

LETTER TO THE EDITOR

# A lensed protocluster candidate at $z = 7.66$ identified in JWST observations of the galaxy cluster SMACS0723–7327

N. Laporte<sup>1,2</sup> , A. Zitrin<sup>3</sup>, H. Dole<sup>4</sup>, G. Roberts-Borsani<sup>5</sup>, L. J. Furtak<sup>3</sup>, and C. Witten<sup>6</sup>

<sup>1</sup> Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK  
e-mail: n1408@cam.ac.uk

<sup>2</sup> Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK

<sup>3</sup> Physics Department, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 8410501, Israel

<sup>4</sup> Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, 91405 Orsay, France

<sup>5</sup> Department of Physics and Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA

<sup>6</sup> Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Received 9 August 2022 / Accepted 3 October 2022

## ABSTRACT

**Context.** According to the current paradigm of galaxy formation, the first galaxies likely formed within large dark matter haloes. The fragmentation of these massive haloes led to the formation of galaxy protoclusters, which are usually composed of one to a few bright objects, surrounded by numerous fainter (and less massive) galaxies. These early structures could have played a major role in reionising the neutral hydrogen within the first billion years of the Universe, especially if their number density is significant.

**Aims.** Taking advantage of the unprecedented sensitivity reached by the *James Webb* Space Telescope (JWST), galaxy protoclusters can now be identified and studied in increasing numbers beyond  $z \geq 6$ . Characterising their contribution to the UV photon budget could supply new insights into the reionisation process.

**Methods.** We analysed the first JWST dataset behind SMACS0723–7327 to search for protoclusters at  $z \geq 6$ , combining the available spectroscopic and photometric data. We then compared our findings with semi-analytical models and simulations.

**Results.** In addition to two bright galaxies ( $\leq 26.5$  AB in  $F277W$ ), separated by  $\sim 11''$  and spectroscopically confirmed at  $z_{\text{spec}} = 7.66$ , we identify six additional galaxies with similar colours within a  $\theta \sim 20''$  radius (corresponding to  $R \sim 60$ – $90$  kpc in the source plane). Using several methods, we estimate the mass of the dark matter halo of this protocluster as  $\sim 3.3 \times 10^{11} M_{\odot}$ , accounting for magnification, consistent with various predictions. The physical properties of all protocluster members are also in excellent agreement with what has been previously found at lower redshifts: star formation main sequence and protocluster size. This detection adds to just a few protoclusters currently known in the first billion years of the universe. These  $z \geq 7$  galaxy protoclusters may play an important role in cosmic reionisation.

**Key words.** galaxies: formation – galaxies: distances and redshifts – galaxies: groups: general

## 1. Introduction

Understanding the formation and evolution of the first population of galaxies a few million years after the Big Bang is one of the most active topics of current extragalactic astronomy. For decades, instruments have been built to push our observational limits ever further. The current most distant and detailed picture of the Universe, the cosmic microwave background (CMB), was obtained by the *Planck* mission (Planck Collaboration VI 2020). It shows that 380 000 years after the Big Bang, the matter density in the Universe was already inhomogeneously distributed, suggesting that small amplitude density fluctuations were taking place in the early phase of the Universe. These fluctuations grew and eventually the denser regions collapsed to form the first bound objects (Bromm & Yoshida 2011). Moreover, the first dark matter haloes underwent a process of fragmentation, suggesting that the most massive galaxies may have formed in overdense regions known as protoclusters (Genel et al. 2010). Recent  $N$ -body simulations and semi-analytic models demonstrate that the first protoclusters may have contributed up to  $\sim 50\%$  of the Cosmic Star Formation Rate Density at  $z \sim 10$

(Chiang et al. 2017). Therefore, determining the number density of  $z \geq 6$  protoclusters could supply new insights into the reionisation process.

A natural method of identifying protoclusters at  $z \geq 6$  is to search for overdensities of photometrically selected dropout galaxies at similar high redshifts. For example, using *Hubble* Frontier Fields data (Lotz et al. 2017), Zheng et al. (2014) identified a likely protocluster of several galaxies at  $z \sim 8$ , one of which was later targeted and found to show UV and far-IR emission lines placing it at  $z = 8.38$  (Laporte et al. 2017). Searching for galaxies with Lyman- $\alpha$  emission at similar redshifts is another successful route to identifying early protoclusters. Since the neutral hydrogen surrounding galaxies at  $z \geq 6$  makes the detectability of this line very low at high redshift (De Barros et al. 2017), a Ly $\alpha$  detection from a galaxy at  $z \geq 6$  suggests that the galaxy is found in a rather large ionised bubble, a result of either high-ionising power of the galaxy itself or the cumulative contribution of several objects at the same redshift (Roberts-Borsani et al. 2022b). Using the *Hubble* Space Telescope (HST) and deep spectroscopic follow-up campaigns, several groups have identified large ionised bubbles up

to  $z \sim 8.68$  (e.g., Castellano et al. 2016, 2022; Tilvi et al. 2020; Leonova et al. 2022; Larson et al. 2022) with the detection of Ly- $\alpha$  in bright galaxies. However, the sensitivity of HST is not sufficient to search for much fainter objects in the local environment of these bright galaxies.

The successful launch of the JWST on December 25, 2021, from Europe's Spaceport in French Guyana has opened a new window for the study of protoclusters. Its unprecedented sensitivity will allow the community to identify and spectroscopically confirm not only bright galaxies, but also fainter galaxies at similar redshifts. On July 12, 2022, the first images and spectra obtained by *Webb* were released, and preliminary analyses show that galaxies had already formed at  $z \geq 12$  (e.g., Finkelstein et al. 2022; Naidu et al. 2022b; Donnan et al. 2022) reinforcing the idea that a protocluster of galaxies may already be in place at  $z \geq 10$ .

In this paper we analyse the first dataset released by the JWST behind the lensing cluster SMACS0723–7327 to search for protoclusters at  $z \geq 6$ . In Sect. 2 we describe our method, which led to the identification of a protocluster at  $z = 7.66$  using NIRSpec, NIRISS, and NIRCам data. Then we determine the physical properties of protocluster members as well as the global properties of the protocluster (Sect. 3). Finally, in Sect. 4 we discuss the implication of our findings on the reionisation process.

