5$′′ ($\gtrsim 50$ kpc) and line luminosities of $L_{\text{Ly}\alpha} \gtrsim 10^{43}$ erg s$^{-1}$ (e.g., Keel et al. 1999; Steidel et al. 2000; Francis et al. 2001; Matsuda et al. 2004; Dey et al. 2005; Smith & Jarvis 2007; Hennawi et al. 2009; Prescott et al. 2009; Yang et al. 2009). Because the large spatial extent of their Lyα-emitting gas implies an interaction between the surrounding intergalactic medium and any embedded galaxies, blobs may signal an important phase of galaxy formation in the early universe, including cold gas accretion (Haiman et al. 2000), or intense radiative feedback from active galactic nuclei (AGNs; Geach et al. 2009). Despite the importance of blobs and the controversy regarding their origins, even basic properties such as their number density, clustering, and large-scale environment are poorly constrained.

To understand into what these mysterious objects will evolve in the present-day universe, measuring their statistics is critical due to the direct connection of number density and field-to-field variance to halo mass in ACDM cosmology. Currently, the halo mass of blobs is unknown. Using the spatial extent and linewidth of the Lyα line, Matsuda et al. (2006) estimate dynamical masses of $(0.5–20) \times 10^{12} M_\odot$ if the extended Lyα emission is from gravitationally bound gas clouds and the resonant scattering of Lyα can be ignored.

The clustering of blobs has not been measured directly in the past work, yet there are hints that it is strong. After surveying over $\sim 4.8$ deg$^2$ in the NOAO Deep-Wide Boötes field (Jannuzi & Dey 1999), Yang et al. (2009) discover just four bright blobs, yet two of them lie within only $\sim 70$′′ of each other (see also the discovery of a Lyα blob near a radio-loud Lyα halo by Matsuda et al. 2009). Some, but not all (e.g., Gronwall et al. 2007; Nilsson et al. 2009), narrowband surveys targeting compact Lyα emitters (LAEs) also detect blobs at similar redshifts. For example, while following up the two bright blobs found by Steidel et al. (2000) in the SSA22 field, Matsuda et al. (2004) discover a spectacular clustering of 35 blobs, in which the two brightest lie at the intersection of filaments traced by compact LAEs (Matsuda et al. 2005). Palunas et al. (2004) discover two additional Lyα blobs in the J2143–4423 region defined by an overdensity of compact LAEs (Francis et al. 2001). In order to directly measure the number density and clustering of Lyα blobs, and thus to constrain their halo mass, a large volume survey is required, particularly over different sight lines, to account for any field-to-field variations. Furthermore, one must apply uniform selection criteria for identifying blobs over the entire survey volume.

To acquire a large, unbiased sample of Lyα blobs at $z = 2.3$, we have pursued two complementary narrowband imaging surveys. The shallow, but wide sky coverage, survey using the Steward Observatory Bok 2.3 m + 90Prime imager targets rare, luminous Lyα blobs ($L_{\text{Ly}\alpha} \gtrsim 2 \times 10^{43}$ erg s$^{-1}$; Yang et al. 2009). In this paper, we report the first results from our deeper, but smaller sky coverage, survey with the NOAO 4 m telescopes and MOSAIC imagers that targets presumably more common intermediate size and luminosity Lyα blobs like...
those discovered by Matsuda et al. (2004). Using blob statistics from the four different 30′ × 30′ survey fields, in the Chandra Deep Field South (CDFS; Brandt et al. 2001), Chandra Deep Field North (CDFN; Giacconi et al. 2002), and two regions of the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007; Koekemoer et al. 2007), we determine the field-to-field variation in the blob number density. We then use a large volume cosmological N-body simulation to constrain their host halo masses for the first time.

In Section 2, we describe our narrowband imaging survey and the data reduction procedures. Section 3 describes the selection of the Lyα blob sample. In Section 4, we present the Lyα blob candidates (Section 4.1), compare their sizes and luminosities to those of compact Lyα sources (Section 4.2), constrain their field-to-field variation and halo masses (Section 4.3), and characterize their large-scale environment using the compact LAEs (Section 4.4). In Section 5, we discuss our conclusions. Throughout this paper, we adopt the cosmological parameters ΩM = 0.3 and ΩΛ = 0.7. All magnitudes are in the AB system (Oke 1974).

2. OBSERVATIONS AND DATA REDUCTION

Using the MOSAIC-I and II CCD imagers on the KPNO Mayall and the CTIO Blanco 4 m telescopes, we obtain deep narrowband images with a custom narrowband filter (hereafter NB403 or NB). This narrowband filter has a central wavelength of λc ≈ 4030 Å, designed for selecting Lyα-emitting sources at z ≅ 2.3. Its bandwidth of ΔλFWHM ≈ 45 Å provides a line-of-sight depth of Δz ≅ 0.037, corresponding to 46.8 Mpc at z = 2.3 in the comoving frame. The CDFS, CDFN, and COSMOS survey fields have extensive ancillary data sets, including the deepest X-ray images for robust identification of AGNs.

We conducted the narrowband imaging observations over nine photometric nights between 2007 January and 2009 February. The MOSAIC I and II cameras have eight 2k × 4k CCDS with a pixel scale of 0.′′7 pixel−1, leading to a sky coverage of 37′ × 37′. We obtain deep narrowband images for four different pointings: CDFS, CDFN, and two COSMOS subfields (hereafter COSMOS1 and COSMOS2). The total exposure time ranges from 7.2 to 10 hr, which consists of individual 20 or 30 minute exposures with a standard dither pattern to fill in the gaps between the eight chips. The seeing ranges from 1′′0 to 1′′3 depending on the fields. Table 1 summarizes our narrowband observations, including the central coordinates of the survey fields, total exposure times, survey areas, seeing, and survey depths.

Identifying line emission objects requires that we subtract the continuum emission underlying the NB403 bandpass. We estimate the continuum using existing, deep broadband (U and B) images. For the CDFS, we use optical images from the

Table 1

| Field  | R.A. (J2000) | Decl. (J2000) | Observing Date | Telescope | Exposure (hr) | Depth\(^a\) | Survey Area | Seeing \(^b\) | Pixel Scale \(^b\) |
|--------|-------------|--------------|----------------|-----------|--------------|------------|-------------|-------------|----------------|
| CDF-S  | 03:32:27.8  | −27:47:56    | 2007 Nov       | Blanco 4 m | 10           | 25.65      | 31′6 × 31′6 | 1′′0         | 0′′27         |
| CDF-N  | 12:36:50.9  | 62:11:48     | 2007 May       | Mayall 4 m | 10           | 25.27      | 29′5 × 29′5 | 1′′3         | 0′′30         |
| COSMOS1| 09:59:16.9  | 01:55:19     | 2009 Feb       | Blanco 4 m | 7.7          | 25.27      | 36′0 × 36′3 | 1′′2         | 0′′27         |
| COSMOS2| 09:59:16.8  | 02:31:19     | 2009 Feb       | Blanco 4 m | 7.2          | 25.25      | 36′0 × 36′3 | 1′′2         | 0′′27         |

Notes.
\(^a\) 5σ detection limit for 2′′ diameter aperture.
\(^b\) Pixel scales of final combined images, which are determined by the largest pixel scale between narrowband and broadband images.

Figure 1. Filter response profiles for the NB403 narrowband (dot-dashed line) and broadband (U and B) filters (solid lines) used in this study. The profiles are normalized to a maximum throughput of 1 and include the transmission of the atmosphere, telescope, camera optics, filter, and detector. Top: ESO 2.2 m U- and B-band filters from MUSYC (Gawiser et al. 2006a, 2006b). Middle: KPNO U and Subaru B filters from the Hawaii-HDF survey (Capak et al. 2004). Bottom: CFHT u′ and Subaru B filters from COSMOS (Capak et al. 2007).

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

Extended CDFS data set from the Multiwavelength Survey by Yale-Chile (MUSYC; Gawiser et al. 2006a, 2006b).\(^4\) For the CDFN, we use the imaging products from the Hawaii Hubble Deep Field North survey (Capak et al. 2004). For COSMOS1 and COSMOS2, we use 25 image tiles for each subfield\(^5\) produced by Capak et al. (2007). For CDFS and CDFN, the broadband images have smaller sky coverage than our narrowband images, limiting our final survey areas (Table 1). Table 2 lists details of the broadband images, including the filter properties, survey depth, seeing, and the instruments used. We also show the narrowband and broadband filter transmission curves for all fields in Figure 1.

\(^4\) The original data were taken with the ESO MPG 2.2 m and Wide Field Imager (WFI) by the ESO Deep Public Survey and COMBO-17 teams (Arnouts et al. 2001; Wolf et al. 2004; Erben et al. 2005; Hildebrandt et al. 2006). In this paper, we use the data products delivered by the MUSYC team.

