


# A population of red candidate massive galaxies ~600 Myr after the Big Bang

<https://doi.org/10.1038/s41586-023-05786-2>

Received: 25 July 2022

Accepted: 2 February 2023

Published online: 22 February 2023

 Check for updates

Ivo Labbé<sup>1,2</sup>, Pieter van Dokkum<sup>3</sup>, Erica Nelson<sup>3</sup>, Rachel Bezanson<sup>4</sup>, Katherine A. Suess<sup>5,6</sup>, Joel Leja<sup>7,8</sup>, Gabriel Brammer<sup>10</sup>, Katherine Whitaker<sup>10,11</sup>, Elijah Mathews<sup>7,8,9</sup>, Mauro Stefanon<sup>12,13</sup> & Bingjie Wang<sup>7,8,9</sup>

Galaxies with stellar masses as high as roughly  $10^{11}$  solar masses have been identified<sup>1–3</sup> out to redshifts  $z$  of roughly 6, around 1 billion years after the Big Bang. It has been difficult to find massive galaxies at even earlier times, as the Balmer break region, which is needed for accurate mass estimates, is redshifted to wavelengths beyond  $2.5 \mu\text{m}$ . Here we make use of the  $1–5 \mu\text{m}$  coverage of the James Webb Space Telescope early release observations to search for intrinsically red galaxies in the first roughly 750 million years of cosmic history. In the survey area, we find six candidate massive galaxies (stellar mass more than  $10^{10}$  solar masses) at  $7.4 \leq z \leq 9.1$ , 500–700 Myr after the Big Bang, including one galaxy with a possible stellar mass of roughly  $10^{11}$  solar masses. If verified with spectroscopy, the stellar mass density in massive galaxies would be much higher than anticipated from previous studies on the basis of rest-frame ultraviolet-selected samples.

The galaxies were identified in the first observations of the James Webb Space Telescope (JWST) Cosmic Evolution Early Release Science (CEERS) program. This program obtained multiband imaging at  $1–5 \mu\text{m}$  with the Near-Infrared Camera (NIRCam) in a ‘blank’ field, chosen to overlap with existing Hubble Space Telescope (HST) imaging. The total area covered by these initial data is roughly  $40 \text{ arcmin}^2$ . The data were obtained from the Mikulski Archive for Space Telescopes (MAST) archive and reduced using the Grizli pipeline<sup>4</sup>. A catalogue of sources was created, starting with detection in a deep combined F277W + F356W + F444W image (see Methods for details). A total of 42,729 objects are in this parent catalogue.

We selected candidate massive galaxies at high redshifts by identifying objects that have two redshifted breaks in their spectral energy distributions (SEDs), the  $\lambda_{\text{Ly}\alpha} = 1.216 \text{ \AA}$  Lyman break and the  $\lambda_{\text{H}\delta}$  of roughly  $3,600 \text{ \AA}$  Balmer break. This selection ensures that the redshift probability distributions are well constrained, have no secondary solutions at lower redshifts and that we include galaxies that have potentially high mass-to-light (M/L) ratios. Specifically, we require that objects are not detected at optical wavelengths, blue in the near-infrared with F150W–F277W is less than 0.7, red at longer wavelengths with F277W–F444W is more than 1.0 and brighter than F444W is less than 27 AB units in magnitude. After visual inspection to remove obvious artefacts (such as diffraction spikes), this selection produced 13 galaxies with the sought-for ‘double-break’ SEDs. Next, redshifts and stellar masses were determined with three widely used techniques, taking the contribution of strong emission lines to the rest-frame optical photometry explicitly

into account<sup>5–15</sup>. We use the EAZY code<sup>16</sup> (with extra strong emission-line templates), the Prospector- $\alpha$  framework<sup>17</sup> and five configurations of the Bagpipes SED-fitting code to explore systematics due to modelling assumptions. The seven individual mass and redshift measurements of the 13 galaxies are listed in the Methods section. We adopt fiducial masses and redshifts by taking the median value for each galaxy. We note that these masses and redshifts are not definitive and that all galaxies should be considered candidates.