Throughout this paper we assume a standard  $\Lambda$ CDM cosmology with parameters from Planck Collaboration VI (2020). All magnitudes are in the AB system (Oke & Gunn 1983).

## 2. Search for protocluster behind SMACS0723

The first JWST dataset included NIRCам images in *F090W*, *F150W*, *F200W*, *F277W*, *F356W*, and *F444W* filters; MIRI images in *F770W*, *F1000W*, *F1500W*, and *F1800W* filters; NIRSpec spectra in *F170LP* and *F190LP*; and NIRISS spectra obtained with *F115W* and *F200W* filters. We first looked at the NIRSpec spectra whose integration time is 2.45 h in each filter. We used the publicly available level 2 data, and visually inspected the 1D spectra using Jdaviz<sup>1</sup>. The spectroscopic redshift was obtained by fitting a list of nebular lines to the brightest detected lines. Two objects among the 35 observed have a similar redshift of  $z = 7.665$  and  $z = 7.659$  (hereafter SMACS0723\_PC\_6 and SMACS0723\_PC\_7) and show several UV lines, such as [OIII] $\lambda\lambda 4959, 5007$  and [OII] $\lambda\lambda 3727, 3729$  (Fig. 1). Our redshift measurements are consistent with the values previously measured by other groups (e.g., Curti et al. 2022; Carnall et al. 2022; Atek et al. 2022; Katz et al. 2022; Trussler et al. 2022).

Theoretical studies demonstrate that at  $z \sim 7$  the size of a protocluster is below 10 comoving Mpc, corresponding to  $\sim 1$  physical Mpc (Chiang et al. 2017). However, the protocluster core, where the large majority of galaxies are expected, is much smaller. For example, Capak et al. (2011) identified a protocluster at  $z = 5.3$ , with a protocluster core size of 0.14 physical Mpc. While it appears complicated to entirely cover a protocluster at such high redshift with the first JWST dataset, one can easily search for a protocluster core. In the following we describe how we searched for galaxies with similar colours in a  $40'' \times 40''$  box centred on the two galaxies identified at  $z = 7.66$ . Moreover, to increase the number of constraints on the SED, we combined JWST/NIRCам and MIRI data with HST/ACS and WFC3 data.

To define the colour–colour criteria needed to select galaxies at  $z \sim 7.66$ , we used BAGPIPES (Carnall et al. 2018)

<sup>1</sup> <https://doi.org/10.5281/zenodo.6824713>

to simulate galaxies whose parameters range from 0.0 to 10 for the redshift,  $10^6$ – $10^{10} M_\odot$  for the stellar mass, and  $-4.0$  to  $-0.5$  for the ionisation parameter. The dust attenuation follows Calzetti et al. (2000), and the dust reddening ranges from  $A_V = 0.0$  to 2.0 mag. Intergalactic medium (IGM) attenuation is included and modelled from Inoue et al. (2014). We assume four different star formation histories (SFHs), constant, burst, delayed, and a combination of a constant and a young burst, and simulate a total of 40 000 galaxies. To account for the depth difference between JWST and HST images, we only defined criteria based on NIRCам filters. We obtain the following criteria to select objects at  $z \geq 7$  (Fig. 2):

$$F090W - F150W > 1.1, \quad (1)$$

$$F150W - F200W < 0.3. \quad (2)$$

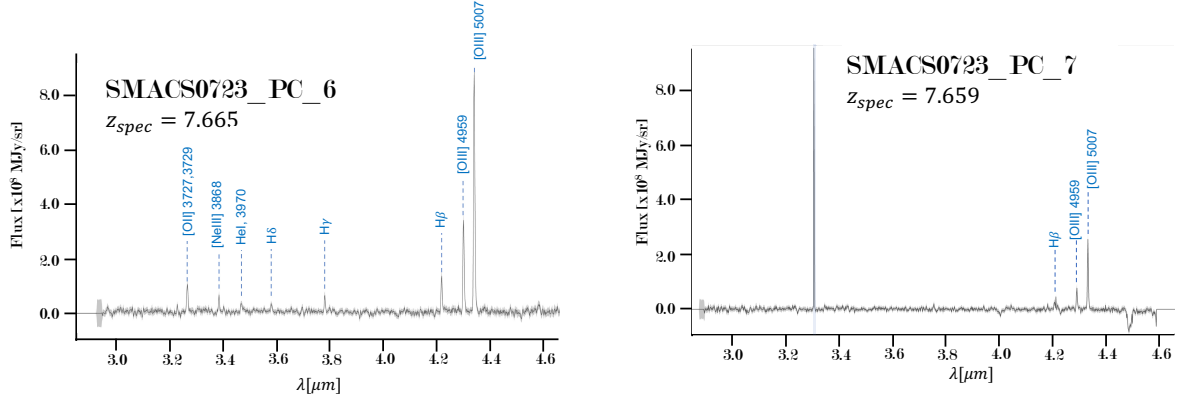
We combined these criteria with non-detection criteria ( $< 2\sigma$ ) in HST/ACS filters *F435W*, *F606W*, and *F814W* and detection criteria ( $> 5\sigma$ ) in *F150W*, *F160W*, and *F200W*. Colours are measured on point spread function (psf)-matched images degraded to the seeing of *F200W* ( $0.07''$ ), whereas non-detection criteria are measured on original images. We identified six more objects in a  $40'' \times 40''$  box (Fig. 3). The photometry of all candidates is reported in Table A.1.