\(^5\) The tile numbers for the lower-left and upper-right corners of COSMOS1 and COSMOS2 are (17, 69) and (65, 117), respectively.
We reduce the narrowband images with the IRAF mscd mosaic data reduction package (Valdes 1998) following the procedures of the NOAO Deep Wide Field Survey team (Jannuzi & Dey 1999). The data are corrected for crosstalk between amplifiers and bias-subtracted. For flat fielding, we use dome flats together with night-sky flats, which are median-combined from unregistered object frames each night. Satellite trails, CCD edges, bad pixels, and saturated pixels are masked. The astrometry is calibrated with the USNO-B1.0 catalog (Monet et al. 2003) using the IRAF ccmap task. The individual images are transformed to have the same pixel scales and world coordinate systems as the reference broadband images. For the COSMOS broadband images with finer pixel scale (0′′.15 pixel$^{-1}$), we resample them with the coarse MOSAIC-II pixel scale (0′′.27 pixel$^{-1}$) to make the reference images. Finally, the projected images are scaled using common stars in each frame and stacked to remove cosmic rays. For flux calibration, the projected images are scaled using common stars in each frame and stacked to remove cosmic rays. For flux calibration, the projected images are scaled using common stars in each frame and stacked to make the final products.

However, residual ghost images from imperfect correction can easily be confused with low surface brightness objects. In particular, for CTIO MOSAIC-II, the crosstalk images of a source at pixel $(x, y)$ on one chip appear at the same pixel coordinates on the other chips (victim chips). Therefore, ghost images always appear at the same location relative to real sources in the dithered exposures, and they will be co-added in the final combined images, leading to false detections mimicking diffuse extended emission.

To weed out false detections, we repeat the entire reduction procedure, doubling the crosstalk coefficients so that the ghost images are over-corrected and appear as negative counts. If a blob candidate is indeed an artifact arising from imperfect correction, it will show up as a negative image in the final combined image. Figure 2 shows an example in which one artifact image appears as a negative mirror image, whereas a real object is not affected by this crosstalk over-correction. Note that the artifact is as bright as the real blob candidate, demonstrating that special care is required to reject false detections for extremely low-surface brightness objects like Lyα blobs.

### 3. SELECTION OF Lyα BLOB CANDIDATES

To find Lyα blob candidates, we construct photometric catalogs in the NB403 narrow band and two broad bands ($U$ and $B$) using SExtractor (Bertin & Arnouts 1996). First, we make “detection” images $(NB+U+B)$ by adding the NB403 and broadband images after scaling them according to their signal-to-noise ratios ($S/N$). After identifying sources in the “detection” images that have least 2 pixels that are 1.5$σ$ above the local sky, we run SExtractor in double-image mode on the NB403, $U$, and $B$ images using these detection images. In other words, we first find the sources in the “detection” images and then obtain photometry at their position in the NB403, $U$, and $B$ images to construct three separate (one narrowband and two broadband) catalogs. We adopt Kron-like elliptical aperture magnitudes (i.e., $\text{Mag}_{\text{UKT0}}$ in SExtractor) to derive photometric properties. Our use of the “detection” images ensures that (1) all the sources detected in either NB or the broadband are included in our catalog and (2) the elliptical apertures of the more extended sources in the $(NB+U+B)$ image are large enough to include all the light from both the NB and broadband images.

Note that this choice of photometric aperture is different from the classical LAE searches that adopt a small circular aperture (a few × FWHM of seeing) to detect fainter sources.

The selection of Lyα blob candidates from the NB and broadband photometric catalogs consists of two steps: (1) selection for line (hopefully, Lyα) emitting objects with large line equivalent widths (EWs) and (2) selection for spatially extended objects with a larger angular extent in line emission than in the broadband bands. In other words, we define a “blob” as an object whose Lyα emission above a certain surface brightness threshold is more extended than its stellar continuum, thus representing light from the intergalactic medium.

First, we choose candidates by requiring that they are detected above the completeness limits of the NB403 images $(NB403 \leq 24.5 \text{mag})$. All candidates must have observed-frame EWs larger than 100 Å ($\text{EW}_{\text{rest}} > 30.3 \text{ Å}$), corresponding to $(UB - NB) > 1.2$, where $UB$ represents the AB magnitude of the average continuum flux density within the NB403 band ($f_{\text{cont}}^{\text{NB}}$) estimated from the $U$- and $B$-band images, $UB \equiv -2.5 \log(f_{\text{cont}}^{\text{NB}}) - 48.60$. We estimate $f_{\text{cont}}^{\text{NB}}$ and the line flux ($f_{\text{line}}$) of these objects using

### Notes.

| Field | Band | Effective Wavelength (Å) | Band Width (Å) | Depth (AB mag) | Seeing (arcsec) | Telescope and Instrument | Reference |
|-------|------|--------------------------|----------------|---------------|----------------|--------------------------|-----------|
| CDF-S | $U$  | 3505                     | 626            | 26.0$^a$      | 1.05           | ESO 2.2 m WFI            | Gawiser et al. (2006b) |
|       | $B$  | 4600                     | 915            | 26.9          | 0.95           | ESO 2.2 m WFI            |           |
| CDF-N | $U$  | 3648                     | 387            | 27.1$^b$      | 1.26           | KPNO MOSAIC I            | Capak et al. (2004)   |
|       | $B$  | 4428                     | 622            | 26.9          | 0.71           | Subaru Suprime-Cam       |           |
| COSMOS| $U$  | 3798                     | 720            | 26.4$^b$      | 0.90           | CFHT Megaprime           | Capak et al. (2007)   |
|       | $B$  | 4460                     | 897            | 27.3          | 0.95           | Subaru Suprime-Cam       |           |

$^a$ CDF-S: 5$σ$ detection limit corrected for infinite aperture.

$^b$ CDF-N and COSMOS 5$σ$ detection limit for 3′′ diameter aperture.
Therefore, we expect that the contamination of our $z = 2.3$ Lyα source catalog by nearby star-forming galaxies is minimal.

Second, we identify those line-emission selected objects that are more spatially extended in Lyα than their continuum counterparts (Figure 4). This selection definition is the same as that adopted by Matsuda et al. (2004) and somewhat different than that of Saito et al. (2006), who select spatially extended objects by requiring the FWHM in their intermediate band to be larger than that in the broadband or continuum image (see also Nilsson et al. 2009).

We measure the spatial extent of the Lyα emission in the continuum-subtracted images. After registering the NB403 and broadband images ($U$ and $B$) at the sub-pixel level and matching their seeing, we construct continuum-subtracted NB403 images by applying the relations in Equation (1) in two dimensions. We measure the isophotal area of the emission region by running SExtractor with a threshold of $5.5 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$, which corresponds to $3\sigma$, $1.8\sigma$, $2.2\sigma$, and $2.1\sigma$ above the local sky in the CDFS, CDFN, COSMOS1, and COSMOS2 fields, respectively. This measurement threshold is $\sim 2.5 \times$ higher than that adopted by Matsuda et al. (2004). However, because our survey redshift ($z = 2.3$) is lower than theirs ($z = 3.1$), we gain a factor of $\sim 2.4$ in surface brightness, thus achieving equivalent surface brightness ($8.6 \times 10^5 L_\odot$ kpc$^{-2}$ compared to their $8.2 \times 10^5 L_\odot$ kpc$^{-2}$). We also estimate local sky background using a fairly large background mesh size ($\sim 60'' \times 60''$), so as not to mistakenly subtract the extended Lyα emission as a local background.

Measuring the size of a low surface brightness feature is always subject to the noise and filtering. After testing the various smoothing filters in SExtractor, we adopt a $5 \times 5$ pixel$^2$ Gaussian kernel with FWHM = 2 pixels to compromise between S/N and over-smoothing. We choose this truncated smoothing kernel such that it can enhance S/N in the low surface brightness wings while not spreading the bright core into the outer part. We note that the measured isophotal area depends strongly on the choice of the filters and recommend specifying the smoothing kernel.
Figure 3. Color–magnitude ($UB-\text{NB}403$) vs. $\text{NB}403$ diagram for all the sources detected in our four survey fields: CDFS, CDFN, and two COSMOS subfields. Here $UB$ represents the AB magnitude of the average continuum flux density within the $\text{NB}403$ band estimated from the $U$- and $B$-band images. Right and top axes show the corresponding EWs in the observed frame and the $\text{NB}$ fluxes, respectively. We select line-emission objects with the criteria $\text{NB} < 24.5$ (solid line) and $(UB-\text{NB}403) > 1.2$ (dashed line; $\text{EW}_{\text{obs}} > 100 \text{ Å}$). Open circles represent the final Ly$\alpha$ blob candidates obtained from Figure 4. There are more open circles than final blob candidates in the CDFS panel because a few blob candidates include multiple sources. (A color version of this figure is available in the online journal.)
Figures 5–7 show postage-stamp images of all 25 blob candidates in the NB403 band, continuum-subtracted Lyα line emission, and three broad bands (UBR) overlaid with the Lyα contour corresponding to a surface brightness of $4 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$.

