



THE CANDIDATE CLUSTER AND PROTOCLUSTER CATALOG (CCPC). II. SPECTROSCOPICALLY IDENTIFIED STRUCTURES SPANNING $2 < z < 6.6$

J. R. FRANCK AND S. S. MCGAUGH

Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106, USA

Received 2016 April 1; revised 2016 September 30; accepted 2016 October 1; published 2016 December 1

ABSTRACT

The Candidate Cluster and Protocluster Catalog (CCPC) is a list of objects at redshifts $z > 2$ composed of galaxies with spectroscopically confirmed redshifts that are coincident on the sky and in redshift. These protoclusters are identified by searching for groups in volumes corresponding to the expected size of the most massive protoclusters at these redshifts. In CCPC1 we identified 43 candidate protoclusters among 14,000 galaxies between $2.74 < z < 3.71$. Here we expand our search to more than 40,000 galaxies with spectroscopic redshifts $z > 2.00$, resulting in an additional 173 candidate structures. The most significant of these are 36 protoclusters with overdensities $\delta_{\text{gal}} > 7$. We also identify three large proto-supercluster candidates containing multiple protoclusters at $z = 2.3$, 3.5 and $z = 6.56$. Eight candidates with $N \geq 10$ galaxies are found at redshifts $z > 4.0$. The last system in the catalog is the most distant spectroscopic protocluster candidate known to date at $z = 6.56$.

Key words: galaxies: clusters: general – galaxies: high-redshift

Supporting material: figure set, machine-readable tables

1. INTRODUCTION

Protoclusters are the diffuse, extended building blocks that will collapse into a galaxy cluster at the current epoch. Structures at high redshift have power as tests of cosmology, as the maximum mass that can collapse and virialize at a given epoch depends on the mass density (Ω_m), the power spectrum (σ_8), and the dark energy equation of state (w) (Bahcall & Fan 1998; Vikhlinin et al. 2009; Mortonson et al. 2011). These primeval systems are also unique laboratories in which we can observe galaxies in the universe assemble into larger collections and evolve in dense environments. Galaxy formation and hierarchical accretion scenarios can be examined by compiling large numbers of high-redshift galaxies at various times (Cooke et al. 2014; Wylezalek et al. 2014). Protocluster galaxies have been found to have enhanced mass assembly (Casey et al. 2015) and evolution (Hatch et al. 2011). Star formation rates of $10^4 M_\odot \text{yr}^{-1}$ for a protocluster (Clements et al. 2014) are in significant excess of hydrodynamic simulation expectations (Contini et al. 2015; Granato et al. 2015). As a result, the search for high-redshift clusters and protoclusters within the astrophysical community has been rapidly expanded in the last few years.

Less than two decades ago, protoclusters were relatively unknown until rich, Lyman-break galaxy (LBG) overdensities were discovered by Steidel et al. (1998) at $z \sim 3$ and Ly α emitters (LAEs) at $z \sim 4$ (Venemans et al. 2002). Until recently, few systematic surveys for $z \gtrsim 2$ systems existed. With the advent of the Clusters Around Radio Loud AGN *Spitzer* survey (Wylezalek et al. 2013), the Search for Protoclusters with *Herschel* (Rigby et al. 2014), data mining the GOODS-N/GOODS-S fields (Kang & Im 2009; Salimbeni et al. 2009; Kang & Im 2015), and finding structures within the COSMOS/UltraVISTA field (Diener et al. 2013; Chiang et al. 2014), the number of photometrically/spectroscopically identified candidate protoclusters is now in the hundreds. Catalogs of pure spectroscopically identified protoclusters in the high-redshift universe, as in Venemans et al. (2007),

typically have fewer than 10 such candidates, however. Most of these previously identified structures are found within $z \sim 3 \pm 1$ (Venemans et al. 2007; Wylezalek et al. 2014; Franck & McGaugh 2016), with a few at $z \sim 6$ (Utsumi et al. 2010; Toshikawa et al. 2014). These objects were identified with varying instruments, selection techniques, and fields of view (FOVs) that rarely encompass the entire system.

It was this understanding that prompted a systematic search of archival data in a simple manner to identify high-redshift ($z > 2$) clusters and protoclusters. This work builds on the original harvest of 43 members in the Candidate Cluster and Protocluster Catalog (CCPC) between $2.74 < z < 3.71$ (hereafter known as CCPC1 Franck & McGaugh 2016). These structures were identified as galaxy overdensities in fixed comoving volumes. In this second paper (hereafter CCPC2), we have extended the redshift range of our search to $2.00 < z < 7.00$ using a similar search method. We eschew the common search technique of using high-redshift radio galaxies as biased tracers of structure (Venemans et al. 2002; Wylezalek et al. 2013; Rigby et al. 2014), but we have recovered overdensities identified in this way. Presented in this work are 173 additional protocluster candidates, 23 of which have been found in the literature.

We have organized CCPC2 as follows: Section 2 discusses the archival data from which these structures are gleaned, the algorithm that identifies structures, calculation of the overdensity of each system, and the explanation of the two mass estimators. In Section 3, we explore the nature of these overdensities, their significance as structures, compare their properties to simulations, and discuss three supercluster candidates.

We assume a standard cosmology, adopting $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, a matter density of $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. The universe is 3.2 Gyr old at the redshift of $z = 2.0$, and has a comoving angular scale of $1.51 \text{ Mpc arcmin}^{-1}$. At $z = 7$, the corresponding age is 0.75 Gyr, while the angular scale is $2.51 \text{ comoving Mpc arcmin}^{-1}$ (Wright 2006).

2. THE CANDIDATE CLUSTER AND PROTOCLUSTER CATALOG (CCPC) II

2.1. Data

In CCPC1, a NASA Extragalactic Database (NED¹) search was used to compile a list of $\sim 14,000$ galaxies with spectroscopic redshifts between $2.74 < z < 3.71$. Many of these galaxies were found in the Hubble Deep Fields, and spectroscopic follow-up of LBGs Steidel et al. (2003). In this work, we have expanded the redshift range to $2.00 < z < 10.23$, which constitutes a galaxy list of 47,000 objects. The upper limit of $z = 10.23$ is an extreme example; only nine galaxies have $z > 7.5$ in our list. The NED database holds published redshift uncertainties for 813 of the ~ 1400 candidate protocluster galaxies, with a mean value of $\sigma = 0.001$. At a redshift of $z = 3$, this represents a comoving distance uncertainty of 1.0 Mpc (Wright 2006).

Nearly 200 sources of spectroscopic measurements were used to identify candidate clusters (listed in Table 4), although many are concentrated in a few catalogs. The single largest source of spectroscopic redshifts that were identified as protocluster galaxies is from Steidel et al. (2003). These galaxies were identified as LBGs initially (with additional color criteria), and then followed up spectroscopically to a limiting magnitude of $R_{AB} = 25.5$ using the Keck telescopes over a FOV of 0.38 square degrees. Many other spectroscopic surveys utilized very deep, multi-wavelength fields (e.g., CANDELS GOODS, Hubble Deep/Ultra Deep) to identify candidate high-redshift galaxies with various color cuts, dropouts, or *Chandra* emission (Noll et al. 2004; Le Fèvre et al. 2005; Reddy et al. 2006; Balestra et al. 2010). These galaxies would then be targeted by VIMOS/FORS/MOIRCS (or a similar instrument) to confirm their redshift.

Another common source of redshifts is from surveys that capitalize on strong line emission from faint star-forming galaxies, either with Ly α or H α . By using a narrowband (NB) filter that selects out one of these lines at a specific redshift, one can effectively identify a large number of sources. These candidates can then be targeted using a large telescope to confirm their redshift. This technique has been used to identify a number of high-redshift protoclusters (Fynbo et al. 2003; Venemans et al. 2007; Cooke et al. 2014). An interesting observation that was noticed by Koyama et al. (2013) is that Ly α and H α are rarely simultaneously observed in star-forming galaxies within a protocluster at $z = 2.16$, suggesting that they may be entirely different populations. They could also have varying amounts and distributions of dust. Simulations suggest that roughly 90% of protocluster galaxies are actively forming stars at these redshifts (Contini et al. 2015).

2.2. Candidate Protocluster Criteria

While the most massive, $z = 0$ clusters have radii on the order of a few Mpc, the components that form these systems (i.e., protoclusters) are much more extended at higher redshift. Previously identified structures at high redshift have observed sizes larger than 50 comoving Mpc (cMpc) (Shimasaku et al. 2003; Matsuda et al. 2005; Lee et al. 2014). These seemingly large sizes have a theoretical basis within Λ CDM simulations as well. Chiang et al. (2013) and Muldrew et al. (2015) mined the Millennium simulation (Springel et al. 2005) for clusters (collapsed halos with $\gtrsim 10^{14} M_{\odot}$) at $z = 0$. Using the semi-analytic model of Guo et al. (2011) to trace the galaxy distributions of the primordial clusters

backwards, they analyzed the evolution of these systems. Muldrew et al. (2015) found that at $z = 2$, 90% of the stellar mass of a protocluster can be found spread across $35 h^{-1}$ cMpc, which grows to more than $40 h^{-1}$ cMpc at $z = 5$. This example is typical for the largest-mass systems (clusters with $M_{z=0} > 10^{15} M_{\odot}$), while a cluster with $M_{z=0} = 10^{14} M_{\odot}$ might have a modest size of $\sim 20 h^{-1}$ cMpc. On the sky, the most massive protoclusters can span more than 0.5° , while the smallest mass clusters are roughly half of that (Muldrew et al. 2015).

This can be problematic when searching for such systems, in that many deep, spectroscopic surveys (with the exception of the CANDELS GOODS-S field) would not encompass the full breadth of the most massive protoclusters. In Franck & McGaugh (2016), we plotted the positions of field and CCPC1 protocluster galaxies within search radii of 20 cMpc to illustrate the distribution of members with respect to the survey widths in the Appendix. Surveys rarely extended beyond the probed search radius. Similarly, Muldrew et al. (2015) make an example of two candidate systems found within the literature (Koyama et al. 2013; Cooke et al. 2014) that have particularly small FOVs which could only target the innermost halo of a protocluster. This imposes a selection effect, in that most ($\sim 90\%$) simulated protoclusters do not have a single, “main” halo at these redshifts (Muldrew et al. 2015).

We took an initial list of $\sim 14,000$ galaxies and identified galaxy groups of three or more within a radius of 2 arcminutes and a $\Delta z < 0.03$ in CCPC1. Individual galaxies are required to be more than $3''$ from one another to be considered unique. From these groups, we expanded our search to a radius of 20 cMpc on the sky and a redshift depth of ± 20 cMpc. Any cylindrical volume that contained four or more galaxies and had a galaxy overdensity of $\delta_{\text{gal}} > 0.25$ (calculation found in the following subsection) was considered a candidate protocluster and included in CCPC1.

These requirements are based on the need for a simple, adaptable, and effective means of identifying galaxy structures from a variety of surveys. The large search radius enables wide surveys to be adequately probed corresponding to the size of the largest protoclusters. The modest richness requirement of four or more galaxies is sensitive to surveys of small volumes with few expected galaxies. These two extremes of survey depths are moderated by the density requirement of $\delta_{\text{gal}} > 0.25$. The $N \geq 4$ galaxy requirement serves primarily as a signpost from which to calculate the overdensity. A more detailed discussion of the algorithm will follow.

In this paper, we have removed the intermediate step of finding groups of three or more galaxies within a 2 arcminute radius. The overdensity of these candidate protoclusters, which may not be accompanied by a dense, central knot of galaxies, is of greater importance than the possible chance alignment of galaxies on the sky. Therefore, any group of four or more galaxies that also exhibits a galaxy overdensity of $\delta_{\text{gal}} > 0.25$ is considered to be part of the CCPC. The following subsection will describe the overdensity calculation in detail. In Section 3.1 we justify the removal of the intermediate step statistically.

Identifying structure in large, high-redshift data sets is not a novel exercise. Diener et al. (2013) found infalling groups of spectroscopic galaxies separated by small physical separations in space in the zCOSMOS field. The GOODS-N and GOODS-S fields have been the subject of many searches for overdensities found with photometric redshifts and supplementary spectroscopic catalogs at $z > 2$ (Kang & Im 2009, 2015; Salimbeni et al. 2009).

¹ <http://ned.ipac.caltech.edu/>

Indeed, we may have recovered a number of these overdensities, which have $0 \leq N \leq 4$ spectroscopic member galaxies in Kang & Im (2015). In CCPC1 (Franck & McGaugh 2016), our candidate protocluster CCPC-z37-001 was identified near a photometric redshift overdensity by Kang & Im (2009), while CCPC-z27-004, CCPC-29-004, and CCPC-z34-001 most likely coincide with overdensities found in Kang & Im (2015). CCPC-z22-001, CCPC-z23-001, and CCPC-z24-001 (all featured in CCPC2) have similar positions and redshifts to the overdensities reported by Salimbeni et al. (2009), while CCPC-z25-004, CCPC-z25-005, CCPC-z40-001, and CCPC-z42-001 may coincide with candidate structures in Kang & Im (2015). CCPC-z28-002, CCPC-z34-002, and CCPC-z37-001 in Franck & McGaugh (2016) are all overdensities originally found in the VIMOS GOODS-S spectroscopic survey by Balestra et al. (2010).

The CCPC is unique in that it applied a single and simple algorithm to all archival spectroscopic data available. The number of galaxy overdensities identified in this paper (hereafter referred to as CCPC2) is 173, the 36 strongest of which are listed in Table 1. The division of these “best” candidates is explained in Section 2.3 (overdensities). The total list of candidates can be found in Table 3. The R.A./decl. coordinates in the table are centered on the galaxy from which the number of members is maximized for a given overdensity. The redshift listed is the mean value of the system members, and may be minimally offset from the search galaxy’s redshift in some cases ($\Delta z \sim 0.0005$ on average). We have included the number of galaxies within $R = 10, 20$ cMpc from the central galaxy, which are equivalent in many instances. This is primarily dependent on the limited FOV of the surveys. In combination with the 43 protocluster candidates in CCPC1, the CCPC contains 216 systems between the redshifts of $2 < z < 7$. To the best of our knowledge, this represents the largest collection of spectroscopic, $z \geq 2$ galaxy overdensities in the literature at the time of writing. Each CCPC candidate has an individual list of redshift measurement references in Table 4.

As in CCPC1, many of the protoclusters have more than the minimum number of members, with a median value of six galaxies per candidate. There are 40 CCPC2 systems with the minimum four galaxies, and nine candidates with 23 or more members (Table 3). There does not appear to be a strong trend in numbers of galaxy members as a function of redshift for the bulk of the CCPC. However, all candidates with more than 20 galaxies are found at $z < 4.5$. The median number of galaxies in protoclusters in the redshift bins of $2 < z < 3$, $3 < z < 4$, $4 < z < 5$, $5 < z < 7$ are 5, 6, 8, and 7, respectively within CCPC2. Although these are only slight variations, one would naively expect that the median numbers would decrease as a function of distance. Perhaps at high redshift we are only identifying the richest overdensities, and thus their median members are larger.

Each candidate structure in CCPC2 has a spectroscopic rating (“ Q ”) associated with it in Tables 1 and 3. In many cases, redshift values cataloged by NED have an accompanying qualifier flag that distinguishes the quality of a given redshift measurement. If a protocluster has four or more member galaxies with no quality flags raised, thus satisfying our criteria as a candidate structure, it is assigned the greatest rating of “1.” If there are four or more galaxies that have either no flags or were identified based on a single line, the system is rated slightly lower with a “2.” A rating of “3” is assigned to collections of galaxies with redshift flags of a somewhat

uncertain measurement or tentative result. Galaxies with redshifts flagged as photometric redshift, modeled from SEDs, or highly uncertain/questionable values, are not used in this work. In CCPC2, there are 135 (out of 173) systems identified with the highest-quality spectroscopic data (RATING = 1), and 32 that have a rank “2” rating.