The SMACS0723 ERO dataset also includes NIRISS direct imaging and wide-field slitless spectroscopy (WFSS) in the *F115W* and *F200W* filters, which affords sufficient wavelength coverage to observe Ly $\alpha$  at  $z > 7.2$  or, for example, H $\beta$ , OIII] $\lambda 5007$ , and/or H $\alpha$  for  $z \sim 2$ – $3$  galaxies. We reduced and analysed the dataset as described by Roberts-Borsani et al. (2022a), ensuring the modelling and subtraction of neighbouring sources and optimal extraction of 1D spectra using the 2D source profile of the galaxies. One galaxy (SMACS0723\_PC\_8) resides outside of the NIRISS WFSS footprint, while SMACS0723\_PC\_2 and SMACS0723\_PC\_1 are too faint to be detected in the pre-imaging. We visually inspected the resulting spectra of the remaining galaxies (SMACS0723\_PC\_3, SMACS0723\_PC\_4, SMACS0723\_PC\_5, SMACS\_PC\_6 and SMACS0723\_PC\_7), and found no clear emission lines indicative of Ly $\alpha$  or low- $z$  interlopers. Given the spectroscopic verification of SMACS0723\_PC\_6 and SMACS0723\_PC\_7, the lack of emission lines suggest the galaxies have not carved out sufficiently large ionised bubbles for Ly $\alpha$  to escape, while the absence of strong lines in SMACS0723\_PC\_6 and SMACS0723\_PC\_7 is supportive of this hypothesis rather than a low- $z$  interloper solution.

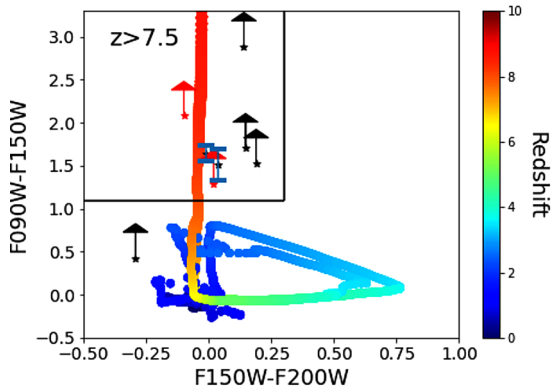
With these numbers in hand we can estimate the overdensity parameter defined in Morselli et al. (2014) as

$$\delta = \frac{\rho}{\dot{\rho}} - 1, \quad (3)$$

where  $\rho$  is the number of objects identified and  $\dot{\rho}$  the number of objects expected from the shape of the UV luminosity function (LF) in our search box. To obtain the latest value, it is necessary to first estimate the surface explored by our survey by masking all bright objects (stars and obvious low- $z$  galaxies) and by correcting each pixel by the magnification by the foreground galaxy cluster SMACS0723–7327. The  $z \sim 8$  UV LF published in Bouwens et al. (2022) is defined between  $z \sim 7.5$  and  $z \sim 8.5$ , and we therefore estimate from the previous effective surface the volume explored within our search box. We find an overdensity parameter of  $\delta = 4.0^{+2.4}_{-1.6}$ . This value is comparable to that computed for protoclusters identified at similar redshift, for example  $\delta = 5.11^{+2.06}_{-1.70}$  for LAGERzOD1 at  $z = 6.9$



**Fig. 1.** One-dimensional NIRSpect spectra of the two galaxies with identical redshift of  $z = 7.66$ . The detected lines are shown with blue dashed lines. A spurious line is seen on the spectrum of SMACS0723\_PC\_6 at  $3.31 \mu\text{m}$  and has been flagged in this figure. Uncertainties on each spectrum are plotted in grey.



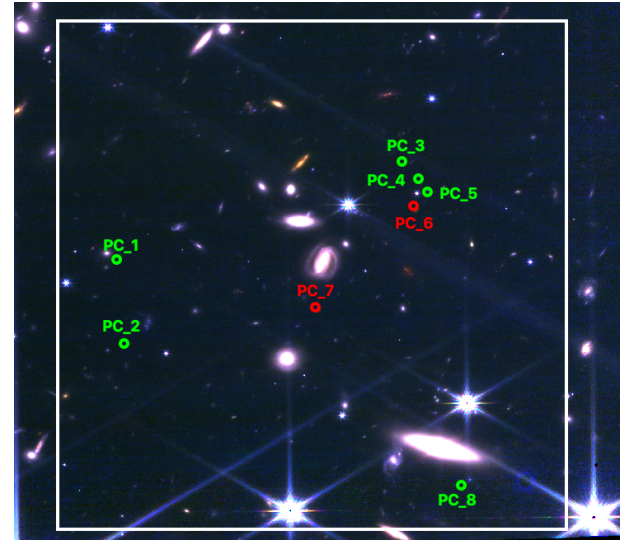
**Fig. 2.** Colour-colour plot used to select galaxies with colour similar to the two galaxies spectroscopically confirmed. The stars show the positions of the eight objects identified in a  $40'' \times 40''$  region. The red dots show the positions of the two spectroscopically confirmed galaxies. Lower limits are computed assuming the  $2\sigma$  depth of the F090W image.

(Hu et al. 2021) and  $\delta \sim 4.5$  for the  $z \sim 8$  overdensity identified in the Brightest of Reionizing Galaxies (BoRG) survey (Trenti et al. 2012). Naidu et al. (2022a) also reported a similar overdensity at  $z = 5.3$  in the Cosmic Evolution Early Release Science (CEERS) Survey data.

### 3. Physical properties of the protocluster

One of the key parameters that must be determined to characterise a protocluster is its total dark matter halo mass. Several methods have recently been used in the literature to probe the total halo mass of confirmed protoclusters. In the following we apply some of these methods to demonstrate that this structure is a convincing protocluster at  $z = 7.66$ .

Before estimating the total halo mass, the stellar mass of each candidate protocluster member needs to be estimated (e.g., Long et al. 2020). We use BAGPIPES (Carnall et al. 2018) and assume several SFHs: constant, burst, delayed, and a combination of a young burst and a constant. We allow a redshift range of  $z \in [0.0:10.0]$ , a stellar mass ranging from  $\log M_\star \in [6.0:12.0]$ , and a reddening range of  $A_V \in [0.0:2.0]$ . The best SED fit is defined as the fit reducing the Bayesian information criterion (BIC; see Laporte et al. 2021 for more details); the results of this



**Fig. 3.** Colour image (F090W, blue; F150W, green; F200W, red) of SMACS0723. The white square displays the protocluster members search box ( $40'' \times 40''$ ), the red circles show the position of the two spectroscopically confirmed galaxies at  $z = 7.66$  and the green circles are at the position of the six other protocluster members candidates identified in this study.