Although we select blob candidates quantitatively, follow-up visual inspection indicates that the spatial extents of four blobs are not clearly larger than the local point-spread function (PSF): CDFN-LAB02, CDFN-LAB03, COSMOS-LAB01, and COSMOS-LAB02. Therefore, we flag them as “marginal.” Excluding them does not affect our conclusion regarding the large field-to-field variation in the blob number density and, in fact, makes that case stronger. The largest blobs (more than 16$''$; the dashed line in Figure 4) are the most robustly extended and are comparable to those identified in other surveys (e.g., Matsuda et al. 2004). Therefore, we treat them separately as the “bright, large” subset of the entire blob sample in the clustering analysis in Section 4.3.

The blob candidates have a wide range of sizes and line luminosities: $A_{\text{iso}} = 10$–$60$ $''$ and $L_{\text{Ly}\alpha} = (0.7–8) \times 10^{43}$ erg s$^{-1}$. They show diverse morphologies ranging from compact (the southwest clump of CDFS-LAB01) to diffuse (CDFS-LAB05) to highly elongated (CDFS-LAB08 and 16). We do not find any bubble-like structures that might be associated with superwinds like those in a few of the blobs identified by Matsuda et al. (2004).

We are following up the entire blob sample spectroscopically. To date, we have observed 5 of the 16 blobs in the CDFS (CDFS-LAB01, 02, 04, 10, 14). We confirm all 5 spectroscopically with their Lyα and/or Hα lines. The details of our spectroscopic campaigns are presented in a forthcoming paper (Y. Yang et al. 2010, in preparation).

The properties of continuum objects associated with blobs provide valuable clues to the source of blob emission (Y. Yang et al. 2010, in preparation). For example, the blob discovered in the CDFS at $z = 3.1$ by Nilsson et al. (2006) is not associated with any continuum source within $\sim 3''$ (but see also Geach et al. 2009). Thus, it is possible that this blob is powered by cooling radiation or cold mode accretion. Although most blob candidates in our sample have clear continuum source counterparts in the rest-frame UV images, two blobs (CDFS-LAB04 and CDFS-LAB05) do not. We are, however, able to identify continuum sources within these blobs in deep, rest-frame UV Hubble Space Telescope (HST) images (GEMS and GOODS-S; Rix et al. 2004; Giavalisco et al. 2004). Preliminary inspection of the multiwavelength images also reveals bright IR or X-ray sources in these blobs. Interestingly, we often find multiple continuum...
sources in a blob, notably in LAB02 and LAB03. Our initial examination of the HST images of the 16 Lyα blobs in the CDFS often resolves these sources as galaxies, suggesting that star formation and/or nuclear activity might play a role in producing the Lyα emission (e.g., Colbert et al. 2006). This apparent clustering of sources within some blobs may indicate that blobs are the progenitors of groups or clusters of galaxies today. We test this possibility quantitatively in Section 4.3.

4.2. Continuous LAE-to-LAB Sequence

There is no evidence of a discontinuity between the properties of our blobs and unresolved, compact Lyα emitting (LAE) galaxies (i.e., the sources below the solid lines in Figure 4). The smoothness of the LAE-to-LAB transition, which has also been observed by Matsuda et al. (2004), suggests that these two types of Lyα sources may not be distinct and that whatever mechanism or mechanisms power blobs work over a wide range of luminosity and spatial extent. Understanding the origin of the extended Lyα emission requires us to probe the host galaxy properties—including the stellar mass, star formation rate, and size of the star-forming region—along the emitter-to-blob sequence. If extended star formation or AGNs are responsible for powering the emission, we would expect that the properties of blobs ( \( A_{\text{iso}}, L_{\text{Ly}\alpha} \) ) are correlated with either star formation rate or X-ray luminosity of the host galaxies. For example, Geach et al. (2005) argue that a correlation between the Lyα and bolometric (actually FIR) luminosity (although weak) suggests that the interaction of an ambient halo of gas with a galactic-scale superwind is responsible for the majority of LABs. We discuss the multi-wavelength properties of galaxies within or near the blobs, considering them as possible energy sources, in a separate paper (Y. Yang et al. 2010, in preparation).

4.3. Significant Differences in Blob Counts per Field

Surprisingly, most blobs (16/25) and all eight of the brightest, largest blobs ( \( A_{\text{iso}} > 16 \, \text{arcmin}^2, L_{\text{Ly}\alpha} \gtrsim 1.5 \times 10^{43} \, \text{erg s}^{-1} \) ) lie in only one of the survey fields: CDFS. Because the depth and seeing of the CDFS images are also superior to the other fields, we first need to verify that this field-to-field variation of the blob number density does not arise from selection effects. For example, although the separation between the fainter/smaller blob candidates and the simulated point-sources is distinct in the CDFS, the separations in other three fields are less clear.

To confirm that the field-to-field variation of the blob population is not due to the different seeing and survey depth, we simulate how the blob candidates in the CDFS field would look if observed in the other three fields. Using the continuum-subtracted postage stamp (27′′ × 27′′) images of the CDFS blob candidates, we first degrade their seeing to that of the CDFN, COSMOS1, and COSMOS2 fields by convolving kernels derived from PSF images. We re-bin the images (if pixel scale is different), add Poisson noise, and place them into empty sky regions in the continuum-subtracted narrowband images of the other fields. We then measure the blob sizes and luminosities in the same way as described in Section 3. We repeat this experiment ∼1000 times to derive the range of recovered luminosities and sizes.
Figure 5. Cut-out images of the 16 Lyα blobs that we discover in the extended CDFS. Images from left to right: $U$, NB403, continuum-subtracted Lyα line ($\lambda_{\text{c}} \simeq 4030$ Å), $B$, and $R$ band. The ticks are spaced in $10''$ intervals. The overlaid (yellow) contours represent a Lyα surface brightness of $4 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$, our $2.2\sigma$ detection limit in the narrowband images. In each case, the Lyα emission is more extended than the broadband counterpart. All five candidates that we have now observed spectroscopically are confirmed as blobs at $z = 2.3$ (Y. Yang et al. 2010, in preparation).

We show the isophotal areas and Lyα luminosities ($A_{\text{iso}} - L_{\text{Ly} \alpha}$) from this recovery test as green squares in the CDFN, COSMOS1, and COSMOS2 panels in Figure 4. If we estimate the number of possible detections as $N = \sum_{f_{\text{recv}} > 50\%} f_{\text{recv}}$, the
number of CDFS-like blobs that should have been detected as bright, large blobs with \( A_{\text{iso}} > 16 \, \mu\text{Jy} \) and \( L_{\text{Ly}\alpha} \gtrsim 1.5 \times 10^{43} \, \text{erg s}^{-1} \) is 6.9, 6.0, and 6.0, respectively, for each of the three other fields. However, we do not find such blobs in other fields, so we conclude that the observed field-to-field variation is real and that the minimum variation in the number counts of the largest blobs \( (A_{\text{iso}} > 16 \, \mu\text{Jy}) \) is at least \( N_{\text{blob}} = (N_{\text{CDFS}}, N_{\text{CDFS-N}}, N_{\text{COSMOS1}}, N_{\text{COSMOS2}}) = (6, 0, 0, 0) \). Throughout the rest of this paper, we refer to this as the “bright, large blob sample” and use it as the default for analyzing blob statistics. When we lower the criteria for the blobs to \( A_{\text{iso}} > 10 \, \mu\text{Jy} \), we expect 14.0, 13.9, and 13.7 blobs in the other three fields, respectively, whereas...
we find 5, 3, and 1 blob candidates. In this case, the observed contrast between the four survey fields is $N_{\text{blob}} = (14, 5, 3, 1)$. We subsequently refer to this as the “entire blob sample.”

We consider the bright, large sample and entire sample field-to-field variations in $N_{\text{blob}}$, which are corrected for the different survey conditions, in deriving the average number density and its variance in the following paragraphs. The blob statistics are summarized in Table 4.

Could this observed field-to-field variation arise solely from the statistical uncertainty? Before we proceed to a detailed analysis, we have to rule out the possibility that this result arises solely from Poisson (shot) noise. For the hypothesis that the surface density of the bright, large Lyα blob sample is $2-20$ blobs deg$^{-2}$ or $\sim 0.5-5$ blobs per survey field, we calculate the probability of observing non-detections in three survey fields and detections of more than six blobs in any of the fields. We are able to rule out a uniform distribution with at least 99.97% confidence ($\sim 3.6\sigma$). For the entire blob sample, a uniform distribution is excluded at the $3.7\sigma$ level. In Section 4.3.3, we compare the observed field-to-field variations with those derived from Poisson statistics and cosmological $N$-body simulations, respectively. We show that the latter better reproduces the observations.