2.3. Candidate Overdensities

To measure the strength of these candidate systems quantitatively, we estimate their galaxy overdensity using the simple formula $\delta_{\text{gal}} = (n_{\text{proto}}/n_{\text{field}}) - 1$. The number density of protocluster galaxies (n_{proto}) is based on the density of galaxies along the line of sight (LOS) within the dispersion (σ_z) from the center of the redshift distribution. Limiting the galaxies in the overdensity to those only within the redshift dispersion likely excludes some objects that are on the outskirts of the candidate structure, and so the δ_{gal} values are conservative estimates. These overdensities and their uncertainties can be found in Tables 1 and 3. Figure 1 illustrates the calculation of the galaxy overdensity.

To calculate the number density of the field (n_{field}), we use the same aperture on the sky as our search criteria ($R = 20$ cMpc) and identify all galaxies along the LOS (using the same galaxy surveys) within a maximum length of $\Delta z = 0.15$ from the protocluster center. The overdense region is excised from the field counts. The choice of the field length is typically an order of magnitude longer than the overdense region to ensure a fair sample. A visual representation of the calculation is shown in Figure 1. It is important to note that in every step of calculating the overdensity (δ_{gal}), we are adopting the most conservative values. The purpose of the CCPC was to investigate protoclusters using methods which did not “cherry-pick” the largest values of overdensities.

The uncertainties in field counts are estimated using the cosmic variance calculator from Trenti & Stiavelli (2008), while the uncertainty within the overdense region is $\sqrt{N_{\text{proto}}}$. For each structure, we input the volume probed by field systems and an assumption of the completeness of the spectroscopic survey. The output is not particularly sensitive to the completeness assumption ($\sigma \pm 0.1$ galaxies between 10% and 90%), and so we have adopted a 50% completeness. The field and protocluster uncertainties are then added in quadrature. On average, the inclusion of cosmic variance and completeness adds 0.3 galaxies to the uncertainty of the CCPC when compared to the Poissonian treatment of uncertainties.

When measuring n_{field} or n_{proto} for the calculation of δ_{gal} , the length (Δz) that the density is computed over is limited by the extent of the data, and not the volume queried (e.g., $\Delta z \leq 0.15$ for the field). For example, NB filters do not extend the full possible width of the field distribution. In $<15\%$ of cases in the CCPC, low numbers of galaxies clustered in redshift space increased n_{proto} or n_{field} to large levels (e.g., $\delta_{\text{gal}} > 50$). We set Δz equal to the protocluster redshift dispersion (σ_z) for these low-richness cases. In some overlapping instances, the field galaxy counts were very low ($N_{\text{field}} < 3$) compared to other candidate structures. For these systems in CCPC1 we injected seven more galaxies into the counts for the field number density n_{field} to reflect the median field counts in the low-richness sample. This effectively decreases the overdensities to more reasonable values. In this work we have neglected this rather un-physical method, but instead put brackets around the overdensity values to reflect their low-richness status in

Table 1
Candidate Cluster and Protocluster Catalog (CCPC) II: Best Candidates

| Candidate Name | R.A. (degree) | Decl. (degree) | Redshift (z_{avg}) | σ_z | N | $N_{R \leq 10}$ cMpc | Overdensity (δ_{gal}) | Cluster Probability (%) | Q | Recovered Reference |
|----------------|---------------|----------------|-------------------------------|------------|-----|----------------------|---------------------------------------|-------------------------|-----|---------------------|
| CCPC-z20-002 | 222.20 | 8.92 | 2.002 | 0.008 | 11 | 7 | 9.38 ± 5.34 | 100.0 | 1 | 8 |
| CCPC-z20-003 | 29.62 | -25.05 | 2.018 | 0.004 | 10 | 10 | 19.43 ± 13.06 | 100.0 | 1 | 1, 2 |
| CCPC-z20-009 | 150.04 | 2.21 | 2.098 | 0.005 | 10 | 4 | 13.15 ± 6.54 | 100.0 | 1 | 4, 5 |
| CCPC-z21-004 | 175.15 | -26.47 | 2.155 | 0.007 | 24 | 24 | 6.34 ± 3.36 | 100.0 | 1 | 9, 1 |
| CCPC-z21-005 | 214.31 | 52.40 | 2.160 | 0.007 | 5 | 3 | 9.07 ± 6.93 | 100.0 | 1 | ... |
| CCPC-z21-006 | 334.35 | 0.32 | 2.172 | 0.005 | 4 | 3 | 18.85 ± 13.55 | 100.0 | 1 | ... |
| CCPC-z21-007 | 356.58 | 12.80 | 2.174 | 0.002 | 7 | 7 | 17.27 ± 10.57 | 100.0 | 1 | ... |
| CCPC-z21-008 | 149.98 | 2.11 | 2.179 | 0.002 | 5 | 1 | 9.41 ± 5.38 | 100.0 | 2 | 4 |
| CCPC-z22-007 | 255.20 | 64.26 | 2.296 | 0.008 | 32 | 30 | 7.77 ± 2.90 | 100.0 | 1 | 6 |
| CCPC-z23-002 | 334.46 | 0.14 | 2.309 | 0.009 | 4 | 3 | 11.45 ± 9.05 | 100.0 | 1 | ... |
| CCPC-z23-003 | 214.39 | 52.49 | 2.333 | 0.008 | 4 | 4 | 11.73 ± 10.66 | 100.0 | 1 | ... |
| CCPC-z23-007 | 258.53 | 50.27 | 2.390 | 0.005 | 7 | 6 | 16.63 ± 14.52 | 100.0 | 1 | 1, 20 |
| CCPC-z24-003 | 164.20 | -3.64 | 2.426 | 0.005 | 7 | 7 | 15.06 ± 8.92 | 100.0 | 1 | ... |
| CCPC-z24-005 | 150.00 | 2.26 | 2.442 | 0.009 | 14 | 8 | 9.27 ± 4.93 | 100.0 | 1 | 4, 10 |
| CCPC-z25-002 | 255.18 | 64.17 | 2.537 | 0.002 | 4 | 4 | 19.86 ± 13.41 | 100.0 | 1 | ... |
| CCPC-z25-003 | 143.36 | 28.77 | 2.548 | 0.003 | 5 | 5 | 10.89 ± 7.70 | 100.0 | 1 | ... |
| CCPC-z25-007 | 216.14 | 22.84 | 2.581 | 0.007 | 5 | 2 | 10.90 ± 6.72 | 100.0 | 1 | ... |
| CCPC-z27-012 | 16.48 | -25.81 | 2.758 | 0.008 | 4 | 3 | 12.83 ± 10.91 | 100.0 | 1 | ... |
| CCPC-z28-011 | 36.36 | -4.32 | 2.820 | 0.006 | 4 | 2 | 9.10 ± 7.53 | 100.0 | 1 | ... |
| CCPC-z28-016 | 36.27 | -4.28 | 2.866 | 0.006 | 5 | 1 | 15.46 ± 11.42 | 100.0 | 1 | ... |
| CCPC-z29-009 | 136.34 | 34.14 | 2.905 | 0.010 | 5 | 5 | 13.33 ± 9.13 | 100.0 | 1 | ... |
| CCPC-z29-011 | 339.95 | 11.87 | 2.925 | 0.008 | 13 | 3 | 9.62 ± 5.14 | 100.0 | 1 | ... |
| CCPC-z29-013 | 13.31 | 12.63 | 2.934 | 0.002 | 5 | 4 | 14.67 ± 10.21 | 100.0 | 1 | ... |
| CCPC-z31-015 | 339.87 | 11.88 | 3.148 | 0.008 | 9 | 2 | 8.15 ± 4.56 | 100.0 | 1 | ... |
| CCPC-z32-007 | 46.15 | -0.19 | 3.233 | 0.007 | 5 | 4 | 11.47 ± 9.32 | 100.0 | 1 | ... |
| CCPC-z33-006 | 150.07 | 2.28 | 3.303 | 0.008 | 4 | 4 | 17.49 ± 13.89 | 100.0 | 2 | ... |
| CCPC-z33-007 | 334.35 | 0.07 | 3.310 | 0.012 | 7 | 5 | 11.52 ± 9.15 | 100.0 | 1 | ... |
| CCPC-z33-010 | 216.12 | 22.83 | 3.379 | 0.009 | 7 | 3 | 10.44 ± 7.10 | 100.0 | 1 | ... |
| CCPC-z36-007 | 34.54 | -5.30 | 3.688 | 0.010 | 5 | 2 | 15.22 ± 16.25 | 90.0 | 2 | ... |
| CCPC-z44-003 | 189.15 | 62.23 | 4.424 | 0.010 | 5 | 3 | 13.79 ± 12.03 | 90.0 | 1 | ... |
| CCPC-z26-001 | 339.84 | 11.81 | 2.617 | 0.003 | 4 | 4 | 7.33 ± 6.16 | 85.6 | 1 | ... |
| CCPC-z26-006 | 255.14 | 64.22 | 2.688 | 0.005 | 5 | 5 | 7.24 ± 5.18 | 85.6 | 1 | ... |
| CCPC-z27-008 | 36.39 | -4.51 | 2.729 | 0.006 | 6 | 2 | 7.27 ± 4.83 | 85.6 | 1 | ... |
| CCPC-z31-008 | 339.89 | 11.88 | 3.104 | 0.007 | 8 | 5 | 7.70 ± 4.58 | 85.6 | 1 | ... |
| CCPC-z31-017 | 36.76 | -4.56 | 3.187 | 0.006 | 11 | 3 | 7.26 ± 3.67 | 85.6 | 1 | ... |
| CCPC-z34-006 | 36.54 | -4.63 | 3.472 | 0.013 | 6 | 1 | 7.82 ± 5.40 | 85.6 | 2 | ... |

Note. The names and positions (first through third columns) of the most overdense candidate protoclusters, with redshifts corresponding to the average value for the system, with their dispersion σ_z . The number of galaxies within $R = 10$ and 20 cMpc from the search center are included as $N_{R \leq 10}$ (column 7) and N (column 6), respectively. Column 8 contains a measure of the galaxy overdensity (δ_{gal}), which is explained in Section 2.3. A probability is assigned to each system that it will collapse into a cluster by $z = 0$. These probability distribution functions at the relevant redshifts can be found in Figure 8 of Chiang et al. (2013) from analysis of protocluster δ_{gal} values in the Millennium simulation. Entries in the table are arranged by their collapse probabilities in descending order. The “ Q ” column provides a rating of the source quality of redshifts as identified within NED. The best-quality redshift systems are listed as “1.” If we recover an overdensity that was previously identified, the discovery references are listed in the last column.

References. (1) Galametz et al. (2012), (2) Galametz et al. (2013), (4) Diener et al. (2013), (5) Yuan et al. (2014), (6) Steidel et al. (2005), (8) Gobat et al. (2013), (9) Pentericci et al. (1997), (10) Chiang et al. (2014), (20) Keel et al. (1999).

Table 3. The resulting overdensities are highly uncertain as a result, and should not be relied on without further observations. Although these overdensity vales are questionable, the average number density of these systems ($n \sim 7 \times 10^{-2} \text{ cMpc}^{-3}$) exceeds that of the mean value for CCPC2. This suggests that these systems are likely overdensities, and need not be removed simply because they lack field counts.

The median value of $\delta_{\text{gal}} = 2.9$ for CCPC2, only slightly larger than for CCPC1 ($\delta_{\text{gal}} \sim 2$). Only 15% of CCPC2 systems have $\delta_{\text{gal}} < 1$. Generally, $\delta_{\text{gal}} \sim 2$ is typical of many protoclusters in the literature (for a summary, see Table 5 in Chiang et al. 2013). These values have considerable breadth, from $0.7_{-0.6}^{+0.8}$ (Venemans et al. 2007) to 16 ± 7 (Toshikawa et al. 2014), which are consistent with the range in the CCPC. The mean number density of the δ_{gal} volume is $n =$

$5.9 \times 10^{-2} \text{ cMpc}^{-3}$ in CCPC2. For a general comparison, the density of LAEs at $z \sim 3$ is $n = 1.5 \times 10^{-3} \text{ cMpc}^{-3}$ (Gronwall et al. 2007), but this value is dependent on the galaxy type. In simulated protoclusters within Millennium, analysis shows that overdensities of $\delta_{\text{gal}} = 1 \pm 1$ from redshifts $z = 2-5$ are consistent with low-mass, $M_{z=0} \approx 10^{14} M_{\odot}$ clusters (Chiang et al. 2013). This assumes that galaxies in the overdensity have $\text{SFR} > 1 M_{\odot} \text{ yr}^{-1}$ and are within boxes of $[25 \text{ cMpc}]^3$. The most massive systems $M_{z=0} > 10^{15} M_{\odot}$ have overdensities roughly a factor of 3 larger, as one would expect (Chiang et al. 2013).

These simulations also show that δ_{gal} expectations for protoclusters are inversely proportional to the volume probed. Steidel et al. (2003) identified LBGs in windows of $R \sim 5 \text{ cMpc}$ on the sky at these redshifts, which is only 25% the size of the maximum CCPC search radius. The LOS

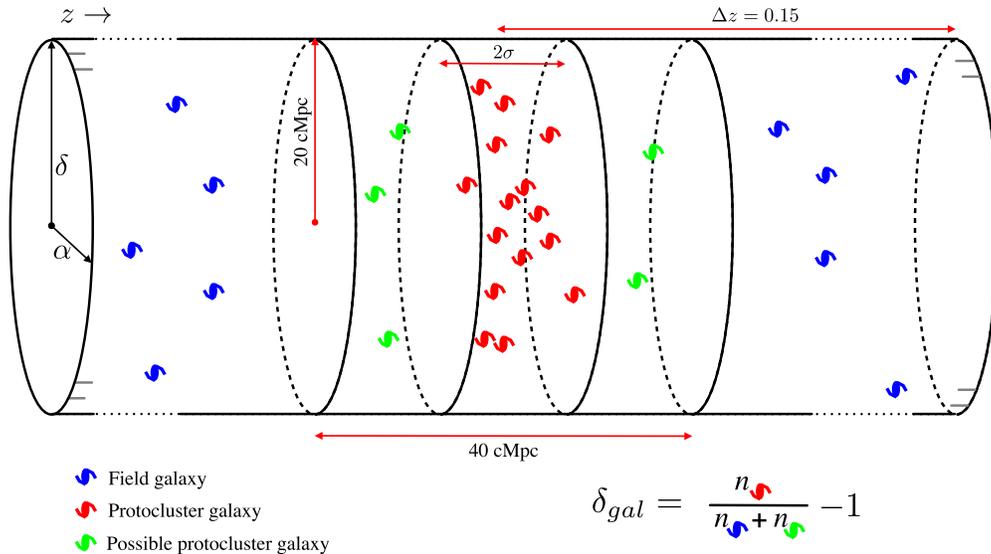


Figure 1. A visual representation of the overdensity (δ_{gal}) measurement. Galaxies within ± 20 cMpc but outside of $z \pm \sigma$ are considered possible protocluster members, but are treated as field galaxies for the computation of the overdensity, so that it is a conservative estimate.

distance probed is not similarly hampered. Figure 3 in Chiang et al. (2013) suggests that the stacked overdensity profiles of Millennium protoclusters can be factors of 6 or more larger when evaluated in such small boxes ($R_e \sim 5$, $z = 3$). Indeed, the CCPC candidates in the regions of Steidel et al. (2003) do have a larger typical overdensity with a median value of $\delta_{gal} > 5$, 20 times larger than our minimum required overdensity and 70% larger than that of the CCPC2 median value. This illustrates the inherent uncertainty when evaluating overdense regions.