As shown in Fig. 1, all 13 objects have photometric redshifts  $6.5 < z < 9.1$ . Six of the 13 have fiducial masses greater than  $10^{10} M_{\odot}$  (Salpeter initial mass function (IMF)) and multiband images and SEDs of these galaxies are shown in Figs. 2 and 3. Their photometric redshifts range from  $z = 7.4$  to  $z = 9.1$ . The model fits are generally excellent, and in several cases clearly demonstrate that rest-frame optical emission lines contribute to the continuum emission. These lines can be so strong in young galaxies that they can dominate the broad-band fluxes redwards of the location of the Balmer break<sup>6–8,14,15</sup> and Spitzer/IRAC detections of optical continuum breaks in galaxies at  $z \geq 5$  have been challenging to interpret<sup>3,5,19–24</sup>. With JWST, this ambiguity is largely resolved due to the dense wavelength coverage of the NIRCam filters and the inclusion of relatively narrow emission-line-sensitive filter F410M (ref. 25), which falls within the F444W band, although the uncertainties are such that alternative solutions with lower masses may exist<sup>14</sup>. The brightest galaxy in the sample, 38094, is at  $z = 7.5$  and may have a mass that is as high as  $M_{*} \approx 1 \times 10^{11} M_{\odot}$ , more massive than the present-day Milky Way. It has two nearby companions with a similar

<sup>1</sup>Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, Victoria, Australia. <sup>2</sup>Department of Astronomy, Yale University, New Haven, CT, USA.

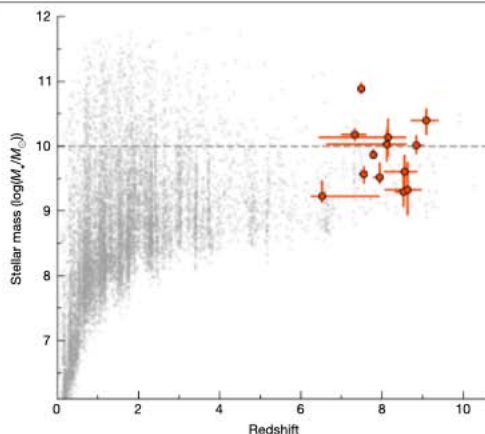
<sup>3</sup>Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO, USA. <sup>4</sup>Department of Physics and Astronomy and PITT PACS, University of Pittsburgh, Pittsburgh, PA, USA.

<sup>5</sup>Department of Astronomy and Astrophysics, University of California, Santa Cruz, Santa Cruz, CA, USA. <sup>6</sup>Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA, USA.

<sup>7</sup>Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA, USA. <sup>8</sup>Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, PA, USA.

<sup>9</sup>Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA, USA. <sup>10</sup>Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Copenhagen N, Denmark. <sup>11</sup>Department of Astronomy, University of Massachusetts, Amherst, MA, USA. <sup>12</sup>Department of Astronomy and Astrophysics, University of Valencia Burjassot, Valencia, Spain.

<sup>13</sup>University Association CSIC ‘Group of Extragalactic Astronomy and Cosmology’ (Institute of Physics at Cantabria University of Valencia), Santander, Spain. <sup>14</sup>e-mail: ilabbe@swin.edu.au



**Fig. 1 | Redshifts and tentative stellar masses of double-break selected galaxies.** Shown in grey circles are EAZY-determined redshifts and stellar masses using emission-line enhanced templates (Salpeter IMF) for objects with SNR > 8 in the F444W band. Fiducial redshifts and masses of the bright galaxies (F444W < 27 AB) that satisfy our double-break selection are shown by the large red symbols. Uncertainties are the 16th–84th percentiles of the posterior probability distribution. All galaxies have photometric redshifts  $6.5 < z < 9.1$ . Six galaxies are candidate massive galaxies with fiducial  $M_* > 10^{10} M_\odot$ .

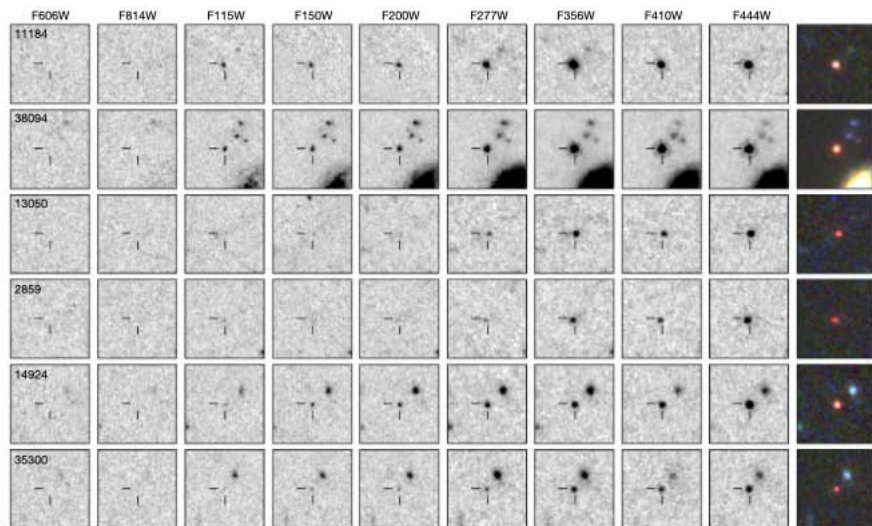
break in their optical to near-infrared SEDs, suggesting that the galaxy may be in a group.