### 4.3.1. Quantifying Blob Field-to-Field Variance

In this section, we estimate the field-to-field variation in the number density of Lyα blobs from the observed number statistics, $N_{\text{blob}}$. According to $\Lambda$CDM cosmology, the number density and variance of a galaxy population are not entirely independent properties, but a function of halo mass. Here, we treat them as independent parameters and aim to measure them as observables by adopting a simple analytic approximation of the underlying fluctuations in blob number density arising from large-scale structure. This method has the advantage that the blob number density and variance can be derived over a wider range of param-
eter space than sampled by simulations. We then compare these properties with the predictions from the simulations to obtain constraints on the halo mass (Section 4.3.3). As will be shown

The simplest way of quantifying the field-to-field variation, \( \sigma_v^2 \), is to adopt the relation (Peebles 1980, Section 36):

\[
\sigma_v^2 = \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle^2} - 1,
\]

where \( \langle N \rangle \) is the average number density per 0.25 deg\(^2\), the typical area of each of our four fields, and \( \sigma_v^2 \) is the fractional variance corrected for Poisson noise. The number density \( n \) can be derived directly from \( N \) by dividing it by the survey volume: \( n \equiv 8.6 \times 10^{-6} N \) Mpc\(^{-3}\). Hereafter, we use surface density to mean surface number density. Thus, \( \sigma_v^2 \) represents the fractional uncertainty in the observational estimate due to finite survey volume. For the bright, large sample, we obtain \( \langle N \rangle \sim 1.5 \) and \( \sigma_v \sim 1.5 \) (150\%). For the entire blob sample, \( \langle N \rangle \sim 5.75 \) and \( \sigma_v \sim 0.76, \) or 76\%. Because of our small number statistics, we choose to adopt a more sophisticated method to better understand the possible range of blob number density and its variance.

To quantify the field-to-field variance of the blob population, we calculate the posterior probability for \( \sigma_v^2 \) and an average surface density per 0.25 deg\(^2\) (\( \bar{N} \)) given our observation (\( D \)) of (6, 0, 0, 0) blobs in the bright, large sample:

\[
p(\sigma_v, \bar{N}|D) \propto \text{prob}(D|\sigma_v, \bar{N}) p(\sigma_v, \bar{N}).
\]

First, we assume that the field-to-field variance follows the lognormal distribution: \( \langle N/\bar{N} \rangle \sim \log \sim \mathcal{N}(0, \sigma_v^2) \); in other words, \( \log(N/\bar{N}) \) follows a normal distribution \( \mathcal{N}(0, \sigma_v^2) \). Here, \( \sigma_v^2 \) is the variance of the lognormal distribution and is related to the actual variance by \( \sigma_v^2 = \exp(\sigma_v^2) - 1 \). Unlike a Gaussian distribution, the lognormal distribution does not allow negative values for \( N \) and naturally introduces a skewness into the distribution. When \( \sigma_v \ll 1 \), the lognormal distribution is similar to a Gaussian, but the distribution becomes skewed toward zero as \( \sigma_v \) increases, effectively mimicking the dark matter fluctuations at the high mass end (Coles & Jones 1991; Bernardeau & Kofman 1995). We choose the lognormal distribution for the simplicity here, but any reasonable functional form capable of representing this skewness can be used.

Second, for a given set of \( \langle \sigma_v, \bar{N} \rangle \), we calculate the probability, \( \text{prob}(D|\sigma_v, \bar{N}) \), of finding six blobs in one field and none in three other fields assuming that the observed number of blobs follows Poisson statistics with a mean of \( \bar{N} \). We adopt logarithmic priors for both \( \sigma_v \) and \( \bar{N} \), which indicate \( p(\sigma_v) \propto 1/\sigma_v \) or \( p(\bar{N}) \propto 1/\bar{N} \), implying that the scale of \( \bar{N} \) and \( \sigma_v \) is unknown, i.e., that the priors are uniform in logarithmic bins. We consider a range of 0.1 < \( \bar{N} \) < 30 and 0.1 < \( \sigma_v \) < 10, i.e., 10%–100% field-to-field variance. We also test other priors including (1) a linear prior for both \( \sigma_v \) and \( \bar{N} \) and (2) a logarithmic prior for \( \sigma_v \) and linear prior for \( \bar{N} \), but the choice of prior does not affect our conclusions.

Figure 8 shows the posterior probability distribution of the average surface density and variance for the bright, large blob sample (left) and the entire sample (right). For the bright, large blobs, the posterior favors high variance (\( \sigma_v \gg 1 \)) as expected. The confidence regions are not closed, allowing us to put only a lower bound on \( \sigma_v \). The lower limits are \( \sigma_v > 1.45 \) and \( \sigma_v > 0.57 \) for the 1\( \sigma \) and 2\( \sigma \) confidence levels, respectively, for the joint distribution (i.e., we attempt to constrain both \( N \) and
The uncertainties of \( \sigma_v \) regardless of the estimated number density, confirming that blobs are highly clustered. For the entire blob survey, the most probable estimate is \((\bar{N}, \sigma) \approx (4.8, 0.9)\). (A color version of this figure is available in the online journal.)

The strong variation in blob counts from one 0.25 deg\(^2\) field to another is consistent with the discovery of a close pair of blobs by Yang et al. (2009). Those blobs were among only four detected in our shallower, but larger (4.8 deg\(^2\)) survey of the NOAO B"ootes field. They are separated by only \( \sim 70'' \). If that survey had been conducted by mosaicking 30' \times 30' fields like those sampled by our imager here, most fields would not contain the pair. On a related note, the clustering of 35 blobs in the SSA22 overdensity (Matsuda et al. 2004) was identified during a follow-up of two giant blobs originally found by Steidel et al. (2000).

4.3.2. Blob Number Density and Luminosity Function

The strong field-to-field variation of the Ly\(\alpha\) blobs presents challenges for the measurement of their number density and luminosity function (LF). In this section, we compare our LFs derived from each survey field with each other and with previous studies. We demonstrate that a large volume survey and/or multiple pointings are critical to constrain the blob LF. For the rest of the paper, we adopt the marginalized number densities from the previous section, which give us the number densities of \( n \approx 1.0_{-0.6}^{+1.8} \times 10^{-5} \) Mpc\(^{-3}\) (from \( N = 1.2 \) per 0.25 deg\(^2\)) for the bright, large blob sample and \( n = 4.1_{-1.6}^{+4.8} \times 10^{-5} \) Mpc\(^{-3}\) (from \( N = 4.8 \) per 0.25 deg\(^2\)) for the entire blob sample.

First, we compare the number density of bright, large blobs derived from Section 4.3.1 with previous measurements (Figure 9). To make a fair comparison with other work, we need to cut the bright, large sample to satisfy the selection criteria used by Yang et al. (2009) when they previously examined the blob number densities among different samples spanning \( z = 0.8-6.6 \). Those criteria were \( L_{\text{Ly}\alpha} > 1.5 \times 10^{43} \) erg s\(^{-1}\) and \( A_{\text{iso}} > 25 \) cm\(^{-2}\) above a surface brightness threshold of \( 5 \times 10^{-18} \) erg s\(^{-1}\) cm\(^{-2}\) arcsec\(^{-2}\) under \( \sim 1''8 \) seeing. We refer readers to Yang et al. (2009) for the details about the estimates at each redshift.

Only four blobs (CDFS-LAB01, 02, 03, and 04) from the sample of six bright, large blobs satisfy the Yang et al. (2009) criteria.
the number density estimates from Matsuda et al. (2004), Saito et al. (2006), but deeper, NOAO 4 m survey (this work), respectively. The filled dots are open squares are from our two narrowband imaging surveys at also lie in an overdense region (see also Palunas et al. 2004; Prescott et al. 2008) and occupy high mass halos. On the other hand, the blob number densities for the other three fields are lower than for the CDFS and the SSA22 field, demonstrating that characterizing the evolution in n with z requires surveys large enough to overcome the field-to-field variance. Because such measurements are not yet available at other redshifts, it is not possible to constrain n(z) at this time.

Note that we are not able to apply the Yang et al. (2009) selection criteria to the higher-z Saito et al. (2006) and Ouchi et al. (2009) blob samples because those authors use selection methods different than ours and Matsuda et al. (2004). Therefore, we plot these higher-z estimates as upper and lower limits, respectively. Unlike the comparison between the z = 2.3 and z = 3.1 samples, comparison of the Saito et al. (2006) and Ouchi et al. (2009) points with our results is perilous. We plot all these points together only to summarize the state of blob surveys.

In addition to the blob statistics obtained from narrowband or intermediate-band imaging techniques, we show the number density estimate (n < 0.5 × 10^{-6} Mpc^{-3}) from GALEX slitless spectroscopy of two galaxy (super)clusters at z ≈ 0.82 (Keel et al. 2009). This n, even though calculated from overdense regions, is well below that of the overdense CDFS and Matsuda et al. (2004) fields, supporting the Keel et al. (2009) claim that Lyα blobs might be high redshift phenomena. Additional surveys at z < 1 are required to confirm this result, as we do not know whether the variance in n at z = 2.3 persists at lower z.