A relatively compact overdensity region can be dwarfed by the diffuse nature of field galaxies along the LOS in some instances (like CCPC-z21-003), resulting in $\delta_{gal} \sim 0$. For four these special cases, we limit the sky aperture of our field counts to the surface area of the overdense region. The overdensities estimated in this way are noted in Table 3. Occasionally, the maximum field length of $\Delta z = 0.15$ will intersect other structures in the same field. As the universe does not contain a simple, smooth distribution of galaxies, it can be quite difficult to measure the “field” surrounding an overdensity. Therefore, the value of δ_{gal} should be taken as an estimate of the strength of the structure, and should not be treated as an absolute metric. As an example of this, we have CCPC-z23-001 and CCPC-z23-005, which are two structures identified by our algorithm whose volumes are fused together in an unbroken ~ 80 cMpc long galaxy distribution. They have 43 and 23 galaxies, respectively, which is significantly larger than the median number of member galaxies in CCPC2 (6). Despite their richness, they have middling δ_{gal} values of 2.0 and 0.7, as their field counts intersect one another. We will discuss the implications of this “superstructure” in Section 3 in more detail.

2.4. Probable Reality of Protoclusters

It is possible that some of the structures we identify are real, and others are chance coincidences. To estimate the probable reality of each overdensity, we refer to simulations of structure formation. Chiang et al. (2013) computed the probability distribution function for overdensities of galaxies becoming a cluster at $z = 0$ (their Figure 8) using the semi-analytic models of Guo et al. (2011) in the Millennium simulation. For different

values of δ_{gal} , they calculated the fraction of volumes that would become halos of $M > 10^{14} M_{\odot}$ (i.e., clusters). They chose to identify galaxy overdensities in the simulation in $[15 \text{ cMpc}]^3$ boxes at redshifts $z = 2-5$ for different galaxy populations. Using their results for galaxies with $M_* > 10^{10} M_{\odot}$ (the most biased tracers of mass), we attribute a conservative probability that each CCPC candidate will collapse into a cluster at $z = 0$ based on its overdensity. If the galaxies within these structures are less biased tracers (e.g., galaxies with $M_* < 10^{10} M_{\odot}$), these probabilities can increase by as much as 40%. These percentages can be found in Tables 1 and 3 next to each CCPC candidate. Table 1 contain systems that have $\geq 85\%$ probability of becoming a cluster at $z = 0$.

There are indications that the overdensities and probabilities we calculate for the CCPC are too conservative. In Section 3, we estimate the number of structures expected within the volume of the CANDELS GOODS-S field. Compared to the candidate protoclusters we identify (and the probability of collapse we assign), the discrepancy is at a minimum of a factor of two smaller, and may be an order of magnitude too low.

It has been noted that a spurious protocluster signal could be the result of the transmission curve of an NB filter (Venemans et al. 2007). A false-positive overdensity is possible if more galaxies are detected at the central wavelength of the filter because of its greater transmittance, while the less responsive tails yield fewer detections. For instance, the FWHM of the Subaru Telescope’s NB filters used to select LAEs at $z = 3.1, 5.7, 6.5$ are only slightly larger than the expected diameter of the most massive protoclusters at these epochs. A smooth distribution of galaxies could appear as an overdensity at the central wavelength with respect to the edges, where “field” galaxies might reside. We find that this is not the case, and the galaxy distributions differ significantly ($> 2\sigma$) from the transmission functions. This serves as a further confirmation that we are detecting actual structure.

To model the possibility of false detections, we adopted a similar approach to Venemans et al. (2007). We ran a Monte Carlo simulation based on the transmission curves of the Suprime-Cam’s NB503, NB816, and NB921 filters with 10^4

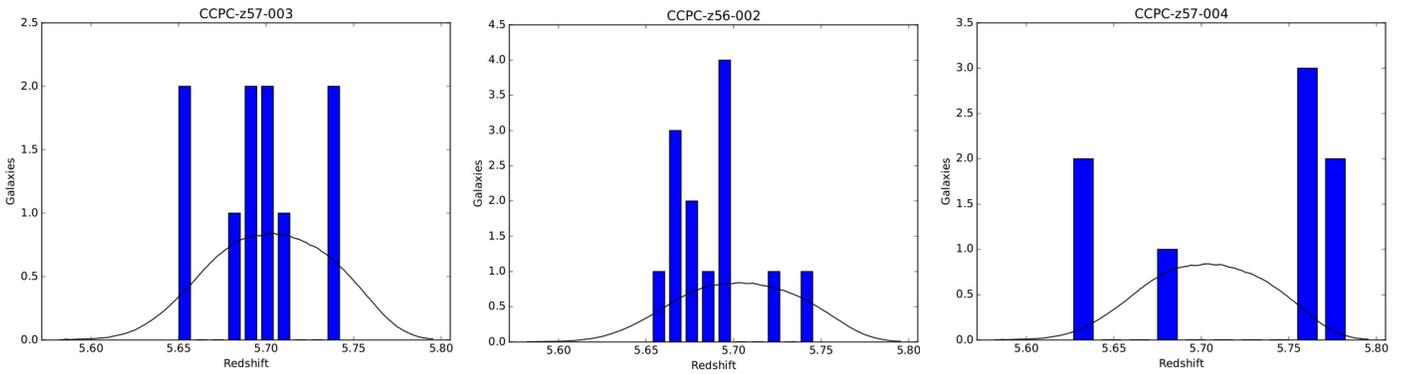


Figure 2. To test the possibility that galaxies observed through a narrowband filter could produce a false-positive protocluster detection from its transmission function, we constructed a Monte Carlo simulation. Shown here are the distribution of LAE galaxies within the NB filter NB816 for three protocluster candidates. Plotted in black is the transmission curve of NB816 at this redshift, arbitrarily normalized to illustrate its shape. The simulated galaxies were distributed following the transmission function of the filter. The simulated distributions were then compared to the mean redshift ($\langle z_{\text{obs}} \rangle - \langle z_{\text{MC}} \rangle$) and dispersion ($\sigma_{z,\text{obs}} - \sigma_{z,\text{MC}}$) of the actual data. The left panel shows one extreme case in which the distribution of real galaxies show only a 0.5σ offset from the Monte Carlo simulation. The center panel is a galaxy distribution in the filter that shares the median value difference (2σ) between the simulated LAEs and the observations, while the right plot shows the most extreme difference example ($\sim 5\sigma$). With the possible exception of the leftmost case, the observed distributions do not follow simply from the shape of the filter transmission.

iterations for all of the 15 protoclusters detected from these observations. For each iteration, the number of galaxies observed within the filter were distributed according to its transmission probability. The difference between the real versus simulated mean redshift ($\langle z_{\text{obs}} \rangle - \langle z_{\text{MC}} \rangle$) and observed versus Monte Carlo redshift dispersions ($\sigma_{z,\text{obs}} - \sigma_{z,\text{MC}}$) were calculated in units of the standard deviation in the simulation. When compared to the actual data of galaxies identified through these filters, the average deviation from the Monte Carlo was 2.4σ (combined $\langle z \rangle$, dispersion deviations). This suggests that the galaxy distribution (and therefore the overdensities) are not merely the result of the NB filter. In Figure 2, the galaxy distributions of three protoclusters (CCPC-z57-004, CCPC-z56-002, CCPC-z57-004) are plotted with respect to the transmittance of NB816. These examples show the full range of the deviations from the Monte Carlo (0.5 – 4.9σ).

2.5. Mass Estimates

The mass of clusters and protoclusters at various redshifts is a stringent probe of cosmological parameters (Ω_m , Ω_Λ , σ_8 and w), as the cluster mass function is tied to these values (Press & Schechter 1974). We attempt to provide two mass estimates in this work. The first is based on the volume and mass overdensity (δ_m) that may collapse by $z = 0$, and the second is a crude virial mass estimate based on the velocity dispersion σ and effective radius of the system. However, these two values do not appear correlated and are highly uncertain.

Steidel et al. (1998) and Venemans et al. (2007) have estimated the expected collapsed masses of protoclusters at high redshift by estimating a galaxy overdensity (δ_{gal}), a volume encompassing the galaxies within the overdensity V , and the critical density of the universe at that redshift ρ_{crit} . Galaxies are assumed to be mere tracers of the dark matter distribution, so it is required to assume a bias parameter b to convert the observed galaxies into a mass overdensity δ_m . Historically, linear bias parameters of $3 \leq b \leq 6$ have been used (Steidel et al. 1998; Venemans et al. 2007), and the matter overdensity can be found via $\delta_m = \delta_{\text{gal}}/b$. In this work, we adopt a bias of $b = 3$. At a redshift of $z = 3$, $\rho_{\text{crit}} = 4.2 \times 10^{10} M_\odot \text{cMpc}^{-3}$ in our assumed cosmology. If

the entire volume is assumed to collapse into a single halo by $z = 0$, that cluster will have an estimated mass of

$$M = \rho_{\text{crit},z} V (1 + \delta_m). \quad (1)$$

Table 2 contains the mass estimates for the protoclusters with the most significant overdensities from Table 1. All system masses are located in Table 5.

There are a number of assumptions that go into this calculation, the most critical of which is the volume estimate. Galaxies on the outskirts of the extended protocluster distribution may not collapse into a single halo by $z = 0$, or be bound to the structure at all, as is seen in simulations (Muldrew et al. 2015). In addition, the SAM used by Contini et al. (2015) to investigate protocluster galaxies suggests that these may be indistinguishable observationally from their field counterparts. Therefore, including these as tracers of the volume that will collapse into a cluster can greatly increase the mass estimate of the system, possibly by orders of magnitude. Some previous works that have utilized this mass estimator include a corrective factor for redshift space distortions (Steidel et al. 1998, 2005). This can result in a difference factor of ~ 2 in some instances. As the volume assumptions can change the mass estimate by orders of magnitude, this space distortion calculation is neglected.

The bias parameter b is an assumed value, and depends on the galaxies used as tracers, in that higher-mass galaxies are more biased tracers of mass. If the galaxies that trace the protoclusters do not have $b = 3$, the overdensity mass estimate will also be systematically affected. Bias parameters can also evolve over time, growing larger at higher redshifts. For LAEs at $z = 3.1$, Gawiser et al. (2007) estimate a value of $b = 1.7$, while LAEs at $z = 4.86$ have $b \approx 3$ (Ouchi et al. 2003). Biases of larger-mass galaxies (LBGs, for instance) can have $b \gtrsim 4$ at $z > 4$ (Ouchi et al. 2004). Based on the no-merger model of Fry (1996), strong evolution of the bias ($\Delta b \gtrsim 3$ from $z = 0$) is predicted for high-redshift, large-bias systems (e.g., $b = 6$ at $z = 5$). For systems of modest bias ($1 < b < 2$), the evolutionary difference is less than unity. If our algorithm is selecting only the most biased sources, the implied mass overdensities may not be sufficient to become clusters. However, if the bias is similar to the roughly constant,

Table 2
CCPC II: Mass Estimates: Best Candidates

| Candidate Name | R_e (Mpc) | σ (km s $^{-1}$) | Virial Mass Estimate ($10^{14} M_\odot$) | δ_m ($b = 3$) | Overdensity Volume (cMpc 3) | Overdensity Mass Estimate ($10^{14} M_\odot$) |
|----------------|-------------|--------------------------|--|------------------------|---------------------------------|---|
| CCPC-z20-002 | 4.5 | 749 | 14.7 | 3.13 | 5361 | 9.8 |
| CCPC-z20-003 | 2.3 | 367 | 1.8 | 6.48 | 125 | 0.4 |
| CCPC-z20-009 | 4.3 | 446 | 4.9 | 4.38 | 589 | 1.4 |
| CCPC-z21-004 | 3.4 | 701 | 9.6 | 2.11 | 519 | 0.7 |
| CCPC-z21-005 | 5.0 | 689 | 13.8 | 3.02 | 799 | 1.5 |
| CCPC-z21-006 | 2.9 | 479 | 3.9 | 6.28 | 58 | 0.2 |
| CCPC-z21-007 | 1.2 | 156 | 0.2 | 5.76 | 13 | <0.1 |
| CCPC-z21-008 | 3.7 | 211 | 1.0 | 3.14 | 176 | 0.3 |
| CCPC-z22-007 | 3.1 | 709 | 8.9 | 2.59 | 3032 | 4.7 |
| CCPC-z23-002 | 2.6 | 783 | 9.3 | 3.82 | 60 | 0.1 |
| CCPC-z23-003 | 2.3 | 711 | 6.8 | 3.91 | 171 | 0.4 |
| CCPC-z23-007 | 1.5 | 442 | 1.7 | 5.54 | 82 | 0.2 |
| CCPC-z24-003 | 2.4 | 399 | 2.2 | 5.02 | 327 | 0.8 |
| CCPC-z24-005 | 4.1 | 770 | 14.1 | 3.09 | 9110 | 15.5 |
| CCPC-z25-002 | 1.3 | 208 | 0.3 | 6.62 | 28 | 0.1 |
| CCPC-z25-003 | 2.2 | 295 | 1.1 | 3.63 | 116 | 0.2 |
| CCPC-z25-007 | 3.1 | 620 | 6.9 | 3.63 | 215 | 0.4 |
| CCPC-z27-012 | 1.3 | 632 | 2.9 | 4.28 | 217 | 0.5 |
| CCPC-z28-011 | 2.6 | 504 | 3.8 | 3.03 | 280 | 0.5 |
| CCPC-z28-016 | 4.0 | 429 | 4.2 | 5.15 | 458 | 1.2 |
| CCPC-z29-009 | 2.2 | 757 | 7.2 | 4.44 | 27 | 0.1 |
| CCPC-z29-011 | 4.1 | 628 | 9.3 | 3.21 | 4579 | 8.0 |
| CCPC-z29-013 | 1.7 | 189 | 0.3 | 4.89 | 56 | 0.1 |
| CCPC-z31-015 | 3.7 | 558 | 6.6 | 2.72 | 2011 | 3.0 |
| CCPC-z32-007 | 2.6 | 504 | 3.8 | 3.82 | 280 | 0.6 |
| CCPC-z33-006 | 1.8 | 578 | 3.5 | 5.83 | 115 | 0.3 |
| CCPC-z33-007 | 3.2 | 818 | 12.3 | 3.84 | 101 | 0.2 |
| CCPC-z33-010 | 3.4 | 611 | 7.3 | 3.48 | 730 | 1.4 |
| CCPC-z36-007 | 3.3 | 634 | 7.7 | 5.07 | 165 | 0.4 |
| CCPC-z44-003 | 1.4 | 575 | 2.6 | 4.60 | 456 | 1.0 |
| CCPC-z26-001 | 0.7 | 245 | 0.2 | 2.44 | 3 | <0.1 |
| CCPC-z26-006 | 2.2 | 366 | 1.7 | 2.41 | 139 | 0.2 |
| CCPC-z27-008 | 3.5 | 463 | 4.3 | 2.42 | 5570 | 7.9 |
| CCPC-z31-008 | 2.6 | 541 | 4.3 | 2.57 | 1038 | 1.6 |
| CCPC-z31-017 | 3.7 | 462 | 4.5 | 2.42 | 10025 | 14.6 |
| CCPC-z34-006 | 4.0 | 878 | 17.7 | 2.61 | 270 | 0.4 |

Note. The mass estimates for the largest overdensities in CCPC2 (same order as Table 1). The effective radii (R_e , in physical units) and velocity dispersions (σ) are used to compute a virial mass estimate, with the caveat that these systems are not expected to be in equilibrium at the relevant redshifts. We obtain the mass overdensity δ_m by assuming galaxies are linearly biased tracers of mass with a slope of $b = 3$. From this and the volume contained within the galaxy distribution we can compute an expected mass of the cluster (Steidel et al. 1998). This assumes that the total volume will collapse into a single halo at $z = 0$.

unevolving value of Gawiser et al. (2007), our candidates might be more significant than we claim.

The mean estimated collapsed mass of candidates in CCPC2 is $M_{z=0} = 1.8 \times 10^{14} M_\odot$, consistent with low-mass clusters found in the local universe. Systems of this mass comprise roughly 70% of the population of clusters in the Millennium simulation (Chiang et al. 2013), while clusters of $M_{z=0} > 10^{15} M_\odot$ should represent only 2% of cases. However, our search methodology is not expected to be mass-blind, and preferentially selects the highest overdensity systems.