We place these results in context by comparing them to previous studies of the evolution of the galaxy mass function to  $z$  of roughly 9. These studies are based on samples that were selected in the rest-frame

ultraviolet (UV) using ultra-deep HST images, with Spitzer/IRAC photometry typically acting as a constraint on the rest-frame optical SED<sup>3,15,26–28</sup>. The bottom panel of Fig. 3 compares the average SED of the six candidate massive galaxies to the SEDs of HST-selected galaxies at similar redshifts. The galaxies we report here are much redder and the differences are not limited to one or two photometric bands: the entire SED is qualitatively different. This is the key result of our study: we show that galaxies can be robustly identified at  $z > 7$  with JWST that are intrinsically redder than previous HST-selected samples at the same redshifts. It is likely that these galaxies also have much higher M/L ratios, but this needs to be confirmed with spectroscopy. We note that the new galaxies are very faint in the rest-frame UV (median F150W of roughly 28 AB), and previous wide-field studies with HST and Spitzer<sup>29</sup> of individual galaxies did not reach the required depths to find this population.

The masses that we derive are intriguing in the context of previous studies. No candidate galaxies with  $\log(M_*/M_\odot) > 10.5$  had been found before beyond  $z$  of roughly 7, and no candidates with  $\log(M_*/M_\odot) > 10$  had been found beyond  $z$  of roughly 8. Furthermore, Schechter fits to the previous candidates predicted extremely low number densities of such galaxies at the highest redshifts<sup>3</sup>. This is shown by the lines in Fig. 4: the expected mass density in galaxies with  $\log(M_*/M_\odot) > 10$  at  $z$  of roughly 9 was roughly  $10^3 M_\odot \text{Mpc}^{-3}$ , and the total previously derived stellar mass density, integrated over the range  $8 < \log(M_*/M_\odot) < 12$ , is less than  $10^4 M_\odot \text{Mpc}^{-3}$ . If confirmed, the JWST-selected objects would fall in a different region of Fig. 4, in the top right, as the JWST-derived fiducial mass densities are far higher than the expected values on the basis of the UV-selected samples. The mass in galaxies with  $\log(M_*/M_\odot) > 10$  would be a factor of roughly 20 higher at  $z$  of around 8 and a factor of roughly 1,000 higher at  $z$  of roughly 9. The differences are even greater for  $\log(M_*/M_\odot) > 10.5$ .