Second, we compare our blob LF with that of Matsuda et al. (2004) in Figure 10. The similar depth and selection criteria of the two surveys make the direct comparison meaningful (see the Appendix). For each of our blobs, we use the total Lyα luminosity (Table 3), which our recovery test demonstrates is reliable. The isophotal flux could be biased due to the different S/N and filtering kernels used in the detection procedure. We construct an average LF as follows. In each Lyα luminosity bin, we divide the number of blobs found in all four of our survey fields by the total survey volume. We also plot the LF obtained from only the 16 blobs in the CDFS. We do not apply a completeness correction using the recovery fraction f_{rec} in Table 3, because we also do not know the completeness of Matsuda et al. (2004) sample. In Table 6, we list our blind survey LFs as well as that of the SSA22 overdense region (Matsuda et al. 2004).

The slope and normalization of the CDFS (squares) and Matsuda et al. (2004; diamonds) LFs agree roughly. The apparent discrepancy in the faintest bin (log L_{Lyα} = 42.8) is likely due to incompleteness in our survey. We also do not find any blobs as bright as those in their brightest bin (log L_{Lyα} = 44.0). Not surprisingly, the normalization of our “average”

\[ n = 0.66_{-0.4}^{+1.2} \times 10^{-5} \text{ Mpc}^{-3} \] (large, open square in Figure 9). The error bar takes into account the field-to-field variation as well as the Poisson noise. This average number density at z = 2.3 is derived from the four NOAO 4 m survey fields is thus consistent with the Yang et al. (2009) measurement at the same redshift (n = 0.25 ± 0.10) × 10^{-5} Mpc^{-3}; large, filled square) obtained from our ~4 x larger volume survey of Boötes.

To illustrate the uncertainties arising from the field-to-field variance, we show the bright, large number density for each of the four fields individually. The small squares represent n = 3.2 (± 1.6) × 10^{-5} Mpc^{-3} for the CDFS, and three upper limits, n < 0.90, 0.62, 0.60 × 10^{-5} Mpc^{-3} for CDFN, COSMOS1, and COSMOS2, respectively. The number density in the CDFS field is consistent with that found by Matsuda et al. (2004) in the SSA22 field (5.8 ± 2.1 × 10^{-5}) within the uncertainties, suggesting that blobs at z = 2.3 in the CDFS also lie in an overdense region (see also Palunas et al. 2004;
LF (filled circles) is ∼4× lower than those of the CDFS and Matsuda et al. (2004) LFs. Once again, we see that comparisons among surveys are difficult without knowledge of the blob clustering strength and that large volume surveys are required to overcome the strong field-to-field variance of the blob counts. The same conclusion is reached from the on-going Subaru survey of blobs (see figure 13 of Goerdt et al. 2009, Y. Matsuda 2010, in preparation).

The clustering of blobs provides a means to discriminate among models for the origin of the extended Lyα emission, including photoionization by AGNs, outflows due to intense star formation, and cooling radiation. This analysis is beyond the scope of this paper, but we note that any blob-producing mechanism must reproduce both the observed LF and its field-to-field variation. For example, it would be interesting to ascertain whether AGNs or sub-millimeter galaxies have clustering strengths similar to blobs. Along these lines, Dijkstra & Loeb (2009) claim that Lyα blobs are cooling radiation arising from cold streams falling onto the embedded galaxies and that a strong variance naturally arises from the underlying variation of dark matter halos (see also Goerdt et al. 2009).

4.3.3. Blob Halo Masses

Because our blob survey is blind and sufficiently large to determine the blob number density and its variance, we have a unique opportunity to constrain the properties of the dark halos in which the blobs reside, and thus to understand what these mysterious objects have evolved into today. For example, based on the blob number density and the discovery of the blob pair in the NOAO Deep-Wide Boötes field, Yang et al. (2009) suggest that blobs are sites for the formation of the brightest galaxies in rich galaxy clusters. However, the small number statistics of that study precluded constraining the mass of blob halos. Although our current survey statistics are still not large enough to directly measure the clustering of blobs via correlation function analysis, we can use \( n, \sigma_v \), and a cosmological \( N \)-body simulation of the \( \Lambda \)CDM universe to identify the most likely halo mass occupied by blobs.

We first select a dark halo (DM) mass in which blobs could reside. All halos above this minimum mass \( M_{\text{min}} \) have a fixed probability of containing a detectable blob, which is labeled the detectability fraction \( f_D \). We choose \( f_D \) such that the halo mass function from the simulation reproduces the observed blob number density from Section 4.3.1. Once the halo mass and detectability fraction are fixed, we can predict the clustering of such halos, i.e., the field-to-field variation, directly from the simulation using counts-in-cells (c-in-c) methodology. Then we compare this prediction to the observed variation in blob counts over the four survey fields to quantify the likelihood that the selected halo mass and detectability fraction reproduce the observed blob statistics in Table 4.

To link the number density and its variance to the DM halo mass, we employ a simple c-in-c analysis using the halo catalog at \( z = 2 \) derived from the ABACUS N-body code (M. Metchnik & P. A. Pinto 2010, in preparation). This simulation has a cubic volume of \( 1 \ h^{-1} \) co-moving Gpc on a side and \( 1024^3 \) dark matter particles with \( n_{\text{DM}} = 1.1 \times 10^{-11} \, \text{M}_\odot \). We adopt the cosmological parameters: \( \Omega_0 = 0.701, n_s = 0.96, \Omega_M = 0.279, \Omega_b = 0.0456, \) and \( \sigma_8 = 0.817 \). Dark halos are defined using a friends-of-friends algorithm with linking length \( b = 0.16 \) in units of the initial particle spacing. The smallest halos used in our analysis consist of 48 particles. Due its large size, this simulation is finely tuned to our problem, which requires sampling many “cells,” ~50comoving Mpc boxes that are roughly the same size and geometry as our survey of each of the four fields.

To constrain the DM halo mass of the blobs, we first consider the observed bright, large blob number density \( N = 1.2 \) per 0.25 \( \text{deg}^2 \) field, or \( n = 1.0 \times 10^{-5} \, \text{Mpc}^{-3} \) (1.5 blobs in a 50 Mpc box). If all halos contain a detectable blob (i.e., the detectability fraction \( f_D = 100% \)), then this number density requires halos with more than 150 DM particles or \( 1.1 \times 10^{13} \, \text{M}_\odot \). We derive the c-in-c distribution of the simulated blobs by counting the number of blobs within 10,000 randomly placed 50 Mpc boxes and by assuming a simple halo occupancy distribution (e.g., Berlind & Weinberg 2002) with \( N_h(M \geq M_{\text{min}}) = 1 + (M/M_1)^\alpha \), where \( M_{\text{min}} = 150 \) DM particles or \( 1.7 \times 10^{13} \, \text{M}_\odot \), \( M_1 = 1000 \) DM particles or 1.1 ×

### Table 6

| log(\( L_{\text{Ly} \alpha} \)) | \( n(> L) \times 10^{-5} \, \text{Mpc}^{-3} \) |
|-----------------------------|------------------------------------------|
| CDFS\(^a\)                  | All\(^b\)                                | SSA22\(^c\)       |
| 43.90                       | ...                                      | 1.54 ± 1.09       |
| 43.70                       | 0.83 ± 0.83                              | 2.31 ± 1.33       |
| 43.50                       | 2.50 ± 1.44                              | 3.08 ± 1.54       |
| 43.30                       | 3.33 ± 1.67                              | 5.38 ± 2.04       |
| 43.10                       | 5.83 ± 2.20                              | 10.77 ± 2.88      |
| 42.90                       | 12.49 ± 3.22                             | 18.46 ± 3.77      |
| 42.70                       | 13.32 ± 3.33                             | 26.92 ± 4.55      |

Notes.

\(^a\) LF from blobs in CDFS.

\(^b\) LF from all four survey fields (CDFS, CDFN, COSMOS1, and COSMOS2).

\(^c\) We also list the LF from Matsuda et al. (2004) for comparison.
We derive the variance of this distribution, \( \sigma_v \). Using Equation (2), we test the null hypothesis that the halos and blobs are not correlated.

The posterior probability is much higher than Poisson, is it statistically acceptable as a halo mass constraint? To test if we can reject the posterior probability of 1.3%, we consider an extreme case that maximizes our posterior statistics: \( P = 4p^3(1 - p) \), the probability of not finding any blob in three survey fields, but finding any number of blobs in the one remaining field. Here, \( p \) represents the probability of finding zero blobs in one survey field and \( P \) has a maximum value of 42.2% when \( p = 0.75 \). Therefore, we should compare our posterior probability of 1.3% with this extreme case, not with 100%. The posterior probability is only \( \sim 34 \times \) smaller than this maximum probability. Therefore, we conclude that our measurements are consistent with the halo model assuming \( M_{\text{halo}} = 10^{13} M_\odot \) and \( f_D = 100\% \).