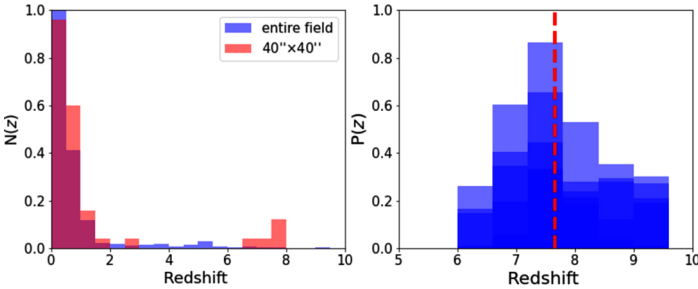
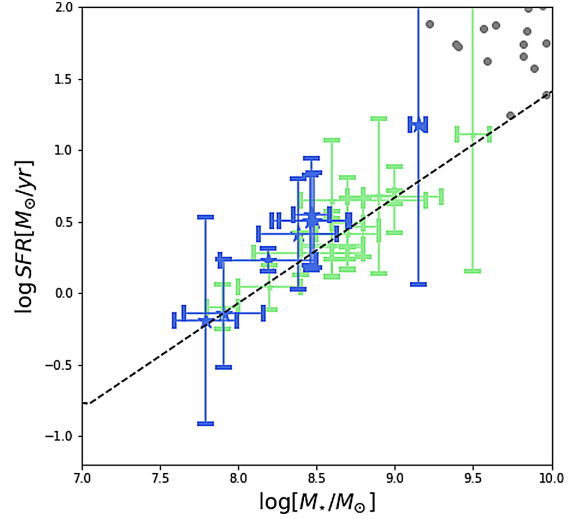
fitting are presented in Table 1. The photometric redshift of the six dropouts identified near the two spectroscopically confirmed galaxies is consistent with these galaxies being at  $z = 7.66$  (Fig. 4, right). We apply the same method to obtain the photometric redshift of all objects in the NIRCам field of view. The distribution in redshift of objects in our search box compared to the distribution in the entire field of view suggests a small excess of objects at  $z \geq 7$ , consistent with the presence of an overdensity in this region (Fig. 4, left). As expected in a protocluster environment (e.g., Muldrew et al. 2015; Lim et al. 2021; Araya-Araya et al. 2021; Gouin et al. 2022), two members are more massive than the others, namely SMACS0723\_PC\_6 and SMACS0723\_PC\_7, with a stellar mass  $\sim 10^9 M_\odot$ .

Figure 5 shows the position of the eight galaxies in our sample on a  $M_\star$  vs. Star-Formation Rate (SFR) diagram compared with previous findings at  $z \geq 7$  (Leethochawalit et al. 2022; Topping et al. 2022). It confirms that all these sources have properties consistent with what is expected in terms of the star



**Table 1.** Physical properties computed with BAGPIPES of the protocluster member candidates identified behind SMACS0723.

| ID    | RA<br>[deg] | Dec<br>[deg] | $z_{\text{phot}}$      | $\log M_{\star}$<br>[ $M_{\odot}$ ] | SFR<br>[ $M_{\odot} \text{ yr}^{-1}$ ] | $A_v$<br>[mag]         | $\mu$ [95% C.I.] |
|-------|-------------|--------------|------------------------|-------------------------------------|----------------------------------------|------------------------|------------------|
| PC_1  | 110.823826  | -73.437880   | $6.68^{+0.04}_{-0.06}$ | $7.79^{+0.20}_{-0.13}$              | $3.37^{+0.45}_{-0.18}$                 | $0.35^{+0.30}_{-0.23}$ | 1.90 [1.49–2.06] |
| PC_2  | 110.820512  | -73.436305   | $7.69^{+1.88}_{-0.85}$ | $8.19^{+0.26}_{-0.31}$              | $1.71^{+0.88}_{-0.77}$                 | $0.11^{+0.11}_{-0.07}$ | 1.79 [1.44–1.92] |
| PC_3  | 110.845907  | -73.435998   | $7.76^{+1.19}_{-1.02}$ | $7.96^{+0.32}_{-0.36}$              | $0.72^{+0.31}_{-0.25}$                 | $0.12^{+0.15}_{-0.09}$ | 1.73 [1.41–1.83] |
| PC_4  | 110.846149  | -73.435474   | $7.20^{+0.55}_{-0.45}$ | $8.57^{+0.22}_{-0.26}$              | $3.25^{+1.09}_{-0.83}$                 | $0.33^{+0.14}_{-0.16}$ | 1.70 [1.39–1.80] |
| PC_5  | 110.847412  | -73.435129   | $8.19^{+0.83}_{-0.57}$ | $8.50^{+0.20}_{-0.23}$              | $3.23^{+1.17}_{-1.18}$                 | $0.19^{+0.12}_{-0.11}$ | 1.68 [1.54–1.77] |
| PC_6* | 110.844634  | -73.435054   | <i>7.665</i>           | $8.59^{+0.19}_{-0.21}$              | $3.55^{+1.08}_{-1.14}$                 | $0.15^{+0.09}_{-0.07}$ | 1.68 [1.38–1.78] |
| PC_7* | 110.834062  | -73.434509   | <i>7.659</i>           | $8.95^{+0.04}_{-0.04}$              | $15.06^{+2.02}_{-1.89}$                | $1.02^{+0.03}_{-0.03}$ | 1.68 [1.41–1.79] |
| PC_8  | 110.835151  | -73.429499   | $7.33^{+1.02}_{-0.09}$ | $8.26^{+0.16}_{-0.08}$              | $2.61^{+2.20}_{-0.57}$                 | $0.54^{+0.04}_{-0.06}$ | 1.50 [1.29–1.57] |