For clusters at low redshift ($z \sim 0$), a traditional method of estimating a mass was to assume the system was virialized, measure the velocity dispersion (σ) and effective radius R_e , and compute the system mass via

$$M = \frac{2\sigma^2}{G} R_{\text{hms}}, \quad (2)$$

where $R_{\text{hms}} \sim 1.25 R_e$. However, this assumption is not expected to hold at higher redshifts in Λ CDM. At 3 Gyr after

the Big Bang ($z \sim 2$), analysis of the Millennium simulation shows that the progenitors of the most massive clusters at $z = 0$ ($M_{z=0} > 10^{15} M_\odot$) have a dark matter halo of $\geq 10^{14} M_\odot$, while the lowest mass clusters at the present day may not have assembled this mass until $z \approx 0.6$ (Chiang et al. 2013). Prior to this epoch, it is likely that only subhalos have virialized.

For each object in the CCPC, we have calculated the effective radius (R_e) in which 50% of the total protocluster members reside, as well as the velocity dispersion of the system σ , and computed a “virial” mass estimate. Mass estimates for the entire CCPC2 list are in Table 5, while the most overdense protoclusters can be found in Table 2. We again urge caution in interpreting these results, as it is unlikely that such extended, diffuse systems are in virial equilibrium. We use the “virial” term only because we utilize the virial mass equation. As there are few protocluster members in some CCPC systems, their diffuse nature can imply large values of both R_e and σ , as is the case for CCPC-z24-007. It has only four galaxy members (the minimum number) and an effective radius of 4.8

(physical) Mpc, a dispersion of 760 km s^{-1} . These yield a virial mass estimate of $1.6 \times 10^{15} M_{\odot}$, a factor of two larger than the average estimate in CCPC2 ($\sim 8 \times 10^{14} M_{\odot}$), despite its minimal richness.

In contrast, there are a few systems (CCPC-z20-005, CCPC-z21-011, CCPC-z22-002) that have a large number of candidate galaxy members ($N \geq 18$) with large implied virial masses ($\geq 10^{15} M_{\odot}$). These systems are at an epoch ($z \leq 2.3$) that can theoretically host virialized clusters (Chiang et al. 2013). With the significant increase in richness, these mass estimates may be at least more physically meaningful than the previous example (CCPC-z24-007), although these would still be an order of magnitude larger than predicted. Further examination and discussion of the mass estimates (via mass overdensities and velocity dispersions) can be found in Section 3.3.

2.6. Objects of Interest

We have compiled a list of targets that are of potential significance, but violate our strict spectroscopic redshift criteria. These objects of interest (OOIs) are generally the result of narrow-band or photometric redshift observations, but otherwise fulfil the requirements of a CCPC target. The last two entries (OOI-z65-001 and -002) are spectroscopic galaxies, but had no field sources with which to calculate a δ_{gal} value. Some of these will likely prove to be fictional if targeted spectroscopically. These are listed in Table 6.

3. DISCUSSION

3.1. Tests of Structure

The method in which we identify structures is relatively simple. In CCPC1 (Franck & McGaugh 2016), we performed a number of tests on the algorithm in which candidate protoclusters were identified. In addition, the significance of the candidates were also evaluated. It is trivial to list positions of galaxy associations, but much more difficult to find physical systems. We have three overarching diagnostics of structures, each with sub-tests for significance: (1) the CCPC systems exist as overdensities, (2) they are statistically distinct spatially when compared to the “field,” and (3) their number densities in deep surveys are comparable to expectations from large simulations. That we recover a number of previously identified structures with our algorithm is an additional confirmation of its fidelity.

One of the criteria for CCPC protoclusters is that candidate systems must show an overdensity of galaxies ($\delta_{\text{gal}} \geq 0.25$) when compared to the local field along the LOS ($\Delta z \pm 0.15$). This criteria is simply a lower limit, and the median overdensity is $\delta_{\text{gal}} = 2.9$. Within the literature, previously identified protoclusters (and simulations) at these redshifts show similar overdensities (Table 5 and Section 3.6, respectively, in Chiang et al. 2013).

Furthermore, these overdensities can be observed visually as spikes of galaxies along the LOS in $N(z)$ plots (Figure 4). These plots were constructed using the full aperture of the sky search radius ($R = 20 \text{ cMpc}$). Many systems, such as CCPC-z20-003 (Figure Set 8), have clear overdensities along the LOS. Other overdensities, particularly those with centrally condensed galaxy distributions on the sky (e.g., CCPC-z20-008), may appear as little more than noise. This underlines the importance of having more than a single protocluster identifier (e.g., only $N(z)$ spikes, δ_{gal} , or a minimum number of galaxies).

It is of interest that the δ_{gal} values and CCPC galaxy members show little correlation. It is logical that spectroscopic surveys which have deeper limiting magnitudes would find more galaxies in both the field and in structures, all galaxies being equal. However, if protoclusters are indeed regions of enhanced mass assembly (Casey et al. 2015), rapid star formation or galaxy evolution (Hatch et al. 2011; Clements et al. 2014), it would be reasonable to assume that some manifestation of this behavior would present itself.

In fact, some of the largest overdensities ($\delta_{\text{gal}} \geq 10$) have fewer than 10 galaxy members within the search volume. For many of these cases (e.g., CCPC-z21-007 and CCPC-z23-007), the Poissonian uncertainties are large ($\geq \pm 10$) because of the small number of galaxies in the overdensity calculation (especially field galaxies). Some cases of these low- N , high- δ_{gal} cases are, interestingly, previously identified structures.

CCPC1 and CCPC2 have nearly identical median overdensities of $\delta_{\text{gal}} \sim 2.0, 2.9$ over the entirety of the redshift range. This is not a perfectly direct comparison, as the CCPC1 criteria required three or more galaxies to be found within 2 arcminutes of the center of the search radius. Potentially, CCPC1 may have been selecting the minority of protoclusters that have dominant main halos (Muldrew et al. 2015), which would represent stronger overdensities. Also, there is an expectation of some δ_{gal} evolution with redshift (Chiang et al. 2013), but with larger variations depending on the galaxies observed as tracers.

Despite these differences in redshift and methodology, the protoclusters at $2.74 < z < 3.71$ in the CCPC2 have a median $\delta_{\text{gal}} \sim 1.7$ and are consistent within the uncertainties of $\delta_{\text{gal}} \sim 2$ for CCPC1 and CCPC2. This is of two-fold importance: (1) there were 66 overdensities left undiscovered in the redshift range of CCPC1 (which contained 43 structures) by requiring an initial core group of galaxies to be a criterion, and (2) that the slightly greater median δ_{gal} between CCPC1 and CCPC2 suggests we are not admitting poorer candidates by removing this step.

Galaxy overdensities should not only exist along the LOS (as observed in $N(z)$ plots and $\delta_{\text{gal}} > 0$), but their spatial distributions should also be distinct from the field. As protoclusters exist as very extended, diffuse systems on the order of tens of cMpc (Figures 1 and 2 in Muldrew et al. 2015), their profiles will naturally look more similar to isolated galaxies than $z = 0$ clusters. However, in CCPC1 we showed that there were significant differences between the surface density and cumulative distribution of field and protocluster galaxies.

In this work, we present similar findings based on the mean surface density of all CCPC2 systems (Figure 3). In order to capture the essence of “field” galaxies, we used the initial list of $> 47,000$ galaxies from which we identified the candidate protoclusters at $z > 2$ to search for any companion galaxies within a radius of 20 cMpc, and $\pm 20 \text{ cMpc}$ along the LOS. This mirrors the search volume of the CCPC algorithm. Included in this galaxy list are all of the CCPC2 member objects. The mean sky surface density (galaxies cMpc^{-2}) is shown as a red dashed line in Figure 3. This is consistent with a single galaxy found within $R \leq 1 \text{ cMpc}$ ($\Sigma = 0.33 \text{ cMpc}^{-2}$). This is labeled “All Galaxies” instead of “Field Galaxies,” as we did not separate out isolated galaxy systems from overdense regions. By comparison, there are more mean galaxies in the central regions of CCPC systems (black line), and a shallower slope

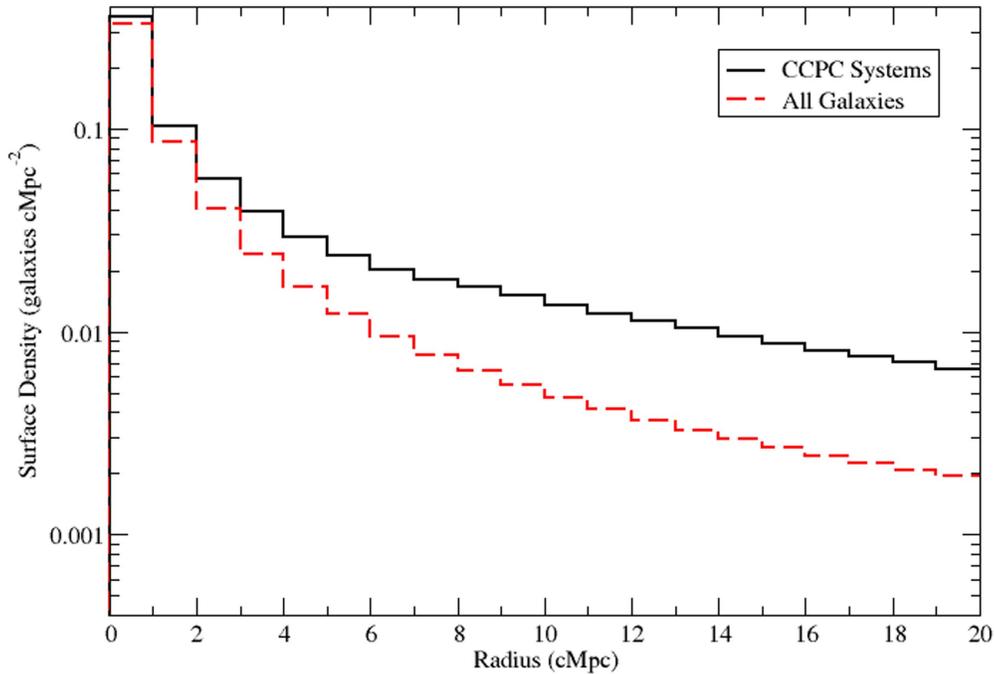


Figure 3. The mean galaxy density on the sky of CCPC2 galaxies is shown in black. In contrast is the mean surface density profile of all galaxies in our initial list ($>40,000$ spectroscopic sources with $z > 2.00$), which is shown as the red dashed profile. The “All Galaxies” distribution contains CCPC2 galaxies. This “field” proxy is consistent with there being a single galaxy in the center of a $R = 20$ cMpc search radius with few companions. A two-sided KS test (0.6) suggests that these are distinct distributions.

(-1.3 in log–log space) compared to the field galaxy slope (-1.7). A two-sample Kolmogorov–Smirnov (KS) test value of 0.6 shows that these are distinct distributions at the 99.93% level.

Figure 4 illustrates the number of galaxies as a function of search radius $N(r)$ for each individual CCPC2 overdensity in the top panel. As in Figure 3, the red dashed line is representative of field galaxies, with clear differences in the distribution and mean number of sources at different comoving radii. The distribution of protocluster galaxies is highly variable, with some overdensities being very concentrated while others have lower central concentrations, but continually add galaxies to large radii. Plotting the number distributions instead of cumulative distributions (as in CCPC1) highlights the differences in survey widths more effectively, as some $N(r)$ profiles flatten out well before the maximum search radius of $R = 20$ cMpc. It is therefore important to not infer a “characteristic” distribution of galaxies in protoclusters from such a plot, as some spectroscopic galaxy surveys will only target the innermost regions of structure, while others will trace out to larger radii. In addition, simulations suggest that there exists a menagerie of galaxy distributions within these high-redshift systems, even among protoclusters that will have the same mass at $z = 0$ (Chiang et al. 2013; Muldrew et al. 2015).

The bottom panel of Figure 4 shows the mean “All Galaxies” distribution as a black dashed line (instead of a red dashed line in the top panel), while CCPC2 galaxies in various redshift bins are shown in blue ($2 < z < 3$), green ($3 < z < 4$), and red ($4 < z < 6.57$). On average, “All Galaxies” will have roughly 1.5 companion galaxies surrounding it within the fixed CCPC volume, while at a minimum a CCPC system will have three other companions, and on average more than seven.

A KS test of the radial distributions of galaxies between “All Galaxies” and the mean CCPC2 galaxies gives a value of 0.75,

more significantly different than the surface density comparison (KS = 0.60). The various redshift bins, within the considerable scatter illustrated by the top panel, show little evolution. Simulations suggest there are large distribution variations of individual systems even at the same mass and redshift (Chiang et al. 2013; Muldrew et al. 2015). When coupled with the heterogeneous spectroscopic data (survey width/depth differences) from which we draw our candidates, the lack of meaningful $N(r)$ variations is not particularly surprising.

The most significant difference in selecting protocluster candidates between CCPC1 and CCPC2 was that the former contained a “group finding” intermediate step. Initial groups were selected by requiring at least three galaxies within $2'$ of the search center, and within a $\Delta z = 0.03$. Requiring this intermediate step does not appear to affect the quality of sources, as the median overdensities between CCPC1 and CCPC2 are equivalent ($\delta_{\text{gal}} \sim 2$). Interestingly, requiring a centrally concentrated group of galaxies does not seem to have a significant effect on the surface density profiles (KS = 0.35 between CCPC1 and 2). Only the inner regions ($R \sim 1$ cMpc) show significant differences, with CCPC1 having a larger surface density (as a criterion) of $\Delta \Sigma = 0.12$ cMpc $^{-2}$. We conclude that the structures identified using the selection methods of CCPC1 and CCPC2 are not significantly different from one another, and both are selecting plausible protocluster candidates.

The CANDELS GOODS-S field is a deep, multi-wavelength field in which 27 CCPC structures have been identified. Of these, nine were identified in Franck & McGaugh (2016) (in the redshift range $2.74 < z < 3.71$). It is the deepest, widest field from which we draw candidate spectroscopic galaxies to identify protoclusters. The majority of the spectroscopic footprint falls on an area of the sky of roughly 0.4° on a side, which corresponds to a box with sides approximately 35 cMpc