We have now established that the observed strong field-to-field variation in blob counts is not surprising if massive dark matter halos with \( M_{\text{halo}} \gtrsim 10^{13} M_\odot \) always produce detectable Ly\( \alpha \) blobs. Here, we investigate whether it is possible to obtain consistency with lower values of halo mass by changing the blob detectability fraction. We derived the maximal halo mass for the observed blob number density by assuming that all halos more massive than \( M_{\text{halo}} = 1.7 \times 10^{13} M_\odot \) have detectable blobs. If we lower \( f_D \) to 50\%, 25\%, and 12.5\%, while increasing the number density of halos capable of hosting blobs by two, four, and eight times to keep the abundance of the observed blobs constant, the threshold halo mass \( M_{\text{min}} \) decreases to 1.2, 0.8, and \( 0.5 \times 10^{13} M_\odot \), respectively. At lower halo mass, we naturally obtain a weaker field-to-field variation and a less prominent high-end tail in the c-in-c distribution (i.e., at \( N > 6 \)) than for \( f_D = 100\% \). To illustrate this trend, we also show the c-in-c distributions for \( f_D = 12.5\% \) (thick solid line) in Figure 11.

Figure 11 shows the simulated bright, large blob c-in-c distribution with an average of 1.46 blobs per cell. The overlaid line represents a Poisson distribution with the same average, a reference case in which the DM halos containing blobs are not clustered. Thus, comparing the bright, large and Poisson distributions tests the null hypothesis that the halos and blobs are not correlated.

As expected from the large field-to-field variation (Section 4.3.1), the low and high tails of the simulated distribution exceed the Poisson counts. Using the Equation (2), we derive the variance of this distribution, \( \sigma_v = 1.04 \), which is consistent with our estimate (\( \sim 1.4\sigma \) lower limit) from the previous section (Section 4.3.1). This variance corresponds to a bias of \( b = \sigma_{\text{blob}}/\sigma_{\text{PM}} \sim 7 \) given the rms fluctuation of mass (\( \sigma_{\text{PM}} = 0.15 \)) within a sphere of 31 Mpc radius with the same volume as our survey box.

Is this simulated distribution consistent with the observed field-to-field variation in \( n \)? For \( f_D = 100\% \) (shaded histogram), the probability of finding no blobs in a cell is as high as 41.5\% (\( \sim 2\times \) greater than the Poisson probability), while the probability of finding six or more blobs is 4.4\% (\( \sim 11 \times \) greater than Poisson). The posterior probability of finding at least six blobs in one survey field with non-detections in three other fields (the case for the bright, large blobs in our sample) is 1.3\% (\( 4 \times 0.415^3 \times 0.044 \)). This posterior probability is \( \sim 65 \times \) times larger than 0.02\% for the Poisson distribution. Therefore, we reject the null hypothesis that the dark matter halos and blobs are uncorrelated.

\( 10^{14} M_\odot \), and \( \alpha = 1 \) as a fiducial value. This cell size is similar to our survey volume, \( 48.7 \times 48.7 \times 46.8 \) Mpc, for the 0.25 deg\(^2\) field of view of the NB403 narrowband filter. While the four survey fields have slightly different survey dimensions, the choice of a 50 Mpc box does not affect our conclusions.

Figure 11 shows the simulated bright, large blob c-in-c distribution with an average of 1.46 blobs per cell. The overlaid line represents a Poisson distribution with the same average, a reference case in which the DM halos containing blobs are not clustered. Thus, comparing the bright, large and Poisson distributions tests the null hypothesis that the halos and blobs are not correlated.

As expected from the large field-to-field variation (Section 4.3.1), the low and high tails of the simulated distribution exceed the Poisson counts. Using the Equation (2), we derive the variance of this distribution, \( \sigma_v = 1.04 \), which is consistent with our estimate (\( \sim 1.4\sigma \) lower limit) from the previous section (Section 4.3.1). This variance corresponds to a bias of \( b = \sigma_{\text{blob}}/\sigma_{\text{PM}} \sim 7 \) given the rms fluctuation of mass (\( \sigma_{\text{PM}} = 0.15 \)) within a sphere of 31 Mpc radius with the same volume as our survey box.

Is this simulated distribution consistent with the observed field-to-field variation in \( n \)? For \( f_D = 100\% \) (shaded histogram), the probability of finding no blobs in a cell is as high as 41.5\% (\( \sim 2\times \) greater than the Poisson probability), while the probability of finding six or more blobs is 4.4\% (\( \sim 11 \times \) greater than Poisson). The posterior probability of finding at least six blobs in one survey field with non-detections in three other fields (the case for the bright, large blobs in our sample) is 1.3\% (\( 4 \times 0.415^3 \times 0.044 \)). This posterior probability is \( \sim 65 \times \) times larger than 0.02\% for the Poisson distribution. Therefore, we reject the null hypothesis that the dark matter halos and blobs are uncorrelated.

Although the posterior probability is much higher than Poisson, is it statistically acceptable as a halo mass constraint? To test if we can reject the posterior probability of 1.3\%, we consider an extreme case that maximizes our posterior statistics: \( P = 4p^3(1 - p) \), the probability of not finding any blob in three survey fields, but finding any number of blobs in the one remaining field. Here, \( p \) represents the probability of finding zero blobs in one survey field and \( P \) has a maximum value of 42.2\% when \( p = 0.75 \). Therefore, we should compare our posterior probability of 1.3\% with this extreme case, not with 100\%. The posterior probability is only \( \sim 34 \times \) smaller than this maximum probability. Therefore, we conclude that our measurements are consistent with the halo model assuming \( M_{\text{halo}} = 10^{13} M_\odot \) and \( f_D = 100\% \).

We have now established that the observed strong field-to-field variation in blob counts is not surprising if massive dark matter halos with \( M_{\text{halo}} \gtrsim 10^{13} M_\odot \) always produce detectable Ly\( \alpha \) blobs. Here, we investigate whether it is possible to obtain consistency with lower values of halo mass by changing the blob detectability fraction. We derived the maximal halo mass for the observed blob number density by assuming that all halos more massive than \( M_{\text{halo}} = 1.7 \times 10^{13} M_\odot \) have detectable blobs. If we lower \( f_D \) to 50\%, 25\%, and 12.5\%, while increasing the number density of halos capable of hosting blobs by two, four, and eight times to keep the abundance of the observed blobs constant, the threshold halo mass \( M_{\text{min}} \) decreases to 1.2, 0.8, and \( 0.5 \times 10^{13} M_\odot \), respectively. At lower halo mass, we naturally obtain a weaker field-to-field variation and a less prominent high-end tail in the c-in-c distribution (i.e., at \( N > 6 \)) than for \( f_D = 100\% \). To illustrate this trend, we also show the c-in-c distributions for \( f_D = 12.5\% \) (thick solid line) in Figure 11.

Table 7 summarizes the c-in-c statistics, the field-to-field variance (\( \sigma_v \)), and the posterior probability of observing six or more blobs in only one of the survey fields for the different \( f_D \) values. For comparison, we also list the c-in-c statistics and the posterior probabilities for the Poisson distributions.

Higher detectability fractions (and more massive halos) produce a field-to-field variance about the observed \( n \) that is more consistent with the observed variance. Halo models with \( \gtrsim 10^{13} M_\odot \) halos and \( f_D \gtrsim 50\% \) work best. Lower detectability fractions (e.g., \( \sim 12\% \)) require somewhat lower halo masses (\( \sim 5 \times 10^{12} M_\odot \)) to reproduce \( n \), but the resulting variance in \( n \) is lower and further from the observed value. However, the effects of lowering the detectability fraction and halo mass are not large enough to put strict lower limits on the halo mass: for the lowest \( f_D \) considered, 12.5\%, we predict \( \sigma_v = 0.76 \) (76\%) and a posterior probability of 0.45\%, which is only \( \sim 3 \times \) lower than in the \( f_D = 100\% \) case. It is possible, but less likely, that the blobs occupy a few \( \times 10^{12} M_\odot \) halos if \( f_D \) is much lower than 12.5\%.

A detectability fraction of \( f_D = 12.5\% \) implies a short blob lifetime, only \( \tau \lesssim 350 \) Myr at \( z = 2–3 \). In principle, we could adopt still lower values (\( f_D \ll 10\% \)), down to the limit where the Ly\( \alpha \) blobs live only a few tens of Myr, and thus lower halo masses (\( \lesssim 10^{12} M_\odot \)), in order to find the point at which we can reject the assumed halo mass. However, the
limited mass resolution of our N-body code does not allow us to resolve smaller halos. Therefore, to put tighter constraints on the halo mass requires improving the blob statistics by extending surveys to larger volumes and/or creating higher resolution simulations.

For now, we conclude that bright, large Lyα blobs are most likely to reside in massive dark halos with $\gtrsim 10^{13} M_\odot$ that have detectable blobs more than $\sim 50\%$ of the time. Interestingly, these halo mass estimates agree with the dynamical masses, $M_{\text{dyn}} = 10^{12} - 10^{13} M_\odot$, derived from the width of Lyα lines in similar blobs (Matsuda et al. 2006). However, special caution is required in using the Lyα line width as a mass proxy because the radiative transfer effects on the line width are poorly understood.