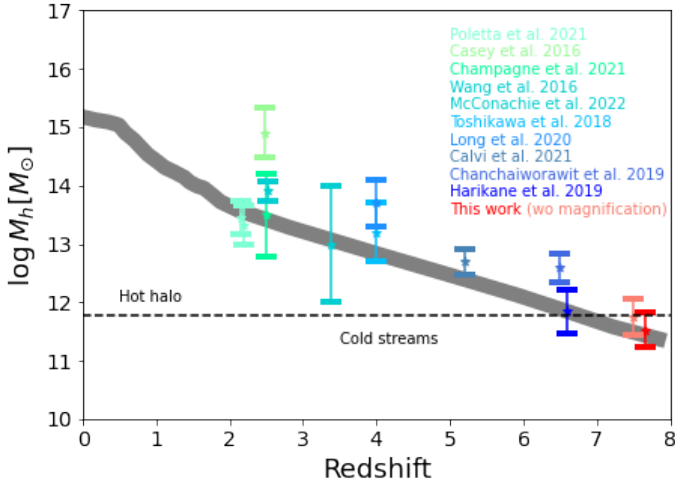
**Notes.** All the values are not corrected for the best-fit magnification. The last column shows the magnification factor estimated by the lens model presented in [Pascale et al. \(2022\)](#), assuming a redshift of  $z = 7.66$  for all objects. \*spectroscopically confirmed at  $z = 7.66$ . Spectroscopic redshift are indicated in italics.**Fig. 4.** Photometric redshift in SMACS0723 *Left*: distribution of the detected objects as a function of redshift in the entire field of view (blue) and in the  $40'' \times 40''$  box around the two spectroscopically confirmed galaxies at  $z = 7.66$ . The histograms are normalised to 1. A small excess of  $z \geq 7$  sources is observed in the protocluster region compared to the entire field of view. *Right*: redshift probability distribution for the six protocluster members candidates identified in this study. The distribution is compatible with a spectroscopic redshift of  $z = 7.66$  (vertical dashed red line).**Fig. 5.** Star formation main sequence for the eight galaxies studied in this paper (blue dots) compared with previously published  $z \geq 7$  sources for which stellar masses are available (green, [Leethochawalit et al. 2022](#); grey, [Topping et al. 2022](#)). The dashed line shows the parametrisation found by [Leethochawalit et al. \(2022\)](#).

formation main sequence currently known at  $z \geq 7$ . We also expand our search over the entire NIRCcam field of view to identify other  $z \sim 7.66$  objects. Twenty-six objects follow our selection criteria (including the eight protocluster member candidates discussed here), but no other region in the field shows an overdensity as large as in the protocluster candidate region. Moreover, no candidate has a stellar mass higher than PC\_7, which confirms that the two spectroscopically confirmed galaxies are the most massive  $z = 7.66$  galaxies in this field of view. We also note that the projected distance between the protocluster core discussed in this paper and the closest  $z \sim 7.66$  galaxy outside of the  $40'' \times 40''$  box is 0.38 Mpc, suggesting that some of these galaxies may also be related to the protocluster.

The individual halo mass can be estimated from the stellar mass of each galaxy using the [Behroozi et al. \(2013\)](#) relationship. The halo masses of the eight galaxies range from  $2 \times 10^{10}$  to  $6 \times 10^{11} M_{\odot}$  corrected for magnification, with a total protocluster halo mass of  $M_h = 3.34^{+0.59}_{-0.50} \times 10^{11} M_{\odot}$ . Another method of estimating the halo mass of a protocluster is to sum all stellar masses ( $M_{\star}^{\text{tot}} = 1.46^{+0.63}_{-0.41} \times 10^9 M_{\odot}$ ) corrected for magnification and to convert into halo mass using the baryonic-to-dark matter fraction measured by [Planck Collaboration VI \(2020\)](#). Following this method we estimate a total halo mass of  $6.52^{+2.82}_{-1.85} \times 10^{10} M_{\odot}$ . Finally, we can also determine the halo mass of the most massive galaxy in our sample from its stellar mass,

which should dominate the total halo mass of the protocluster. SMACS0723\_PC\_7 has a stellar mass of  $5.31^{+0.54}_{-0.41} \times 10^8 M_{\odot}$  corrected for magnification, which converts into a minimum halo mass of  $6.86^{+2.51}_{-2.42} \times 10^{10} M_{\odot}$ .

In an independent analysis, we use the observed stellar-to-halo mass ratios measured by [Shuntov et al. \(2022\)](#) with the COSMOS2020 catalogue, which represents the most complete deep galaxy catalogue to date ([Weaver et al. 2022](#)), to compute an additional measurement of the halo mass of our protocluster candidate. We fit the redshift-evolution of the [Shuntov et al. \(2022\)](#) stellar-to-halo mass ratio and extrapolate it out to the redshift of our cluster ( $z_{\text{spec}} = 7.66$ ) using a Markov chain Monte Carlo (MCMC) analysis to rigorously propagate the uncertainties. The resulting halo masses of the eight cluster member galaxies range from  $2 \times 10^{10}$  to  $1 \times 10^{11} M_{\odot}$ , which broadly agrees with the range found using the [Behroozi et al. \(2013\)](#) relation above. Summing over these halo masses, we find a total halo mass of  $M_h = 2.07 \pm 1.51 \times 10^{11} M_{\odot}$ , which is higher than our previous estimates but agrees within the uncertainties ( $1\sigma$ ). We note that with this conversion the halo mass of the most massive



**Fig. 6.** Evolution of protocluster halo mass as a function of redshift. The grey shaded region displays the expected halo mass evolution of a Coma-like cluster (from Chiang et al. 2013) assuming a smooth evolution above  $z \geq 7$ . The dashed line shows the typical threshold mass for a stable shock in a spherical infall, below which the flows are predominantly cold and above which a shock-heated medium is present (from Dekel & Birnboim 2006). Overplotted are the halo mass of recently studied protoclusters at  $z \geq 2$  from Poletta et al. (2021), Casey (2016), Champagne et al. (2021), Wang et al. (2016), McConachie et al. (2022), Toshikawa et al. (2018), Long et al. (2020), Calvi et al. (2021), Chanchaiworawit et al. (2019), and Harikane et al. (2019). Our halo mass estimate (red dot, the sum of halo masses of individual members) is fully consistent with that expected for a Coma-like cluster. The pink point shows the total halo mass without accounting for magnification (shifted to slightly lower redshift for purposes of clarity).

galaxy in the protocluster is also higher than our previous estimate, with a halo mass of  $10.32 \pm 9.78 \times 10^{10} M_{\odot}$ , though it also has much higher uncertainties. Moreover, Chiang et al. (2013) studied the evolution of a Coma-like cluster from  $z \sim 7$  to  $z = 0$ . Assuming a smooth evolution at  $z \geq 7$ , the expected halo mass of a protocluster at  $z = 7.66$  is  $\sim 3 \times 10^{11} M_{\odot}$ . This value agrees well with the total halo mass we estimated for our protocluster candidate.