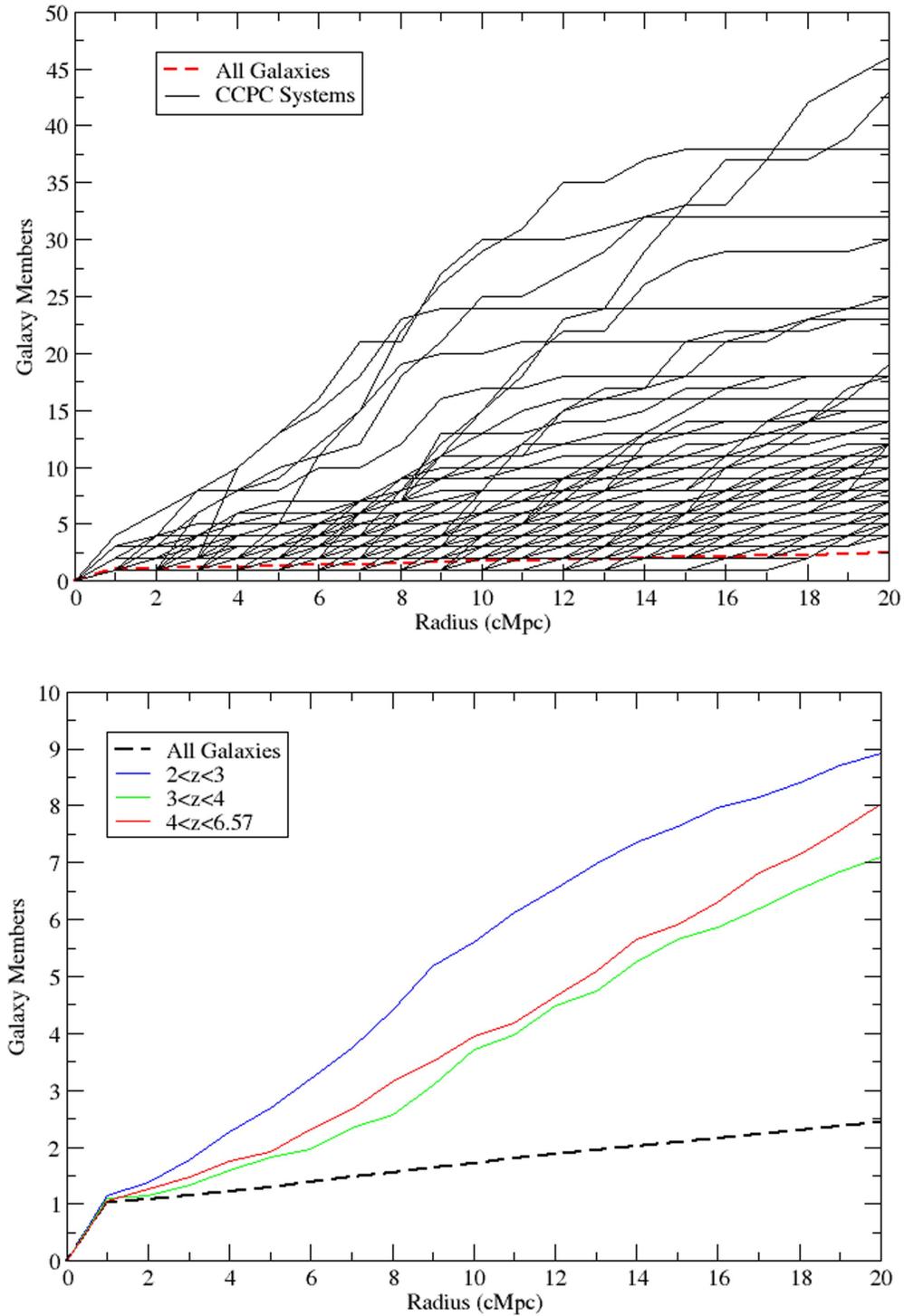


Figure 4. Top: the distribution of galaxies for each of the 173 CCPC2 systems are shown in black as a function of distance from the center of the search radius in comoving Mpc. The mean distribution of 44,000 spectroscopic galaxies from which we identified structures is shown as a dashed red line. For the CCPC systems, there is a clear difference in both the number and distribution of galaxies. In many cases, the survey edge can be seen as a flat line at $R < 20$ cMpc. Bottom: at different bins in redshift, we have plotted the mean number of galaxies as a function of radius, compared to the black dashed line representing all galaxies at $z \geq 2$. As shown by the top panel, there is considerable variation in the distribution of galaxies. The difference in survey widths also prevents meaningful analysis of the comoving distribution of the galaxies, especially at radii > 10 cMpc. Primarily, this plot is effective at illustrating the difference between field galaxy distributions and those within protoclusters at all redshifts. A KS test of the two mean distributions (“All Galaxies” vs. CCPC galaxies) gives a value of 0.75, indicative that these two populations are not from the same group.

wide at $z = 2$. At $z = 5.7$, the box side length increases to ~ 55 cMpc in our assumed cosmology, with a length along the LOS ($z = 2 \rightarrow 5.7$) of almost 3 cGpc. This length is the expanse of the field in which we identify protoclusters.

From this pencil beam survey, we estimate the volume that we have probed to be $\sim 5.89 \times 10^6$ cMpc³. The Millennium Simulation has a volume of $[500 h^{-1} \text{cMpc}]^3$ (Springel et al. 2005), which is almost two orders of magnitude larger

Table 3
Candidate Cluster and Protocluster Catalog (CCPC): All Candidates

| Candidate Name | R.A. (degree) | Decl. (degree) | Redshift (z_{avg}) | σ_z | N | $N_{R \leq 10}$ cMpc | Overdensity (δ_{gal}) | Cluster Probability (%) | Q | Recovered Reference |
|----------------|---------------|----------------|-------------------------------|------------|-----|----------------------|---------------------------------------|-------------------------|-----|---------------------|
| CCPC-z20-001 | 189.00 | 62.18 | 1.997 | 0.006 | 30 | 15 | 5.07 ± 1.52 | 79.4 | 1 | ... |
| CCPC-z20-002 | 222.20 | 8.92 | 2.002 | 0.008 | 11 | 7 | 9.38 ± 5.34 | 100.0 | 1 | 8 |
| CCPC-z20-003 | 29.62 | -25.05 | 2.018 | 0.004 | 10 | 10 | 19.43 ± 13.06 | 100.0 | 1 | 1, 2 |
| CCPC-z20-004 | 36.62 | -4.52 | 2.023 | 0.006 | 6 | 1 | 2.23 ± 1.67 | 31.3 | 1 | ... |
| CCPC-z20-005 | 52.97 | -27.80 | 2.033 | 0.008 | 25 | 8 | 0.99 ± 0.25 | 2.5 | 1 | ... |
| CCPC-z20-006 | 189.19 | 62.29 | 2.031 | 0.008 | 8 | 6 | 0.48 ± 0.22 | 2.5 | 1 | ... |
| CCPC-z20-007 | 149.24 | 69.65 | 2.049 | 0.006 | 4 | 4 | $[2] \pm 4.16$ | 11.3 | 1 | 3 |
| CCPC-z20-008 | 188.97 | 62.23 | 2.086 | 0.007 | 15 | 9 | 1.30 ± 0.52 | 11.3 | 1 | ... |
| CCPC-z20-009 | 150.04 | 2.21 | 2.098 | 0.005 | 10 | 4 | 13.15 ± 6.54 | 100.0 | 1 | 4, 5 |
| CCPC-z21-001 | 246.49 | 26.75 | 2.107 | 0.007 | 5 | 4 | 4.17 ± 3.02 | 75.6 | 1 | ... |

Notes.

Please refer to the machine-readable table for a detailed description of each column.

^a Counts of field galaxies were limited to the surface area of the overdense region.

(This table is available in its entirety in machine-readable form.)

than the GOODS-S data. Chiang et al. (2013) identified 2832 clusters at $z = 0$ with masses $M > 10^{14} h^{-1} M_{\odot}$ within their analysis of the simulation. This corresponds to a number density of $7.8 \times 10^{-6} \text{ cMpc}^{-3}$. Therefore, one would expect to find 46 clusters at $z = 0$ in the volume of the GOODS-S field. As we have only identified 27 systems in this region, our algorithm is probably not over-identifying structure. Using the number density of low-mass clusters ($M_{z=0} < 3 \times 10^{14} M_{\odot}$) in Millennium, there should be an estimated 32 such systems in this deep field (Chiang et al. 2013), while 14 Virgo-mass or larger ($M_{z=0} \geq 3 \times 10^{14} M_{\odot}$) protoclusters are expected within this volume. It is probable that our algorithm is only identifying some of the richest overdensities that will collapse into Virgo or larger mass systems at $z = 0$, while missing many of the smaller protoclusters in this field.

In all cases of protocluster candidates within the GOODS-S field, the probability that these objects will collapse into clusters by $z = 0$, based on the overdensities of Chiang et al. (2013), is considerably less than 100% (see Table 3). If we sum their fractional collapse estimates (e.g., a candidate's 10% chance of collapse can be approximated as $\frac{1}{10}$ of a cluster), for the entire CCPC we have identified only ~ 3.7 clusters from $2 < z < 5.7$ in the CANDELS GOODS-S survey. As we expect more than 40 structures in this probed volume, there exists a serious discrepancy. One explanation could be that our conservative estimates for both δ_{gal} and its application in determining a collapse probability from the work of Chiang et al. (2013) is much too stringent. This is a probable scenario, as explained in the outline of overdensities and probabilities within Section 2. Another option could be that the number density estimates are for all clusters, while we are mainly identifying higher-mass systems. However, 14 Virgo-like protoclusters are expected in this volume, and our probabilities are an order of magnitude smaller than that. It is also plausible that the CANDELS GOODS-S field is not accurately represented within the Millennium simulation. Indeed, a combination of these elements are probable to span the gap of the excess number of simulated systems.

As mentioned earlier, many of these deep, pencil-beam surveys from which we identify protoclusters have sky widths smaller than the expected size ($\sim 20 \text{ cMpc}$) of the most massive protoclusters (Muldrew et al. 2015). Therefore, if the

overdensity does not significantly fall within the footprint of the spectroscopic data, it may be missed by our search method. Intriguingly, Kang & Im (2015) delved into both the GOODS-S and GOODS-N combined spectroscopic and photometric data and found an excess of structures in the redshift range $0.6 \leq z \leq 4.5$ when compared with the Millennium Simulation protoclusters. While in the CCPC we have limited our comparison to the number density, they included the mass of their structures, finding a factor of 5 more systems with $M > 7 \times 10^{13} M_{\odot}$. Kang & Im (2015) provide some plausible explanations for the overabundance (elements of the input physics in the models could be simply incorrect, Λ CDM may not be an accurate representation of the universe, over-estimation of masses), but no definitive diagnosis for this complex problem.

That we derive the opposite conclusion from the same data set, similar search methodology, and identical simulation is puzzling. It is possible that the precision of photometric redshift measurements produces an increase in false-positive structure detections. This seems unlikely, as we found more overdensities (24 versus 9) in the shared redshift space with Kang & Im (2015) in the GOODS-S field. We find it more likely that the authors' method is sufficiently distinct from our own so as to make it difficult to compare results.

3.2. Poissonian Expectation Model

A further confirmation that these associations are more than chance groups of galaxies can be found by estimating the number of false positives that should be expected in a smooth density field (i.e., lacking structure) along the LOS that could arise from Poissonian fluctuations. We can utilize the analytical formula

$$N_M = N_g e^{-2nl} [1 - e^{-nl}]^{M-1} \quad (3)$$

from the toy model of Sheth (2001), which approximates the number (N) of Poissonian fluctuations of M galaxies that would manifest as protoclusters. This model uses the number density along the LOS, n , for N_{gal} total galaxies, separated by a linking length l . The CANDELS GOODS-S field is approximately a 1D pencil beam survey in the context of this work ($\sim 40 \times 40 \times 3000 \text{ cMpc}$ from $2 < z < 5.7$).

The number density of spectroscopic sources varies from $2 < z < 5.7$, which makes it a linearly decreasing density field as a function of redshift, unlike the toy model. At $2 < z < 3$, $n = 0.76$ galaxies cMpc^{-1} , while over the full range ($2 < z < 5.7$), this decreases to $n = 0.40$ galaxies cMpc^{-1} . We adopt the latter value ($n = 0.40$ galaxies cMpc^{-1}) as the estimate of the number density.

We require a suitable estimate of the linking length l between galaxies in this survey. We can measure the strength of clustering at different length scales by using the auto-correlation function (Davis & Peebles 1983).

$$\xi(l) = \frac{N_{DD} n_R}{N_{DR} n_D} - 1. \quad (4)$$

This compares the number of pairs of actual galaxies (N_{DD}) found within shells separated by $l \pm \Delta l$ in the GOODS-S field, to the number of Data-Random pairs of galaxies (N_{DR}). When the value of $\xi(l)$ crosses $\xi = 1$, the clustering strength at length l is said to transition from strong to weak. At $l = 1.1\text{--}1.75$ this occurs within the spectroscopic data set. This is similar to the linking length measured between $2.74 < z < 3.71$ found in CCPC1 (Franck & McGaugh 2016).

The CCPC algorithm does not utilize linking lengths to identify structure, as l would vary from survey to survey. To similarly match the fluctuations in the toy model of Sheth (2001), we identified structures using a one-dimensional friends-of-friends (FoF) algorithm (Huchra & Geller 1982). We also ran 500 Monte Carlo simulations of the field and computed the mean number of groups found using the FoF algorithm as a check on the toy model. For $l = 1.2\text{--}1.75$ cMpc, an excess number of groups are identified above the Poissonian expectation for any choice of the minimum member galaxies ($M \geq 2$). There were 15 FoF structures found within the CANDELS data set using separation $l = 1.2$ cMpc with $M \geq 6$ galaxies and 6 associations of $M \geq 8$ members. The toy model predicts that 4 and 0.6 such groups of galaxies should exist (respectively) from Poissonian fluctuations, while the Monte Carlo finds a mean of 1.4 and 0.08 systems of such richness. Choosing a larger separation of $l = 1.75$ cMpc reveals an excess of 21 FoF groups over the expectation of Sheth (2001) for associations of 6 or more members, and an excess of 25 compared to the Monte Carlo value. For groups of 10 or more galaxies with separation of $l = 1.75$ cMpc, there should be fewer than one chance fluctuation in the Poissonian and Monte Carlo models, while 5 systems are identified in the GOODS-S volume.

The clear excess of FoF groups using a range of values for l and M compared to what would be expected stochastically illustrates that there are physical associations within these data. Values of N_m and l are unique to this survey and do not apply to CCPC protoclusters outside of the CANDELS GOODS-S sample. This example is only used to illustrate that an excess of clustering is found over Poissonian fluctuations in a simple toy model.

3.3. Mass Estimates

In Section 2, we outlined two distinct methods to estimate the mass of protoclusters at these high redshifts. Unlike the methods utilized for “nearby” clusters at $z \leq 1$ which use

signatures from a massive $M > 10^{14} M_\odot$, virialized halo (e.g., SZ effect, X-ray emission), protoclusters at $z \geq 2$ are not expected to have main halos of this magnitude (Chiang et al. 2013; Muldrew et al. 2015). More uncertain means must be employed to provide some metric of the mass of these systems.

One method used by Steidel et al. (1998) and Venemans et al. (2002) is to calculate the volume of the overdensity that will collapse into a cluster at $z = 0$. The mass is simply the density of the universe (ρ_{crit}) multiplied by the mass overdensity ($\delta_m = \delta_{\text{gal}}/b$) and volume of the system (Equation (1)). The mean overdensity mass is $M_{z=0} \sim 1.8 \times 10^{14} M_\odot$ for the CCPC2 catalog. As the majority of clusters ($\sim 70\%$) in the universe are expected to be of this mass (Chiang et al. 2013), this appears reasonable. We have also made the most conservative estimates of the volume and overdensities (δ_m) which would enhance this expectation.

There does not appear to be a significant trend in decreasing mass with increasing redshift. There is considerable scatter in the $\delta_{\text{gal}} - z$ distribution, which subsequently persists into the mass estimates. Most of the sources with the highest probability of collapsing into a structure (Table 2) are at low redshift and can have masses in excess of $M_{z=0} > 5 \times 10^{14} M_\odot$. A large mass estimate for CCPC-z65-005, the highest redshift system, is an example of the lack of mass evolution. Located at a redshift of $z = 6.56$, its large volume ($\sim 7000 \text{ cMpc}^3$) and overdensity imply a collapsed mass of nearly $M_{z=0} \approx 6 \times 10^{14} M_\odot$. The CCPC1 has a mean mass estimate of $2.5 \times 10^{14} M_\odot$, only slightly larger than CCPC2.

It should be noted that the δ_{gal} estimator relies on a number of assumptions. Primarily, the tracer of both the volume and matter overdensity are galaxies, which are not a significant contributor to the density of the universe in Λ CDM. We must assume that all galaxies, especially those at the outskirts that define the volume of the overdensity, will collapse by $z = 0$. Numerous simulations have shown that many galaxies within the comoving volume of the structure may not be bound to the cluster by $z = 0$ (Muldrew et al. 2015). They also lack physical properties (e.g., enhanced SFRs, stellar colors, etc.) that could distinguish them as outliers observationally in the SAMs of Contini et al. (2015). We also must adopt a linear bias parameter b that translates a galaxy overdensity into a mass overdensity ($\delta_m = \delta_{\text{gal}}/b$). This bias value is assumed to be in the range of 3–6 at these redshifts (Steidel et al. 1998), and is dependent on the types of galaxies observed. Furthermore, the values of δ_{gal} are highly dependent on how the field is defined and the scales at which the overdensity exists, as we have pointed out here and in CCPC1.