For the entire blob sample (with counts of 14, 5, 3, 1 for the four survey fields), the required halo mass for $f_D = 100\%$ is $M_{\text{min}} = 0.8 \times 10^{13} M_\odot$. The fractional variance is $\sigma_v = 0.76$, again consistent with the observed value. Because the halo mass function is steep at the high mass end, $M_{\text{min}}$ for the entire sample is similar to that derived for the bright, large blobs alone. Therefore, the resulting halo mass is insensitive to the blob selection cuts and the statistics for our entire sample are consistent with blobs occupying halos of $\sim 10^{13} M_\odot$.

Halos of $\sim 10^{13} M_\odot$ lie in the high mass tail of the halo mass distribution at $z = 2.3$. How massive will these halos be today? Because our simulations were not run beyond $z = 1$, we consider the N-body simulation of Maccio et al. (2008), which extends to $z = 0$. Their $M_{\text{halo}} \gtrsim 10^{13} M_\odot$ halos grow in mass by $2$–$10 \times$ (with an average factor of $5.2 \pm 2.4$ increase) from $z = 2.3$ to now. Therefore, it is likely that blobs are sites for the formation of the brightest galaxies in what will become the typical halos of rich clusters ($M_{\text{halo}} \sim 10^{14} M_\odot$) at the present epoch.

### 4.4. Large-scale Environment of Blobs

The clustering of Lyα blobs in the CDFS and the inferred large mass of their individual halos imply that blobs inhabit overdense regions. Here, we test this hypothesis using the much larger population of compact LAEs in the CDFS to trace large-scale structure over tens of comoving Mpc and thus characterize the blob environment independently. Matsuda et al. (2005) show that their blobs lie near the intersection of large-scale filamentary structures in the SSA22 overdensity. Prescott et al. (2008) use the surface number density of LAEs to show that a giant blob, originally identified via its strong MIR emission (Dey et al. 2005), resides in a region $\sim 3 \times$ more dense than the edge of their survey field. We apply c-in-c methodology to the LAE spatial distribution in the CDFS to quantify the scale over which any structure is coherent. We then identify over-densities of LAEs relative to their average number density in the field and compare them to the spatial distribution of blobs.\(^8\)

We select a sample of LAEs in the CDFS ($N \sim 200$ with $NB < 25.0$, excluding the blobs) following the first step in our blob selection procedure (Section 3), but using a $2' \times 2'$ diameter aperture to maximize the S/N of fainter point sources. In Table 1, we list the $5\sigma$ limiting magnitudes, which are determined by measuring fluxes within the randomly placed apertures in the sky background region.

Using this LAE catalog, we test whether there is structure over various spatial scales. For each scale, we count the number of LAEs within circular cells of that scale radius randomly placed over the field. Then we compare this c-in-c distribution with a Poisson distribution using Kolmogorov–Smirnov statistics to check if this distribution deviates from Poisson noise. We repeat this test with different cell sizes ranging from 1' to 8'. The c-in-c distributions deviate significantly (at $> 95\%$ confidence) from Poisson noise for cells larger than 5', indicating coherent large-scale structure over scales of at least $\sim 8$ comoving Mpc.

Figure 12 shows the LAE distribution in the CDFS field. To estimate the overdensity of LAEs, we overlay the surface overdensity $\delta_{\Sigma} = (\Sigma - \bar{\Sigma})/\Sigma$ contours on the map. Here $\Sigma$ represents the average surface density estimated from all the LAE candidates over the entire field. The overdensity maps are smoothed using the adaptive kernel smoothing method developed by Ebeling et al. (2006). The FWHMs of the Gaussian filters adopted by this algorithm range from 2' to 5' depending on the local surface density of LAEs within in the field. We identify a belt-like large-scale structure elongated from north to south. The blob candidates are preferentially located within or at the boundary of this structure. Therefore, we conclude that the Lyα blobs in the CDFS indeed trace overdensities in the early universe.

Recently, Salimbeni et al. (2009) systematically searched for galaxy overdensities within the GOODS-South field up to $z \sim 2$ using photometric redshifts that have an rms accuracy of $\Delta z / (1 + z) = 0.03$. In Figure 12, we show their three overdensities (crosses) with redshifts similar to our survey ($z_{\text{phot}} \simeq 2.23$–$2.28$). While the GOODS-South field does not include the densest part of our surface density map, their overdensities each include 19–23 members and are located near the boundary of the LAE overdensity. Spectroscopic confirmation of this system is required to relate these systems with the structures revealed by our Lyα blobs and emitters.

### Table 7

| $f_D$ | $M_{\text{min}}(M_\odot)$ | $\sigma_v$ | $P(N = 0)$ | $P(N \geq 6)$ | Probability | $P(N = 0)$ | $P(N \geq 6)$ | Probability |
|-------|--------------------------|------------|------------|---------------|-------------|------------|---------------|-------------|
| 100%  | $1.66 \times 10^{13}$     | 1.03       | 41.5%      | 4.40%         | 12.5%       | 23.3%      | 0.440%        | 0.022%      |
| 50%   | $1.16 \times 10^{13}$     | 0.93       | 38.5%      | 3.65%         | 0.83%       | 22.9%      | 0.370%        | 0.018%      |
| 25%   | $7.97 \times 10^{12}$     | 0.83       | 35.8%      | 3.27%         | 0.60%       | 23.4%      | 0.300%        | 0.015%      |
| 12.5% | $5.32 \times 10^{12}$     | 0.76       | 35.6%      | 2.51%         | 0.45%       | 23.8%      | 0.370%        | 0.020%      |

**Notes.** Column 1: Detectability fraction, Column 2: dark matter halo mass, Column 3: variance (Equation (2)), Column 4: probability of non-detection, Column 5: probability of finding six or more blobs, Column 6: posterior probability of finding six or more blobs in one field and none in the three other fields, Columns 7–9: same as Columns 4–6 but for Poisson distribution.

\(^8\) We do not detect any significant large-scale structure in the other three survey fields. Because survey depth and seeing vary over the fields, it is not clear whether the non-detections arise from a real absence of structure or poorer sensitivity.
Salimbeni et al. (2009) using photometric redshifts (traced by the LAEs. The three crosses represent overdensities identified by density of the whole region. The Ly$\alpha$ of at least $\sim 10^{14} M_\odot$ blobs are discovered only in CDFS, indicating a strong field-to-field variation. Using a simple analytic approximation for the underlying fluctuations of the blob number density, we find that these large, bright blobs have a field-to-field variance of $\sigma_n \gtrsim 1.5 (150\%)$ about their number density of $n \sim 1.0^{+1.8}_{-0.6} \times 10^{-3} \text{Mpc}^{-3}$. This variance is large, significantly higher than that of unresolved Ly$\alpha$ emitters ($\sigma_n \sim 0.3$ or 30%).

To constrain the mass of the dark matter halo around each Ly$\alpha$ blob, we compare the number density and clustering of blobs with the in-c distribution of halos predicted from a 1 $h^{-1}$Gpc cosmological N-body simulation. At $z = 2.3$, $n$ implies that bright, large blobs could occupy halos of $M_{\text{halo}} \gtrsim 10^{13} M_\odot$, if most halos contain a detectable blob, $n$, the detectability fraction is $\gtrsim 50\%$. The predicted variance in $n$ is consistent with that observed and corresponds to a bias of 7. Lower detectability fractions (e.g., $\sim 10\%$) require somewhat lower halo masses ($\sim 5 \times 10^{12} M_\odot$) to reproduce $n$, but the resulting variance is lower and further from the observed value. Blob halos lie at the high end of the halo mass distribution at $z = 2.3$ and are likely to evolve into the $\sim 10^{14} M_\odot$ halos typical of galaxy clusters today.

The clustering and the inferred halo mass of blobs suggest that they lie in overdense environments. The spatial distribution of LAEs confirms this hypothesis: a c-in-c analysis of the CDFS reveals coherent large-scale structure over scales of at least $\sim 8$ comoving Mpc where both the LAEs and blobs cluster.

Given the strong field-to-field variance of Ly$\alpha$ blobs, one must be cautious in comparing blob number densities and LFs among different surveys. We construct a reliable LF of Ly$\alpha$ blobs from a deep, blind narrowband survey. Larger volume blob surveys, combined with large volume and/or higher resolution N-body simulations, will improve the constraints on blob halo mass and detectability fraction, thus discriminating among the possible energy sources of the extended Ly$\alpha$ emission.

We thank an anonymous referee for helpful comments. Y.Y. thanks the KPNO/CTIO staffs for observing support and the NOAO graduate student travel fund. We thank Yuichi Matsuda for providing the narrowband images of his Ly$\alpha$ blobs, Andrea Macciò for the halo growth function data, and Masami Ouchi for his LF of LAEs. We gratefully acknowledge the assistance of Toshihiko Kimura at Asahi Spectra for his help in manufacturing our narrowband filter. A.I.Z. acknowledges support from the NSF Astronomy and Astrophysics Research Program through grant AST-0908280 and from the NASA Astrophysics Data Analysis Program through grant NNX10AD47G.