In addition to the halo mass of the protocluster and the main sequence, another relevant physical property is its size. As seen in Fig. 3, the protocluster constituent galaxies are all found within a  $40'' \times 40''$  box, or within a radius of  $\sim 20''$  from the two brighter, central protocluster galaxies. We use the analytic mass model presented in Pascale et al. (2022) to estimate the magnification of the protocluster. In Table 1 we list the magnifications of each of the eight galaxies associated with the protocluster. The magnifications vary between 1.5 and 2, roughly, with a typical  $\mu \approx 1.7$  at the centre of the protocluster. We do not find a strong shear at the location of the protocluster so that a circle of radius  $\sim 20''$  around the observed protocluster galaxies, is delensed into an ellipse of axis ratio 1:1.35 (approximately), located a few arcseconds closer to the bright central galaxy. Taking the typical magnification into account ( $\mu \approx 1.7$ ), we estimate the size of the protocluster to be about 120 kpc in diameter. This size is consistent with what is expected from the Chiang et al. (2017) simulations. It is also similar with what was observed at  $z = 5.3$  in Capak et al. (2011) and at  $z = 6.5$  by Chanchaiworawit et al. (2019), with protocluster core size of 0.14 Mpc and 0.45 Mpc.

#### 4. Implications for cosmic reionisation

The transmission and escape of Lyman- $\alpha$  and Lyman continuum photons through the intergalactic medium around a galaxy depends exponentially on the optical depth (Furlanetto et al. 2004; Miralda-Escudé et al. 2000). This means that even small amounts of neutral hydrogen will absorb potentially ionising photons. As a result, smaller and fainter galaxies can only very locally ionise their surroundings, if at all. This leads to a strong bias towards detecting Lyman- $\alpha$  from brighter and more massive galaxies in the reionisation era (e.g., Stark et al. 2011) since typically only these galaxies radiate strongly enough to ionise the surrounding hydrogen and create a sufficiently large bubble to allow these photons to escape (e.g., Leonova et al. 2022; Larson et al. 2022). Hence, smaller galaxies located close to these more massive and brighter galaxies are effectively in an ionised bubble, and their UV photons can travel large distances to help reionise the Universe (photons that would otherwise be absorbed in the local vicinity of the galaxies). Therefore, protoclusters may play a significant and important role in the reionisation process.

Chiang et al. (2017) demonstrated, using  $N$ -body simulations and semi-analytical models that protoclusters could have contributed up to  $\sim 50\%$  of the cosmic star formation rate density at  $z = 10$ . They also show that a protocluster core at  $z = 7.66$  can itself represent  $\sim 10\%$  of the total ionising budget. Furthermore, Ma et al. (2021) analyse in their simulations a protocluster with a halo mass similar to the protocluster we report in this Letter ( $M_h \sim 10^{11} M_{\odot}$ ), and conclude that the escape fraction for this type of protocluster could reach  $f_{\text{esc}} \sim 20\%$ , a golden number to explain the end of the reionisation process by  $z \sim 6$ .

Moreover, the latest constraints on the shape of the UV luminosity function at  $z \geq 6$  show that the number density of galaxies within the first billion years of the Universe is  $\leq 10^{-6} \text{ Mpc}^{-3}$  (e.g., Bouwens et al. 2022; Finkelstein & Ryan 2015), which is similar to the number density of rich clusters at  $z = 0$  ( $\geq 10^{15} M_{\odot}$ ; e.g., Wen et al. 2010). This could suggest that the brightest  $z \geq 6$  galaxies detected in deep HST surveys may lie in overdense regions with the vast majority of sources well below the detection limit of *Hubble*. If this is the case, and assuming that each  $z \geq 6$  protocluster has an escape fraction of  $\sim 20\%$ , as suggested by simulations, it could open a new route to solve the UV photon deficit observed during the reionisation process.

Figure 6 demonstrates that SMACS0723\_PC could be seen as the progenitor of a Coma-like cluster ( $M_h \geq 10^{15} M_{\odot}$ ), and illustrates the importance of studying galaxy protoclusters from cosmic dawn to cosmic noon (e.g., in this  $M_h$ - $z$  plane). Previous telescopes were only able to detect the brightest members of such structures. Even if recent HST studies start to see overdensities near extremely bright objects at  $z \geq 7$  (e.g., Leonova et al. 2022; Castellano et al. 2022), the density of protoclusters within the first billion years of the Universe is totally unknown. The identification of a protocluster candidate at  $z = 7.66$  in less than 15 h with JWST is encouraging. Wider and deeper data are now needed to determine how many bright  $z \geq 7$  sources lie in overdense regions, which could give new insights into the cosmic reionisation.

#### 5. Summary

In this Letter we report the detection of a protocluster candidate at  $z = 7.66$  behind the lensing cluster SMACS0723-7327 observed with the *James Webb* Space Telescope. In addition to the spectroscopic confirmation of two  $z = 7.66$  sources, we

identified six more galaxies with similar colours, whose SED fitting suggests that they may lie at the same redshift, and with physical properties comparable to galaxies at the same redshift. Assuming they are all part of the same structure, we estimate an overdensity parameter  $\delta = 4.0^{+2.4}_{-1.6}$ , consistent with previous values found for other protoclusters at  $z \geq 6$ . Based on several methods, we estimate the total dark matter halo mass of this protocluster candidate to be  $M_h = 3.6^{+13.3}_{-2.8} \times 10^{11} M_\odot$ . This value agrees perfectly with what is expected for progenitors of a Coma-like cluster. Furthermore, the star formation main sequence at  $z \geq 7$  and the estimated size are in line with expectations.