The “virial” mass estimate also has a number of systematic uncertainties. These systems are the diffuse, primordial manifestations of clusters and are not expected to be in equilibrium in Λ CDM. In the Millennium Simulation, Chiang et al. (2013) found that the first $M \geq 10^{14} M_\odot$ halos are not present until $z \leq 2.3$, and may not be virialized until a dynamical time later. However, there are some indications that subhalos in protoclusters can be virialized, as observed by Venemans et al. (2007), Shimakawa et al. (2014) and Topping et al. (2016). These subhalos can be seen as bimodal distributions of protocluster galaxies along the LOS. Intriguingly, Wang et al. (2016) recently identified a cluster core emitting X-rays at $z = 2.5$. The cosmological implications of this discovery are uncertain.

Virial equilibrium is unlikely to hold, but we can query the data to see what mass is implied. Equation (2) is the familiar virial mass estimate that only requires an effective radius ($R_{\text{rms}} = 1.25R_e$) and velocity dispersion (σ) for a system in equilibrium. Based on the position of galaxies in the plane of the sky, we have calculated the radius in which 50% of the member galaxies can be found. The mean R_e is 2.8 ± 1.0 (physical) Mpc. The relatively small number of galaxies (minimum of four) in our criteria is a further source of significant uncertainty, in that values of R_e can be based on only two systems. The dispersion velocity is simply the standard deviation from the mean redshift of the system, and is on average 653 km s^{-1} . This is typical of dispersions found for other protoclusters, as compiled in Table 5 by Chiang et al. (2013). We note that the few galaxy members involved in the calculation of the dispersion of these systems could introduce a large bias. Some candidates with $N = 4$ members have σ_z values based on only two galaxies. Richness alone does not appear to be a significant driver, however. The average velocity dispersion for systems with $N \geq 23$ members is 661 km s^{-1} , the same as the mean value of the entire group.

Interestingly, these large velocity dispersions, like those at $z \geq 3$ as found in Venemans et al. (2007), CCPC1, and here, appear to be larger than those predicted by simulations. Chiang et al. (2013) report that dispersions along the LOS for overdensities at $z = 3$ in the Millennium Simulation are $400 \pm 60 \text{ km s}^{-1}$ for the progenitors of $M_{z=0} \geq 10^{15} M_\odot$ clusters. Our typical redshift uncertainty ($\sigma \sim 0.001$) represents a velocity uncertainty of 75 km s^{-1} (at $z = 3$), which cannot account for the $+100 \text{ km s}^{-1}$ dispersion excess with respect to the simulated systems. Similarly, Venemans (2005) compared the protocluster dispersions in their sample to simulated dark matter halos of clusters within a similar window size. For $\sigma_{z=0} = 1000 \text{ km s}^{-1}$ systems (e.g., the Coma cluster), the dark matter velocity dispersion was found to be systematically lower than what was observed in the protoclusters at nearly all redshifts sampled.

It could be possible that galaxies not bound to the protocluster, or those at the outskirts, are boosting the dispersion significantly above expectations. Contini et al. (2015) plot the velocity field for simulated protocluster galaxies in their Figure 2 for a system with $M_{z=0} \geq 10^{15} M_\odot$. At $z = 3$, the members near the central galaxy ($R < 5 \text{ cMpc}$) have velocities of $200\text{--}400 \text{ km s}^{-1}$. Galaxies more than 20 cMpc away can have velocities in excess of 1000 km s^{-1} (Contini et al. 2015). Survey sizes within the CCPC1 and 2 vary in both width and depth, yet this phenomenon persists. In our smallest survey volumes, the inner regions should contain the brightest, most massive galaxies and the velocities are expected to be the smallest (Contini et al. 2015). Greater spectroscopic completeness within some of the richest overdensities might shed light on this mystery. This would allow a more apt comparison to simulated protoclusters, which “observe” all galaxies and not just the brightest sources.

As an attempt to mitigate the weight of outliers for small numbers of galaxies when calculating a velocity dispersion for protoclusters, Venemans et al. (2007) utilized the biweight location estimator (Beers et al. 1990). The dispersions discussed previously, both here and in CCPC1, were simple standard deviations. We employed this biweight estimator in the AstroPy package (Astropy Collaboration et al. 2013) and iteratively solved for the velocity dispersion for each CCPC system. The median value of these biweight dispersions across the CCPC2

catalog is 717 km s^{-1} , which is $< 1\sigma$ larger than a simple median value of all dispersions (677 km s^{-1}). Outlier galaxies do not appear to be biasing our results significantly.

We note that unlike CCPC1, this work did not require a knot of ≥ 3 galaxies to exist near the central search regions. This had the occasional effect of “centering” the galaxy distribution. For CCPC2, we sought only to maximize the number of galaxy members within the search radius of $R \leq 20 \text{ cMpc}$. Therefore, some R_e values could possibly be reduced if a different protocluster center were chosen (as in CCPC-z28-017), whereas others would be unaffected by this method change. However, for the bulk of the CCPC2 population, this is unimportant. The mean $R_e = 2.2 \text{ Mpc}$ for CCPC1 is only marginally smaller than CCPC2 and within its standard deviation. The velocity dispersions are essentially equivalent for CCPC1 and 2 as well (668 and 653 km s^{-1} , respectively).

By limiting the distance we probe for structure along the LOS, the calculated velocity dispersions are decreased, as one would expect. For example, restricting the search redshift range to half its original length ($\Delta z \pm 10 \text{ cMpc}$), only 32 protocluster candidates exceed the 400 ± 60 value for the expected dispersion of a $M_{z=0} \geq 10^{15} M_\odot$ cluster at $z = 3$ (Chiang et al. 2013). These anomalous systems also have only ~ 6 galaxies, on average. To establish an expectation of the role window size affects the observed dispersion, we designed a Monte Carlo simulation of a protocluster at $z = 3$. $N = 6, 10$, and 200 galaxies were randomly distributed 10^4 times with varying LOS distances ($1 < d < 20 \text{ cMpc}$). We measured the mean dispersion and standard deviation of the simulation for each LOS distance. The Monte Carlo produces a linear relationship between LOS window size and velocity dispersion. When compared to the CCPC2 dispersion values in different window sizes, as well as the expectation value from the Millennium Simulation (Chiang et al. 2013), the Monte Carlo results were not distinct at a statistically significant level. These results suggest that the velocity dispersion excess could merely be the result of window size rather than a physical characteristic of the system.

The Monte Carlo simulation is only an effective modeling tool if we assume the correct, underlying distribution of galaxies and their velocities. For instance, we can make a simple assumption that all galaxies in the simulation have a true distance corresponding to $z = 3$, and then imprint a Gaussian distribution of velocities of $\sigma = 400 \text{ km s}^{-1}$. These results can fit the data point of CCPC2 with a window size of $\pm 10 \text{ cMpc}$, as well as the expectation value of Chiang et al. (2013). We can modify the simulation again by allowing galaxies to be normally distributed within a specified radius, and then applying a random velocity (a pseudo-virial distribution) to each one individually. This can provide a fit to the CCPC2 for the full window size of the redshift distribution, but no other data points. Applying only infalling velocities to spatially dispersed galaxies can also match the observations. These examples are meant to illustrate the underlying degeneracy of the simulation and the complexity of the problem. Fundamentally, the menagerie of simulated protocluster sizes, velocity fields, and evolutionary states (Chiang et al. 2013; Contini et al. 2015; Muldrew et al. 2015) make the prospect of a direct comparison of dispersions to real data a daunting task, even at a specific redshift and structure mass.

The average virial mass estimate is $M = 8 \times 10^{14} M_\odot$ in CCPC2. Included are 61 structures (plus 13 in CCPC1) that have masses $M > \times 10^{15} M_\odot$ at $z > 2$. The Millennium Simulation analysis by Chiang et al. (2013) suggest that for

the single most massive halo in a protocluster, none should have masses in excess of $\sim 10^{14}$ prior to $z = 2.3$. Likewise, Mortonson et al. (2011) predict that at $z \geq 2$, there should not be a single collapsed structure in a Λ CDM universe with $M > 6 \times 10^{14} h^{-1} M_{\odot}$. If any of these systems are in virial equilibrium, it could pose a serious challenge to the concordance cosmology.

Chiang et al. (2013) also computed the radii (in comoving units) for the main halos in protoclusters at redshifts $z = 2-5$ (their Figure 2). As noted earlier, our search algorithm was not designed to minimize the R_e of candidate systems, and so the radii listed in Table 2 may be an overestimate. Subsequently, these large radii could bias our anomalously large virial mass estimates. To test this, we limited our entire sample to those with R_e values less than the expected range for the most massive protoclusters (Chiang et al. 2013). There were 66 such sources, with an average radius of 6.5 cMpc and redshift of $z \sim 3.4$ (but spread between $2.00 < z < 6.5$). The average mass of these systems is $M_{\text{vir}} = 2.7 \times 10^{14} M_{\odot}$. At $z = 2.0$, this is approximately the mass expected for the largest main halos of the universe at this epoch. Observational mass estimates for structures at $z \sim 2$ from SZ/X-ray emission are consistent with this result (Gobat et al. 2011; Mantz et al. 2014). However, at larger redshifts, especially those at $z \geq 6$, the discrepancy is on the order of 10^2 from the predictions of Chiang et al. (2013). For at least this subsample, the radii do not appear to be the aberrant property in determining the mass.

It would seem that these discrepancies can be directly attributed to the inherent uncertainties in the velocity dispersions. Gobat et al. (2011) identified a (proto)cluster at $z \sim 2$ (identified here as CCPC-z20-002) and calculated its mass via X-ray emission to be $M_{\text{X-ray}} = (5-8) \times 10^{13}$. In a follow-up spectroscopic study of the galaxies in the system, they obtained a velocity dispersion of 1300 km s^{-1} (Gobat et al. 2013), which would suggest a much larger mass than inferred by the emission from the intracluster medium. Considerable work needs to be done to establish reliable methods for estimating protocluster masses at high redshift.

The overdensity and virial mass estimates (Equations (1) and (2), respectively) do not correlate with one another in the CCPC. The two estimators are ostensibly measuring two different properties, so taken at face value this lack of a link is not too surprising. The overdensity method is attempting to quantify the amount of mass that will collapse at $z = 0$, while the virial estimate is a representation of the current mass of the system in dynamical equilibrium. Despite the fact that even cluster progenitors of the same $z = 0$ mass can exist in a variety of evolutionary stages at $z \geq 2$ (Chiang et al. 2013; Muldrew et al. 2015), one would expect some relationship to exist. The conclusion that can be drawn from this comparison, in addition to the previous paragraphs discussing both estimators, is that these mass values are highly uncertain; no reliable mass indicator is available at present. While each may represent some aspect of the physical nature for these systems, we caution the reader not to rely on these values.

3.4. Protocluster Groups

CCPC-z23-001 and CCPC-z23-005: These overdensities were identified as two distinct systems by our search algorithm. After inspection, it was realized that these objects appear to be a single, extended system at $z = 2.33$. Figure 5 shows that the galaxy footprints overlap one another on the sky. However,

the LOS distribution of the galaxies appear to show a continuous distribution of sources of roughly 70 cMpc in length. The $N(z)$ plot exhibits a single, wide spike in galaxy counts (i.e., not a bimodal distribution). This spike was also identified using a combined catalog of primarily photometric with some spectroscopic galaxies by Salimbeni et al. (2009) as well as Kang & Im (2015). On the sky, the system is approximately 22×31 cMpc in R.A./decl., respectively. This is a large structure, but it is not unique. For example, Lee et al. (2014) found three overdensities at $z = 3.78$ within a $75 \times 75 \times 25 \text{ cMpc}^3$ volume. Dey et al. (2016) provided follow-up spectroscopic coverage to that extended system, which may be part of a filament stretching ~ 170 cMpc. More recently, Zheng et al. (2016) identified four protoclusters of LAEs with volumes in excess of $[15 \text{ cMpc}]^3$ around an overdensity discovered by Balestra et al. (2010) and listed in the CCPC1 as CCPC-z28-002.

The galaxy overdensities of CCPC-z23-001 and CCPC-z23-005 are relatively modest ($\delta_{\text{gal}} = 2, 0.7$ respectively). However, this can be partially attributed to the strong ‘‘field’’ counts from the neighboring system. CCPC-z23-001 is one of the richest systems in the combined CCPC ($N = 43$), while together they boast 66 member galaxies. These galaxies fill a volume of $4.6 \times 10^4 \text{ cMpc}^3$ and have a combined overdensity of $\delta_m = 0.52$. To compute the mass overdensity, we found the number density of galaxies within $z = 2.329 \pm 0.03$ and a field length $\Delta z = 0.15$. The volume is simply the rectangular region that encapsulates all galaxy members multiplied by the length of the box in comoving units. This volume and overdensity imply a mass of $3 \times 10^{15} M_{\odot}$ if the superstructure were to collapse. However, this large a volume is not expected to become one system by $z = 0$ in a Λ CDM universe. Within this volume, assuming a number density of clusters to be $7.8 \times 10^{-6} \text{ cMpc}^{-3}$ (Chiang et al. 2013), we would expect 0.3, $M \geq 10^{14} M_{\odot}$ clusters to be found. If they are each $\geq 10^{15} M_{\odot}$ systems, as their overdensity and volume implies, the cluster number density decreases to $1.6 \times 10^{-7} \text{ cMpc}^{-3}$. Within the estimated $5.893 \times 10^6 \text{ cMpc}^3$ volume in which we identify structure in the GOODS-S field ($2 < z < 5.7$), we would expect to find fewer than one $\geq 10^{15} M_{\odot}$ protocluster. That we find two systems in a volume 8×10^{-3} smaller is intriguing. The nature and fate of this system(s) is not yet understood.

CCPC-z65-004 and CCPC-z65-005: This object is similar to the previous example in that the algorithm detected these two objects as separate sources. However, the redshift at which this object is found makes it even more interesting. Along the LOS, these two systems essentially overlap (right panel of Figure 6), are much shorter than the $\Delta z = 40$ cMpc search length (< 20 cMpc), and are thin. On the sky, their respective search centers are merely 40 cMpc offset from one another. This could mean that this is a single, very massive protocluster ($\delta_{\text{gal}} \sim 4$, $M > 2 \times 10^{15} M_{\odot}$) with a geometric center in between the two coordinates listed. Of the two other $z \sim 6$ protoclusters known (Utsumi et al. 2010; Toshikawa et al. 2014), this combined system has an overdensity mass estimate at least 5 times larger.

Regardless of whether CCPC-z65-004 and CCPC-z65-005 are a single system or two separate, high-mass protoclusters, this detection represents the highest redshift association of galaxies that has been spectroscopically confirmed to the best of our knowledge. Trenti et al. (2012) and Ishigaki et al. (2015) have both identified protocluster candidates at $z \sim 8$ based on strong overdensities of Y -dropout galaxies, but these have yet to be spectroscopically confirmed.