Facilities: Mayall (MOSAIC I), Blanco (MOSAIC II)

APPENDIX

COMPARISON WITH MATSUDA ET AL. Ly$\alpha$ BLOB SAMPLE

It is difficult to compare the properties of Ly$\alpha$ blob samples among different surveys because of non-uniform selection criteria and different imaging depth. Here, we compare our $z = 2.3$ sample to that of Matsuda et al. (2004) at $z = 3.1$. We repeat the procedures for the recovery test described in Section 3 (see also Yang et al. 2009), pasting thumbnail images provided by Matsuda et al. (2004) into our CDFS image. Because the narrowband filter bandwidths of the two surveys are similar, we do not make any correction for the difference in filter transmission, but we scale the apparent size and the surface brightness accounting for the different redshifts.

Further studies are required to establish that the CDFS overdensity and similar overdensities observed by other authors (e.g., Prescott et al. 2008) are in fact “proto-clusters.” For now, we note that structure in the CDFS at $z = 2.3$ is coherent over at least eight comoving Mpc and is likely to contain tens of $\sim 10^{13} M_\odot$ blob halos, or at least $\sim 10^{14} M_\odot$ worth of mass. This over-density grows to typically and virialize by $z = 0$, then its halo mass today would be like that of the rich Coma cluster, i.e., $\sim 10^{15} M_\odot$.

5. CONCLUSIONS AND SUMMARY

To understand what Ly$\alpha$ blobs will become in the present-day universe requires that we first constrain their number density, clustering, and large-scale environment. In order to obtain unbiased measures of these quantities, we target four clusters, and large-scale environment. In order to obtain unbiased measures of these quantities, we target four clusters, and large-scale environment. In order to obtain unbiased measures of these quantities, we target four clusters, and large-scale environment. In order to obtain unbiased measures of these quantities, we target four clusters, and large-scale environment.
Figure 13. Comparison in $L_{\text{Ly}\alpha}$ and $A_{\text{iso}}$ between our 16 CDFS blobs (horizontal bars) and the 35 Matsuda et al. (2004) blobs in the SSA22 field (filled circles). The vertical and horizontal dashed lines indicate the size selection criteria for the extended sources adopted by Matsuda et al. (2004) and by our work, respectively. The dotted lines represent the average ratios ($\sim 0.61$ and 1.0) between originally reported and the recovered values in $A_{\text{iso}}$ and $L_{\text{Ly}\alpha}$. The recovery test of the Matsuda et al. (2004) blobs when pasted into our CDFS field shows that we can detect all blobs like theirs in our survey. (A color version of this figure is available in the online journal.)

Figure 13 shows the Ly$\alpha$ luminosity and isophotal area recovered from this test as a function of the input $L_{\text{Ly}\alpha}$ and $A_{\text{iso}}$ for the 35 Matsuda et al. (2004) blobs. Due to the slightly shallower depth of our survey, the Matsuda et al. blobs would look smaller by a factor of 61% than the originally reported sizes. Note that one expects an 86% decrease in area purely from the differences in angular diameter distance between two survey redshifts. The vertical and horizontal dashed lines indicate the size selection criteria for the extended sources adopted by Matsuda et al. (2004) and by our work, respectively. Most Matsuda et al. (2004) blobs are recovered as larger than 10 $\alpha''$, confirming the capability of our survey for detecting them. The line luminosities are also recovered well, so there is no bias in the luminosity measurements.

REFERENCES

Arnouts, S., et al. 2001, A&A, 379, 744
Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587
Bernardeau, F., & Kofman, L. 1995, ApJ, 443, 479
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blanton, W. N., et al. 2001, AJ, 122, 2810
Capak, P., et al. 2004, AJ, 127, 180
Capak, P., et al. 2007, ApJS, 172, 99
Colbert, J. W., Teplitz, H., Francis, P., Palunas, P., Williger, G. M., & Woodgate, B. 2006, ApJ, 637, L89
Coles, P., & Jones, B. 1991, MNRAS, 248, 1
Dey, A., et al. 2005, ApJ, 629, 654
Dijkstra, M., & Loeb, A. 2009, MNRAS, 400, 1109
Ebeling, H., White, D. A., & Rangarajan, F. V. N. 2006, MNRAS, 368, 65
Erben, T., et al. 2005, Astron. Nachr., 326, 432
Fardal, M. A., Katz, N., Gardner, J. P., Hernquist, L., Weinberg, D. H., & Davé, R. 2001, ApJ, 562, 605
Francis, P. J., et al. 2001, ApJ, 554, 1001
Gawiser, E., et al. 2006a, ApJS, 162, 1
Gawiser, E., et al. 2006b, ApJ, 642, L13
Geach, J. E., et al. 2005, MNRAS, 363, 1398
Geach, J. E., et al. 2009, ApJ, 700, 1
Giacconi, R., et al. 2002, ApJS, 139, 369
Giavalisco, M., et al. 2006, ApJ, 640, L123
Goerdt, T., Dekel, A., Sternberg, A., Ceverino, D., Teyssier, R., & Primack, J. R. 2010, MNRAS, 933
Gronwall, C., et al. 2007, ApJ, 667, 79
Haiman, Z., & Jones, B. 2006, ApJ, 637, L89
Hayashino, T., et al. 2004, AJ, 128, 2073
Hayashino, T., et al. 2004, AJ, 128, 2073
Hennawi, J. F., Prochaska, J. X., Kollmeier, J., & Zheng, Z. 2009, ApJ, 693, L49
Hildebrandt, H., et al. 2006, A&A, 452, 1121
Hogg, D. W., Cohen, J. G., Sandford, R., & Pakstis, A. 1998, ApJ, 504, 622
Hogg, D. W., Cohen, J. G., & Pakstis, A. 1998, ApJ, 504, 622
Jannuzi, B. T., & Dey, A. 1999, in ASP Conf. Ser. 171, Photometric Redshifts and the Detection of High Redshift Galaxies, ed. R. Weymann et al. (San Francisco, CA: ASP), 111
Keel, W. C., Cohen, S. H., Windhorst, R. A., & Waddington, I. 1999, AJ, 118, 2547
Keel, W. C., White, R. E., Chapman, S., & Windhorst, R. A. 2009, ApJ, 138, 986
Koekemoer, A. M., et al. 2007, ApJS, 172, 196
Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, MNRAS, 391, 1940
Matsuda, Y., Yamada, T., Hayashino, T., Yamauchi, R., & Nakamura, Y. 2006, ApJ, 640, L123
Matsuda, Y., et al. 2004, ApJ, 618, 569
Matsuda, Y., et al. 2005, ApJ, 634, L125
Matsuda, Y., et al. 2009, MNRAS, 400, L66
Munet, D. G., et al. 2003, AJ, 125, 984
Nilsson, K. K., Fynbo, J. P. U., Møller, P., Sonner-Larsen, J., & Ledoux, C. 2006, A&A, 452, L23
Nilsson, K. K., Tapken, C., Møller, P., Freudling, W., Fynbo, J. P. U., Meisenheimer, K., & Laursen, P., & Ostri, G. 2009, A&A, 498, 13
Oke, J. B. 1974, ApJS, 27, 21
Ouchi, M., et al. 2008, ApJS, 176, 301
Ouchi, M., et al. 2009, ApJ, 696, 1164
Palunas, P., Teplitz, H. I., Francis, P. J., Williger, G. M., & Woodgate, B. E. 2004, ApJ, 602, 945
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton, NJ: Princeton Univ. Press)
Pescott, M. K. M., Dey, A., & Jannuzi, B. T. 2009, ApJ, 702, 554
Pescott, M. K. M., Kashikawa, N., Dey, A., & Matsuda, Y. 2008, ApJ, 678, L77
Rix, H.-W., et al. 2004, ApJS, 152, 163
Saito, T., Shimasaku, K., Okamura, S., Ouchi, M., Akiyama, M., & Yoshida, M. 2006, ApJ, 648, 54
Salimbeni, S., et al. 2009, A&A, 501, 865
Scoville, N., et al. 2007, ApJS, 172, 1
Shioya, Y., et al. 2009, ApJ, 696, 546
Smith, D. J. B., & Jarvis, M. J. 2007, MNRAS, 378, L49
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2000, ApJ, 532, 170
Taniguchi, Y., & Shioya, Y. 2000, ApJ, 532, L13
Valdes, F. G. 1998, in ASP Conf. Ser. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco, CA: ASP), 53
Wolf, C., et al. 2004, A&A, 412, 913
Yan, Y., Zabludoff, A. I., Davé, R., Eisenstein, D. J., Pinto, P. A., Katz, N., Weinberg, D. H., & Barton, E. J. 2006, ApJ, 640, 539
Yan, Y., Zabludoff, A. I., Tremonti, C., Eisenstein, D., & Davé, R. 2009, ApJ, 693, 1579