Simulations predict that protoclusters may have played an important role in the cosmic reionisation, with up to a 50% contribution to the cosmic star formation rate density. They also suggest that protoclusters with the dark matter halo mass of SMACS0723\_PC could have an escape fraction as high as 20%. The main unknown parameter is their number density. However, the number density of bright galaxies at  $z \geq 6$  is of the same order of magnitude as the number density of rich clusters at  $z \sim 0$ , suggesting that they may be linked. If this is confirmed with the detection of protocluster candidates around many bright galaxies at  $z \geq 6$  with the JWST, and given the efficiency at which these early structures ionise the neutral hydrogen, it could give new insight into the reionisation process.

**Acknowledgements.** We thank Francesco D'Eugenio for providing NIRSpectra and Richard Ellis for interesting comments on our manuscript. NL acknowledges support from the Kavli foundation. AZ and LF acknowledge support by Grant No. 2020750 from the United States-Israel Binational Science Foundation (BSF) and Grant No. 2109066 from the United States National Science Foundation (NSF), and by the Ministry of Science & Technology, Israel. CW acknowledges support from the Science and Technology Facilities Council (STFC) for a Ph.D. studentship. NL and HD acknowledge the Astr'Auvergne astronomy festival (and InfiniSciences and Astro-Jeunes volunteers) as it all started there. This work is based in part on observations made with the NASA/ESA/CSA *James Webb* Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. The authors acknowledge the ERO team (program #2736) for developing their observing program with a zero-exclusive-access period.

## References

- Araya-Araya, P., Vicentin, M. C., Sodré, L. Jr., Overzier, R. A., & Cuevas, H. 2021, *MNRAS*, **504**, 5054
- Atek, H., Shuntov, M., Furtak, L.J., et al. 2022, *MNRAS*, submitted [arXiv:2207.12338]
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, **770**, 57
- Bouwens, R. J., Illingworth, G. D., Ellis, R. S., Oesch, P. A., & Stefanon, M. 2022, *A&AS*, submitted [arXiv:2205.11526]
- Bromm, V., & Yoshida, N. 2011, *ARA&A*, **49**, 373
- Calvi, R., Dannerbauer, H., Arrabal Haro, P., et al. 2021, *MNRAS*, **502**, 4558
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, *Nature*, **470**, 233
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, *MNRAS*, **480**, 4379
- Carnall, A. C., Begley, R., McLeod, D. J., et al. 2022, *MNRAS*, submitted [arXiv:2207.08778]
- Casey, C. M. 2016, *ApJ*, **824**, 36
- Castellano, M., Dayal, P., Pentericci, L., et al. 2016, *ApJ*, **818**, L3
- Castellano, M., Pentericci, L., Cupani, G., et al. 2022, *A&A*, **662**, A115
- Champagne, J. B., Casey, C. M., Zavala, J. A., et al. 2021, *ApJ*, **913**, 110
- Chanchaiworawit, K., Guzmán, R., Salvador-Solé, E., et al. 2019, *ApJ*, **877**, 51
- Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2013, *ApJ*, **779**, 127
- Chiang, Y.-K., Overzier, R. A., Gebhardt, K., & Henriques, B. 2017, *ApJ*, **844**, L23
- Curti, M., D'Eugenio, F., Carniani, S., et al. 2022, *MNRAS*, in press [arXiv:2207.12375]
- De Barros, S., Pentericci, L., Vanzella, E., et al. 2017, *A&A*, **608**, A123
- Dekel, A., & Birnboim, Y. 2006, *MNRAS*, **368**, 2
- Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2022, *MNRAS*, submitted [arXiv:2207.12356]
- Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., et al. 2015, *ApJ*, **810**, 71
- Finkelstein, S.L., Bagley, M.B., Arrabal Haro, P., et al. 2022, *ApJ*, in press [arXiv:2207.12474]
- Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, *ApJ*, **613**, 1
- Genel, S., Bouché, N., Naab, T., Sternberg, A., & Genzel, R. 2010, *ApJ*, **719**, 229
- Gouin, C., Aghanim, N., Dole, H., Polletta, M., & Park, C. 2022, *A&A*, **664**, A155
- Harikane, Y., Ouchi, M., Ono, Y., et al. 2019, *ApJ*, **883**, 142
- Hu, W., Wang, J., Infante, L., et al. 2021, *Nat. Astron.*, **5**, 423
- Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, *MNRAS*, **442**, 1805
- Katz, H., Saxena, A., Cameron, A. J., et al. 2022, *MNRAS*, in press [arXiv:2207.13693]
- Laporte, N., Ellis, R. S., Boone, F., et al. 2017, *ApJ*, **837**, L21
- Laporte, N., Meyer, R. A., Ellis, R. S., et al. 2021, *MNRAS*, **505**, 3336
- Larson, R. L., Finkelstein, S. L., Hutchison, T. A., et al. 2022, *ApJ*, **930**, 104
- Leethochawalit, N., Trenti, M., Santini, P., et al. 2022, *ApJ*, accepted [arXiv:2207.11135]
- Leonova, E., Oesch, P. A., Qin, Y., et al. 2022, *MNRAS*, **515**, 5790
- Lim, S., Scott, D., Babul, A., et al. 2021, *MNRAS*, **501**, 1803
- Long, A. S., Cooray, A., Ma, J., et al. 2020, *ApJ*, **898**, 133
- Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *ApJ*, **837**, 97
- Ma, X., Quataert, E., Wetzel, A., Faucher-Giguère, C.-A., & Boylan-Kolchin, M. 2021, *MNRAS*, **504**, 4062
- McConachie, I., Wilson, G., Forrest, B., et al. 2022, *ApJ*, **926**, 37
- Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 2000, *ApJ*, **530**, 1
- Morselli, L., Mignoli, M., Gilli, R., et al. 2014, *A&A*, **568**, A1
- Muldrew, S. I., Hatch, N. A., & Cooke, E. A. 2015, *MNRAS*, **452**, 2528
- Naidu, R. P., Oesch, P. A., Setton, D. J., et al. 2022a, *ApJ*, submitted [arXiv:2208.02794]
- Naidu, R.P., Oesch, P.A., van Dokkum, P., et al. 2022b, *ApJ*, submitted [arXiv:2207.09434]
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Pascale, M., Frye, B. L., Diego, J., et al. 2022, *ApJ*, submitted [arXiv:2207.07102]
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Polletta, M., Soucail, G., Dole, H., et al. 2021, *A&A*, **654**, A121
- Roberts-Borsani, G., Morishita, T., Treu, T., et al. 2022a, *ApJ*, submitted [arXiv:2207.11387]
- Roberts-Borsani, G., Treu, T., Mason, C., et al. 2022b, *ApJ*, submitted [arXiv:2207.01629]
- Shuntov, M., McCracken, H. J., Gavazzi, R., et al. 2022, *A&A*, **664**, A61
- Stark, D. P., Ellis, R. S., & Ouchi, M. 2011, *ApJ*, **728**, L2
- Tilvi, V., Malhotra, S., Rhoads, J. E., et al. 2020, *ApJ*, **891**, L10
- Topping, M. W., Stark, D. P., Endsley, R., et al. 2022, *MNRAS*, **516**, 975
- Toshikawa, J., Uchiyama, H., Kashikawa, N., et al. 2018, *PASJ*, **70**, S12
- Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2012, *ApJ*, **746**, 55
- Trussler, J. A. A., Adams, N. J., Conselice, C. J., et al. 2022, *ArXiv e-prints* [arXiv:2207.14265]
- Wang, T., Elbaz, D., Daddi, E., et al. 2016, *ApJ*, **828**, 56
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, *ApJS*, **258**, 11
- Wen, Z. L., Han, J. L., & Liu, F. S. 2010, *MNRAS*, **407**, 533
- Zheng, W., Shu, X., Moustakas, J., et al. 2014, *ApJ*, **795**, 93