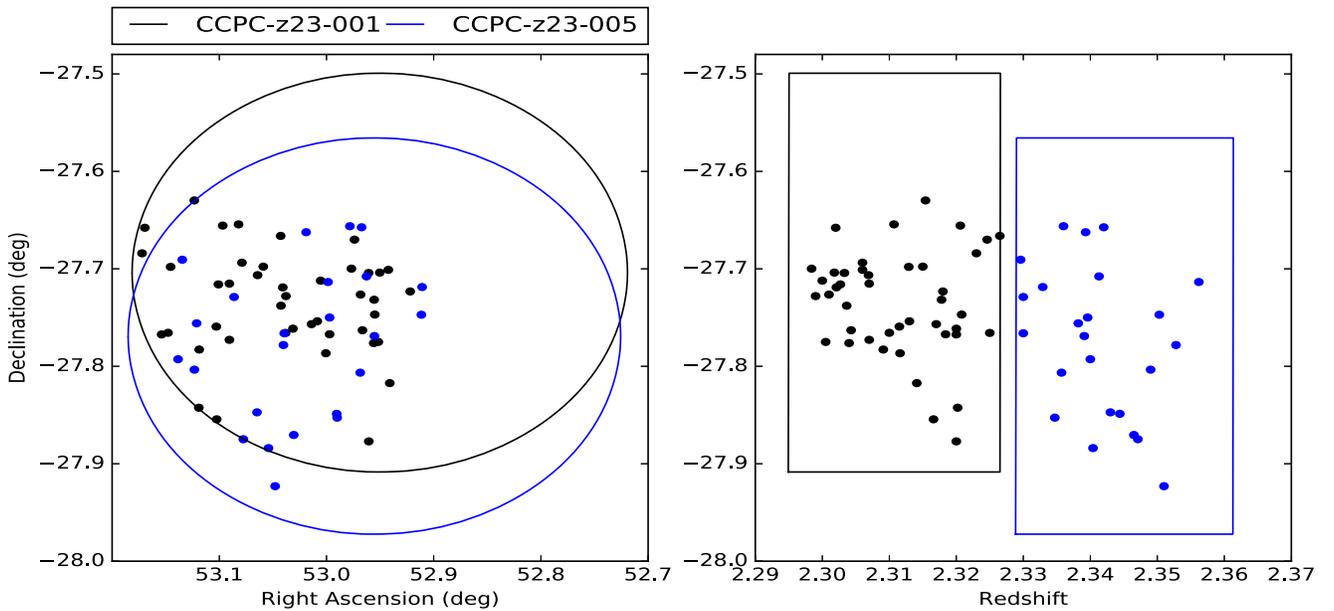


Figure 5. At $z \sim 2.33$, our search algorithm selected the two galaxy associations (CCPC-z23-001 and CCPC-z23-005) as being strong protocluster candidates. Left: sky plot showing the overlap of the two structures on the sky, with the ellipses illustrating the boundary of our search volume ($R = 20$ cMpc). Right: LOS distribution of galaxies in the two candidates, which appear as a single, unbroken distribution. The boxes are the boundaries of Δz corresponding to ± 20 cMpc. Further analysis shows that these objects appear as one continuous structure of volume 4.6×10^4 cMpc³. The comoving volume is calculated by assuming a box with a length corresponding to the minimum and maximum redshifts, and width bounded by the galaxy positions in the sky plot (left panel). There is also no break in the $N(z)$ distribution that would indicate two separate structures. A mass estimate based on its overdensity is $3 \times 10^{15} M_{\odot}$, which would be one of the largest structures known.

CCPC-z34-005, CCPC-z34-006, CCPC-z35-002, and CCPC-z35-003 Complex: These overdensities, although individually unremarkable, are part of a linked superstructure. These four separate systems almost touch on the sky in Figure 7. CCPC-z34-005, CCPC-z34-006, and CCPC-z35-003 all exist along the same LOS in a chain of ≤ 120 cMpc in length. The center of CCPC-z35-002 is separated from the chain by approximately 40 cMpc. This complex may be a proto-supercluster in the process of assembly.

There are a number of other systems in CCPC2 that may be associated with one another. CCPC-z27-006 and CCPC-z27-010 appear to be nearly touching along the LOS and overlap one another on the sky. They appear similar to CCPC-z23-001 and CCPC-z23-005 (as seen in Figure 5), but are not nearly as rich. CCPC-z28-015 and CCPC-z28-016, and CCPC-z45-001-CCPC-z45-002 are more similar in nature to the distribution of galaxies in Figure 6 with similar redshifts but separated on the sky by roughly the search radius of the algorithm. However, the associations are not as strong as that of CCPC-z65-004-CCPC-z65-005. Nevertheless, it appears that it is not uncommon for protocluster candidates to reside in very large-scale associations.

We emphasize again that the volumes estimated in this manner are highly uncertain. Candidate member galaxies on the outskirts of a distribution, whether projected on the sky or along the LOS, can significantly enhance the volume implied. These volumetric tracers may not be bound to the structure at all, or may be sufficiently separated that they will not reside within the cluster’s halo at $z = 0$.

4. SUMMARY

We have extended the CCPC to redshifts $2 < z < 6.6$ in CCPC2, adding 173 protocluster candidates to the 43 in CCPC1. In CCPC2, we identified galaxy overdensities ($\delta_{\text{gal}} > 0.25$) of four or more galaxies within a search radius of $R = 20$ cMpc and within a Δz of ± 20 cMpc. In Table 3, all candidate protoclusters

are listed. The 36 systems that have the largest collapse probabilities ($\delta_{\text{gal}} > 6$) are found in Table 1. The median number of galaxy members is six in CCPC2, and nine candidates have $N > 23$ galaxies. Above a redshift of $z > 4$, we have identified 40 structures. Prior to this work, fewer than 10 had been identified. At the time of writing, this list includes the most distant spectroscopically confirmed protocluster known (CCPC-z65-005). The combined CCPC is the largest known list of high-redshift, spectroscopic protoclusters to date.

Following the examples of Steidel et al. (1998) and Venemans et al. (2002), we estimate the strength of these protoclusters by computing their overdensities with respect to field counts of galaxies at similar redshifts. These structures contain a median overdensity of $\delta_{\text{gal}} \sim 2.9$, slightly larger than the median value found in CCPC1 and typical of the overdensities found in the literature (summarized in Table 5 of Chiang et al. 2013), as well as expectations from simulations. These overdensities and their visual counterparts ($N(z)$ plots for each CCPC member in the Appendix) are two pieces of evidence suggesting that these galaxy associations are indeed structures and not merely coincident on the sky. We emphasize that these overdensities are conservative estimates (Figure 1).

In Section 3, we compare the spatial distribution and surface densities of CCPC systems to field galaxies. Using the list of $\geq 40,000$ galaxies from which we identify overdensities, we built the mean surface density of “all galaxies” as a proxy for the field in Figure 3. Compared to the mean distribution of CCPC galaxies out to $R = 20$ cMpc, there are significantly fewer galaxies at all radii in the field. A KS test of these two distributions shows a statistically significant difference (KS = 0.60). Figure 4 illustrate a similar trend in the total number of galaxies as a function of the distance from search center. There are fewer field galaxies, at all redshifts and radii, than in CCPC systems. We estimate the expected number of protoclusters that could be found within the CANDELS

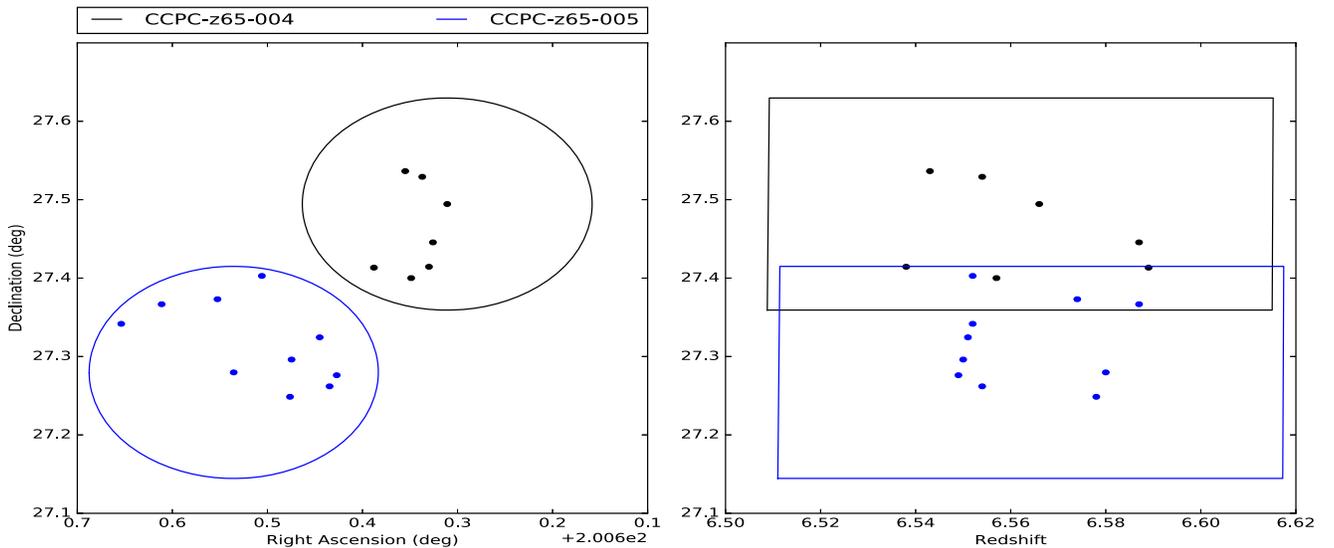


Figure 6. CCPC-z65-004 and CCPC-z65-005 are two systems separated by ~ 40 cMpc on the sky (left panel), with their search volumes (ellipses) nearly touching. Indeed, a search center located at the midpoint (R.A.: 201° , decl.: $+27^\circ 4'$) would capture the majority of galaxies in the two distributions. In the right panel, the Δz plots corresponding to a length of 40 cMpc, almost perfectly overlap, showing the galaxies along the LOS. These could be two distinct associations, or one system with a geometric center in between the two groups. The mass of this system is estimated to be $M > 2 \times 10^{15} M_\odot$. These two (sub)protoclusters are the highest spectroscopically identified protoclusters at the time of writing.

GOODS-S pencil beam survey volume based on the number density of clusters in Millennium (Chiang et al. 2013). In CCPC1 and CCPC2, we found 27 candidate structures, whereas 46 clusters of $M \geq 10^{14} M_\odot$ were expected in a similar volume of the Λ CDM N -body simulation. This suggests that we are not over-identifying structure using our method, and may be recovering only the most significant overdensities. For systems of Virgo-mass and larger ($M \geq 3 \times 10^{14} M_\odot$), 14 are expected in the GOODS-S volume.

We applied an analytic toy model by Sheth (2001) to estimate the expected number of galaxy groups stochastically produced in a smooth density field. We then compared this expectation value to the number of FoF groups identified in CANDELS GOODS-S survey and found an excess of FoF groups compared to what would be expected stochastically from Poissonian fluctuations and Monte Carlo simulations. This test furthers the notion that these are physical structures and not mere chance overdensities.

These tests were to ensure that the structures identified with the heterogeneous data available were legitimate structures. We have shown that: (1) these objects exist as galaxy overdensities (median $\delta_{\text{gal}} \sim 2.9$), (2) have $4\times$ larger number densities than LAEs (Section 2), (3) differ spatially from field spectroscopic catalogs in Section 3, (4) are not the result of Poissonian fluctuations in a smooth density field (Sheth 2001), and (5) we do not over-identify structures in the deepest survey used (CANDELS GOODS-S). It is statistically unlikely that these tests, when taken together, would fail to distinguish the field from the candidates if a significant number of these systems are not actual structures.

For each CCPC member, we estimated the mass of the structure using two distinct methods. The first technique uses a linear bias parameter b to transform the galaxy overdensity δ_{gal} into a mass overdensity δ_m . If one assumes that the volume (V) traced by galaxies will collapse to a cluster, the mass of the system at $z = 0$ can be approximated by the product of the volume and mass density (Steidel et al. 1998). The average overdensity mass estimate for the CCPC2 is $1.8 \times 10^{14} M_\odot$. We also computed the virial mass for each system based on its

effective radius R_e and velocity dispersion σ , making the crude assumption that these systems are in virial equilibrium. The mean “virial” mass is $M = 8.4 \times 10^{14} M_\odot$ calculated for all CCPC2 members. With the combined CCPC1 and CCPC2 catalog, a total of 74 structures have mass estimates of $M \geq 10^{15} M_\odot$. Such large masses are not expected at redshifts $z > 2.3$ in N -body simulations (Chiang et al. 2013). There is little agreement between the volumetric and virial mass estimators which emphasizes their uncertainty and the dubiousness of the necessary assumptions (e.g., virial equilibrium).

There does appear to be a discrepancy between the observed and predicted velocity dispersions (Chiang et al. 2013; Contini et al. 2015). The median value is $\sigma \geq 650 \text{ km s}^{-1}$ in both CCPC1 and CCPC2, with some systems having dispersions as large as 900 km s^{-1} , here and in Venemans et al. (2007). This is more than double the expected value at the relevant redshifts for the most massive protoclusters ($400 \pm 60 \text{ km s}^{-1}$ in Chiang et al. 2013). This is important because no assumption of equilibrium is made. However, it may not be possible to make an apples-to-apples comparison, as objects near the edge of the protocluster have large ($\sim 1000 \text{ km s}^{-1}$) simulated velocities (Contini et al. 2015). By reducing the LOS window size to ± 10 cMpc generally removes this observed excess to expected values. We note that it is difficult to map “observations” of simulations to real data. The observed excess of velocity dispersion in the CCPC is most likely not a physically relevant result.

The CCPC2 also has three groups of protoclusters (Figures 5–7) that may be primordial superclusters based on their small (≤ 100 cMpc) separations. CCPC-z23-001 and CCPC-z23-005 appear to be physically connected in a chain of length ~ 70 cMpc. Its nature is not understood currently, as it appears to be a single, extended structure of volume $\geq 4 \times 10^4 \text{ cMpc}^3$ and mass $\geq 10^{15} M_\odot$. CCPC-z65-004 and CCPC-z65-005 are the two highest-redshift spectroscopically confirmed protoclusters known to date, and are separated by a mere 40 cMpc. They may be a single, massive structure with a geometric center offset from their galaxy distributions, or two disparate protoclusters very near to one another. CCPC-z34-005, CCPC-z34-006, CCPC-z35-002, and

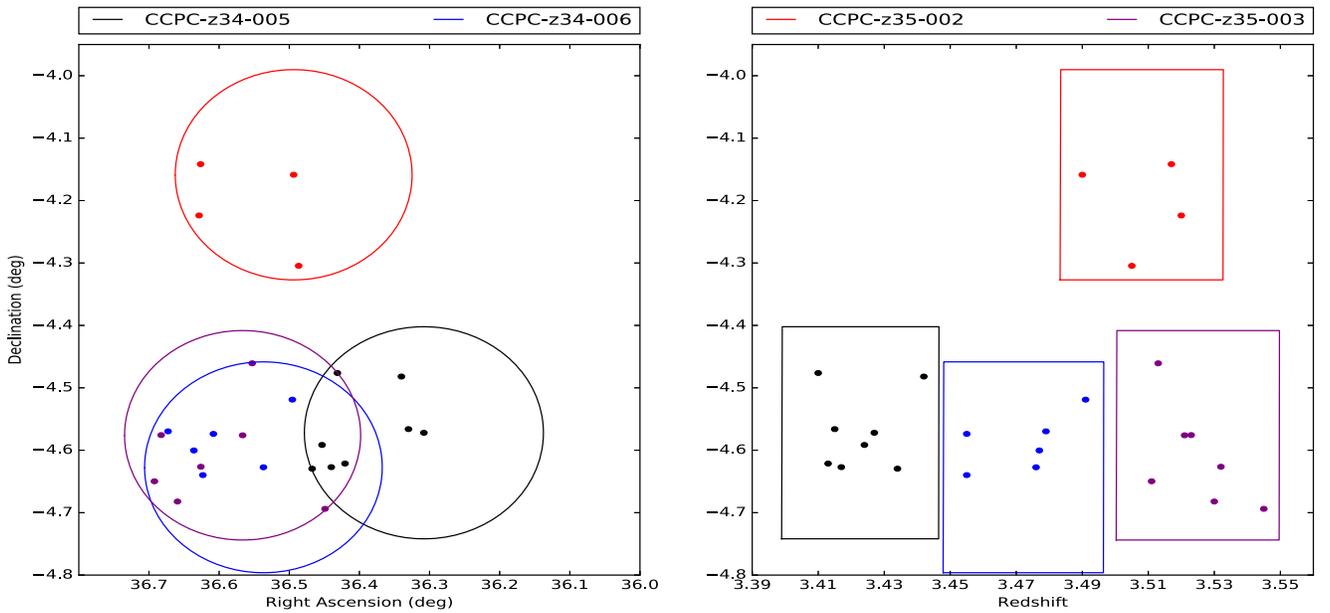


Figure 7. A group of four protocluster candidates at $z = 3.5$ that are proximate in space. The left panel is the sky association of the candidates, while the right panel shows their galaxy distribution along the LOS. The ellipses/boxes show the search volume boundaries ($R_{\text{search}} = 20$ cMpc, $\Delta z \pm 20$ cMpc). Three of these (CCPC-z34-005, CCPC-z34-006, CCPC-z35-003) exist in a chain along the LOS stretching ≤ 120 cMpc. This may become a supercluster-sized structure at $z = 0$.

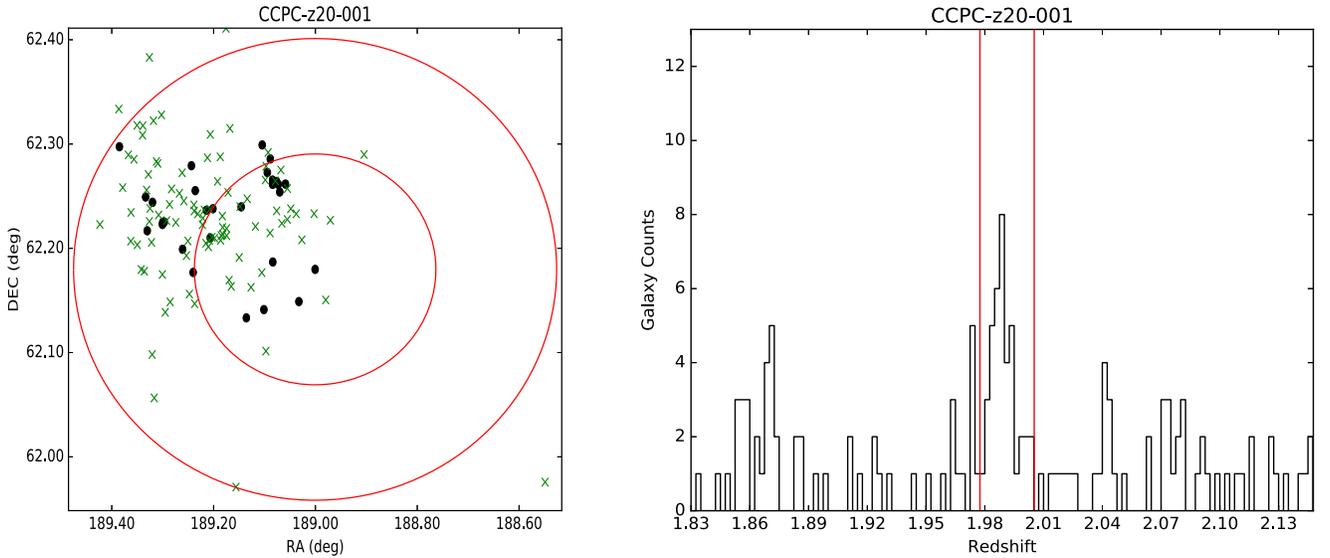


Figure 8. Left: the outer and inner red circles represents 20 and 10 cMpc from the center of the protocluster, respectively. Black points are the protocluster galaxies and green \times symbols indicate “field” sources within $\Delta z \pm 0.15$ of the redshift of the structure. This serves as a proxy of the FOV of sources in the combined surveys. Right: number of galaxies plotted as a function of redshift ($N(z)$ distribution). The red vertical lines represent ± 20 comoving Mpc (the search criteria) from the mean redshift of the system. The calculation of the overdensity δ_{gal} is restricted to $z \pm \sigma$ and is typically a factor of a few smaller.

(The complete figure set (173 images) is available.)

CCPC-z35-003 form a protocluster complex, with three existing in a chain along the LOS.

In total, we have identified 173 protocluster candidates that appear to be genuine, spectroscopically confirmed, physical associations of galaxies at high redshift. Some of these systems reside in close proximity, as if part of (proto) super-clusters. Candidate protoclusters of high confidence are found up to $z \approx 6.5$. There appears to be a rich amount of structure still to be revealed in the high-redshift universe.

We would like to thank the anonymous referee for comments that improved this work. JRF is grateful to Idit Zehavi for

helpful discussion on Poissonian expectations for a smooth density field, and Ji Hoon Kim for providing Subaru NB filter curves. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013). This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

APPENDIX

Table 4
CCPC: Member Redshift Reference List

| Candidate Name | Redshift References |
|----------------|---|
| CCPC-z20-001 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 |
| CCPC-z20-002 | 15, 16 |
| CCPC-z20-003 | 17, 18 |
| CCPC-z20-004 | 19, 20, 21 |
| CCPC-z20-005 | 22, 23, 24, 25, 26 |
| CCPC-z20-006 | 27, 13, 28, 3 |
| CCPC-z20-007 | 29, 30 |
| CCPC-z20-008 | 31, 3, 4, 32 |
| CCPC-z20-009 | 33, 34, 35, 36 |
| CCPC-z21-001 | 37, 27, 38, 39 |

Note.

Please refer to the machine-readable table for a detailed description of each column.

(This table is available in its entirety in machine-readable form.)

Table 5
CCPC: Mass Estimates

| Candidate Name | R_e (Mpc) | σ (km s $^{-1}$) | Virial Mass Estimate ($10^{14} M_\odot$) | δ_m ($b = 3$) | Overdensity Volume (cMpc 3) | Overdensity Mass Estimate ($10^{14} M_\odot$) |
|----------------|-------------|--------------------------|--|------------------------|---------------------------------|---|
| CCPC-z20-001 | 4.2 | 563 | 7.7 | 1.69 | 2945 | 3.5 |
| CCPC-z20-002 | 4.5 | 749 | 14.7 | 3.13 | 5361 | 9.8 |
| CCPC-z20-003 | 2.3 | 367 | 1.8 | 6.48 | 125 | 0.4 |
| CCPC-z20-004 | 4.9 | 619 | 10.8 | 0.74 | 152 | 0.1 |
| CCPC-z20-005 | 5.0 | 773 | 17.4 | 0.33 | 11341 | 6.5 |
| CCPC-z20-006 | 3.4 | 827 | 13.3 | 0.16 | 1249 | 0.6 |
| CCPC-z20-007 | 2.6 | 565 | 4.8 | 0.67 | 40 | <0.1 |
| CCPC-z20-008 | 4.3 | 671 | 11.3 | 0.43 | 2126 | 1.4 |
| CCPC-z20-009 | 4.3 | 446 | 4.9 | 4.38 | 589 | 1.4 |
| CCPC-z21-001 | 3.4 | 689 | 9.3 | 1.39 | 690 | 0.7 |

Note.

Please refer to the machine-readable table for a detailed description of each column.

(This table is available in its entirety in machine-readable form.)

Table 6
CCPC II: Objects of Interest (OOIs)

| Candidate Name | R.A. (degree) | Decl. (degree) | Redshift (z_{avg}) | N | $N_{R \leq 10}$ cMpc | R_e (Mpc) | Reference |
|----------------|---------------|----------------|------------------------|-----|----------------------|-------------|-----------|
| OOI-z22-001 | 30.199 | 1.749 | 2.2 | 11 | 10 | 2.65 | 1 |
| OOI-z22-002 | 30.569 | 1.488 | 2.2 | 14 | 4 | 4.45 | 1 |
| OOI-z22-003 | 30.581 | 1.928 | 2.2 | 10 | 2 | 5.1 | 1 |
| OOI-z22-004 | 150.852 | 0.181 | 2.2 | 17 | 11 | 2.4 | 1 |
| OOI-z22-005 | 151.318 | 0.697 | 2.2 | 19 | 9 | 3.2 | 1 |
| OOI-z22-006 | 197.999 | 42.715 | 2.2 | 4 | 3 | 1 | 1 |
| OOI-z23-001 | 325.525 | -44.580 | 2.38 | 10 | 4 | 3.15 | 2 |
| OOI-z23-002 | 325.843 | -44.218 | 2.38 | 10 | 2 | 4.1 | 2 |
| OOI-z23-001 | 325.525 | -44.580 | 2.38 | 10 | 4 | 3.15 | 2 |
| OOI-z23-002 | 325.843 | -44.218 | 2.38 | 10 | 2 | 4.1 | 2 |
| OOI-z57-001 | 149.654 | 1.539 | 5.7 | 6 | 1 | 1.8 | 3 |
| OOI-z57-002 | 149.746 | 2.752 | 5.7 | 7 | 2 | 1.75 | 3 |
| OOI-z57-003 | 150.398 | 2.773 | 5.7 | 8 | 4 | 1.35 | 3 |
| OOI-z57-004 | 230.905 | -0.144 | 5.7 | 5 | 3 | 1.45 | 4 |
| OOI-z57-005 | 150.168 | 2.318 | 5.7 | 5 | 3 | 1.4 | 5 |
| OOI-z65-001 | 314.069 | -4.630 | 6.50 | 4 | 4 | 0.75 | 6 |
| OOI-z65-002 | 34.4896 | -5.146 | 6.57 | 4 | 1 | 1.85 | 7 |

Note. Candidate systems found within NED as spectroscopic sources. Further inspection revealed that these galaxies are mainly from NB imaging and do not have spectroscopic information, but fulfil all other CCPC criteria. The last two systems are spectroscopic sources, but did not have any field galaxies with which to compute an overdensity. The references of their estimated redshifts are included.

References. (1) Matsuda et al. (2011), (2) Francis et al. (2004), (3) Murayama et al. (2007), (4) Martin et al. (2008), (5) Mallery et al. (2012), (6) de Diego et al. (2013), (7) Jiang et al. (2013).

REFERENCES

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Bahcall, N. A., & Fan, X. 1998, *ApJ*, **504**, 1
- Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, **512**, A12
- Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, **100**, 32
- Casey, C. M., Cooray, A., Capak, P., et al. 2015, *ApJL*, **808**, L33
- Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2013, *ApJ*, **779**, 127
- Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, *ApJL*, **782**, L3
- Clements, D. L., Braglia, F. G., Hyde, A. K., et al. 2014, *MNRAS*, **439**, 1193
- Contini, E., De Lucia, G., Hatch, N., Borgani, S., & Kang, X. 2015, arXiv:1512.01561
- Cooke, E. A., Hatch, N. A., Muldrew, S. I., Rigby, E. E., & Kurk, J. D. 2014, *MNRAS*, **440**, 3262
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, **267**, 465
- de Diego, J. A., De Leo, M. A., Cepa, J., et al. 2013, *AJ*, **146**, 96
- Dey, A., Lee, K.-S., Reddy, N., et al. 2016, arXiv:1604.08627
- Diener, C., Lilly, S. J., Knobel, C., et al. 2013, *ApJ*, **765**, 109
- Francis, P. J., Palunas, P., Teplitz, H. I., Williger, G. M., & Woodgate, B. E. 2004, *ApJ*, **614**, 75
- Franck, J. R., & McGaugh, S. S. 2016, *ApJ*, **817**, 158
- Fry, J. N. 1996, *ApJL*, **461**, L65
- Fynbo, J. P. U., Ledoux, C., Möller, P., Thomsen, B., & Burud, I. 2003, *A&A*, **407**, 147
- Galametz, A., Stern, D., De Breuck, C., et al. 2012, *ApJ*, **749**, 169
- Galametz, A., Stern, D., Pentericci, L., et al. 2013, *A&A*, **559**, A2
- Gawiser, E., Francke, H., Lai, K., et al. 2007, *ApJ*, **671**, 278
- Gobat, R., Daddi, E., Onodera, M., et al. 2011, *A&A*, **526**, A133
- Gobat, R., Strazzullo, V., Daddi, E., et al. 2013, *ApJ*, **776**, 9
- Granato, G. L., Ragone-Figueroa, C., Domínguez-Tenreiro, R., et al. 2015, *MNRAS*, **450**, 1320
- Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, *ApJ*, **667**, 79
- Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, **413**, 101
- Hatch, N. A., Kurk, J. D., Pentericci, L., et al. 2011, *MNRAS*, **415**, 2993
- Huchra, J. P., & Geller, M. J. 1982, *ApJ*, **257**, 423
- Ishigaki, M., Ouchi, M., & Harikane, Y. 2015, arXiv:1509.01751
- Jiang, L., Egami, E., Mechtley, M., et al. 2013, *ApJ*, **772**, 99
- Kang, E., & Im, M. 2009, *ApJL*, **691**, L33
- Kang, E., & Im, M. 2015, *JKAS*, **48**, 21
- Keel, W. C., Cohen, S. H., Windhorst, R. A., & Waddington, I. 1999, *AJ*, **118**, 2547
- Koyama, Y., Kodama, T., Tadaki, K.-I., et al. 2013, *MNRAS*, **428**, 1551
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, *A&A*, **439**, 845
- Lee, K.-S., Dey, A., Hong, S., et al. 2014, *ApJ*, **796**, 126
- Mallery, R. P., Mobasher, B., Capak, P., et al. 2012, *ApJ*, **760**, 128
- Mantz, A. B., Abdulla, Z., Carlstrom, J. E., et al. 2014, arXiv:1401.2087
- Martin, C. L., Sawicki, M., Dressler, A., & McCarthy, P. 2008, *ApJ*, **679**, 942
- Matsuda, Y., Smail, I., Geach, J. E., et al. 2011, *MNRAS*, **416**, 2041
- Matsuda, Y., Yamada, T., Hayashino, T., et al. 2005, *ApJL*, **634**, L125
- Mortonson, M. J., Hu, W., & Huterer, D. 2011, *PhRvD*, **83**, 023015
- Muldrew, S. I., Hatch, N. A., & Cooke, E. A. 2015, *MNRAS*, **452**, 2528
- Murayama, T., Taniguchi, Y., Scoville, N. Z., et al. 2007, *ApJS*, **172**, 523
- Noll, S., Mehlert, D., Appenzeller, I., et al. 2004, *A&A*, **418**, 885
- Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2003, *ApJ*, **582**, 60
- Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004, *ApJ*, **611**, 685
- Pentericci, L., Roettgering, H. J. A., Miley, G. K., Carilli, C. L., & McCarthy, P. 1997, *A&A*, **326**, 580
- Press, W. H., & Schechter, P. 1974, *ApJ*, **187**, 425
- Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, *ApJ*, **653**, 1004
- Rigby, E. E., Hatch, N. A., Röttgering, H. J. A., et al. 2014, *MNRAS*, **437**, 1882
- Salimbeni, S., Castellano, M., Pentericci, L., et al. 2009, *A&A*, **501**, 865
- Sheth, R. K. 2001, *NYASA*, **927**, 1
- Shimakawa, R., Kodama, T., Tadaki, K.-I., et al. 2014, *MNRAS*, **441**, L1
- Shimasaku, K., Ouchi, M., Okamura, S., et al. 2003, *ApJL*, **586**, L111
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, *ApJ*, **492**, 428
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, *ApJ*, **592**, 728
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2005, *ApJ*, **626**, 44
- Topping, M. W., Shapley, A. E., & Steidel, C. C. 2016, arXiv:1605.02079
- Toshikawa, J., Kashikawa, N., Overzier, R., et al. 2014, *ApJ*, **792**, 15
- Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2012, *ApJ*, **746**, 55
- Trenti, M., & Stiavelli, M. 2008, *ApJ*, **676**, 767
- Utsumi, Y., Goto, T., Kashikawa, N., et al. 2010, *ApJ*, **721**, 1680
- Venemans, B. P. 2005, PhD thesis, Leiden Observatory, Leiden Univ.
- Venemans, B. P., Kurk, J. D., Miley, G. K., et al. 2002, *ApJL*, **569**, L11
- Venemans, B. P., Röttgering, H. J. A., Miley, G. K., et al. 2007, *A&A*, **461**, 823
- Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al. 2009, *ApJ*, **692**, 1060
- Wang, T., Elbaz, D., Daddi, E., et al. 2016, arXiv:1604.07404
- Wright, E. L. 2006, *PASP*, **118**, 1711
- Wylezalek, D., Galametz, A., Stern, D., et al. 2013, *ApJ*, **769**, 79
- Wylezalek, D., Vernet, J., De Breuck, C., et al. 2014, *ApJ*, **786**, 17
- Yuan, T., Nanayakkara, T., Kacprzak, G. G., et al. 2014, *ApJL*, **795**, L20
- Zheng, Z.-Y., Malhotra, S., Rhoads, J. E., et al. 2016, arXiv:1606.07073