## Appendix A: Photometry of the protocluster members candidates

**Table A.1.** Photometry of the selected protocluster member candidates.

| ID   | <i>F</i> 435W<br>ACS/HST | <i>F</i> 606W<br>ACS/HST | <i>F</i> 814W<br>ACS/HST | <i>F</i> 090W<br>NIRCam/JWST | <i>F</i> 105W<br>WFC3/HST | <i>F</i> 125W<br>WFC3/HST | <i>F</i> 140W<br>WFC3/HST | <i>F</i> 150W<br>NIRCam/JWST | <i>F</i> 160W<br>WFC3/HST | <i>F</i> 200W<br>NIRCam/JWST | <i>F</i> 277W<br>NIRCam/JWST | <i>F</i> 356W<br>NIRCam/JWST | <i>F</i> 444W<br>NIRCam/JWST |
|------|--------------------------|--------------------------|--------------------------|------------------------------|---------------------------|---------------------------|---------------------------|------------------------------|---------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| PC_1 | > 27.22                  | > 27.84                  | > 26.80                  | 29.22 ± 0.47                 | > 27.63                   | > 27.49                   | > 27.77                   | 27.71 ± 0.13                 | > 27.40                   | 27.67 ± 0.11                 | 27.80 ± 0.24                 | 28.19 ± 0.33                 | 28.37 ± 0.56                 |
| PC_2 | > 27.72                  | > 28.47                  | > 27.55                  | > 29.93                      | > 27.51                   | > 28.06                   | > 27.10                   | 28.47 ± 0.15                 | > 28.00                   | 28.76 ± 0.18                 | 29.19 ± 0.48                 | 28.71 ± 0.31                 | 29.33 ± 0.72                 |
| PC_3 | > 27.67                  | > 28.21                  | > 27.68                  | > 29.69                      | > 27.38                   | 27.78 ± 0.54              | 27.00 ± 0.25              | 27.38 ± 0.07                 | 26.85 ± 0.25              | 27.19 ± 0.05                 | 27.25 ± 0.14                 | 27.16 ± 0.12                 | 27.48 ± 0.23                 |
| PC_4 | > 27.59                  | > 27.99                  | > 27.47                  | > 29.65                      | 27.56 ± 0.51              | > 26.77                   | 26.94 ± 0.28              | 27.19 ± 0.06                 | 27.08 ± 0.31              | 27.05 ± 0.05                 | 27.30 ± 0.18                 | 27.28 ± 0.14                 | 27.28 ± 0.18                 |
| PC_5 | > 27.76                  | > 28.31                  | > 27.63                  | 28.21 ± 0.13                 | 27.34 ± 0.37              | > 25.89                   | 26.70 ± 0.20              | 26.57 ± 0.03                 | 26.53 ± 0.18              | 26.58 ± 0.03                 | 26.68 ± 0.08                 | 26.75 ± 0.07                 | 26.79 ± 0.11                 |
| PC_6 | > 27.66                  | > 28.23                  | > 27.72                  | > 29.73                      | 27.46 ± 0.43              | 26.20 ± 0.43              | 26.26 ± 0.14              | 26.02 ± 0.02                 | 26.04 ± 0.12              | 25.88 ± 0.02                 | 25.74 ± 0.04                 | 25.49 ± 0.03                 | 24.60 ± 0.02                 |
| PC_7 | > 27.76                  | > 28.07                  | > 27.69                  | > 29.83                      | 26.95 ± 0.33              | 27.06 ± 0.24              | 26.54 ± 0.35              | 26.82 ± 0.04                 | 26.78 ± 0.20              | 26.92 ± 0.04                 | 26.87 ± 0.07                 | 26.85 ± 0.07                 | 26.09 ± 0.05                 |
| PC_8 | > 27.70                  | > 28.22                  | > 27.65                  | > 28.23                      | OOF                       | OOF                       | OOF                       | 27.62 ± 0.19                 | OOF                       | 27.60 ± 0.14                 | 27.23 ± 1.10                 | 27.37 ± 0.33                 | >27.19                       |

**Notes.**  $2\sigma$  non-detections are measured at the position of the candidate in a  $0.3''$  radius aperture. Object PC\_8 is not covered by current HST data, and is marked as out of the field (OOF).