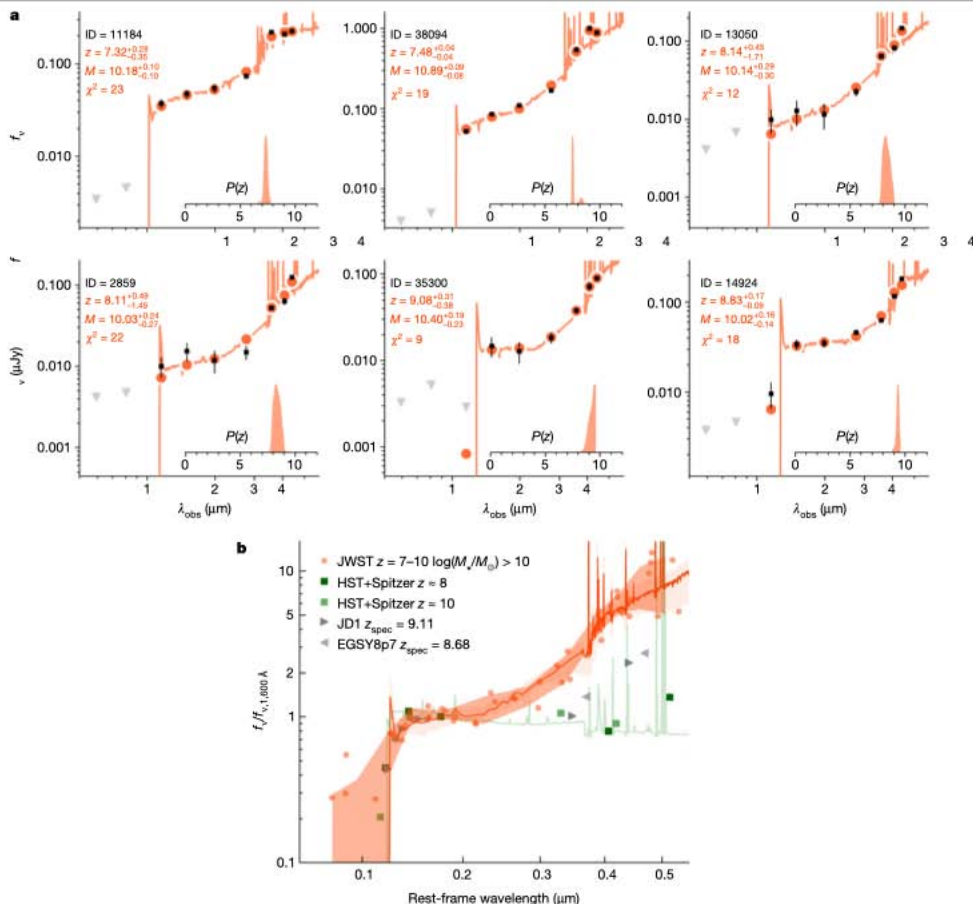
We infer that the possible interpretation of these JWST-identified ‘optical break galaxies’ falls between two extremes. If the redshifts and fiducial masses are correct, then the mass density in the most massive galaxies would exceed the total previously estimated mass



**Fig. 2 | Images of the six galaxies with the highest apparent masses as a function of wavelength.** The fiducial stellar masses of the galaxies are  $\log(M_*/M_\odot) > 10$ . Each cut-out has a size of  $2.4'' \times 2.4''$ . The filters range from the 0.6  $\mu\text{m}$  F606W filter of HST/ACS to the 4.4  $\mu\text{m}$  F444W JWST/NIRCam filter.

The galaxies are undetected in the optical filters, blue in the short-wavelength NIRCam filters and red in the long-wavelength NIRCam filters. The colour stamps show F150W in blue, F277W in green and F444W in red.





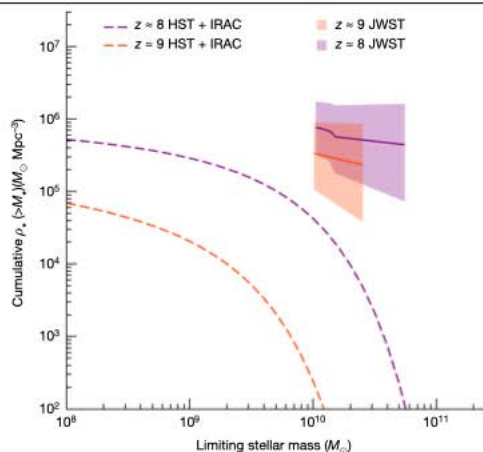
**Fig. 3 | SEDs and stellar population model fits. a**, Photometry (black squares), best-fitting EAZY models (red lines) and redshift probability distribution  $P(z)$  (grey filled histograms) of six galaxies with apparent fiducial masses  $\log(M_*/M_\odot) > 10$ . The flux density units are  $f_\lambda$ . Uncertainties and upper limits (triangles) are  $1\sigma$ . Fiducial best-fit stellar masses and redshifts are noted. The SEDs are characterized by a double break: a Lyman break and an upturn at more than  $3 \mu\text{m}$ . Emission lines are visible in the longest wavelength bands in several cases. **b**, Average rest-frame SED of the six candidate massive galaxies

(red dots) and the 16th–84th percentile of the running median (shaded area). The red line is the best-fit median EAZY model. Green squares and the green line show average rest-frame UV-selected galaxies at  $z = 8, 10$  from HST + Spitzer<sup>31,35</sup>. Grey triangles show two spectroscopically confirmed galaxies at  $z$  of roughly 9 (refs. 23, 34, 35). The double-break selected galaxies are notably redder than previously identified objects at similar redshifts. This may be due to high  $M/L$  ratios or effects that are not included in our modelling, such as AGN or exotic lines.

density (integrated down to  $M_* = 10^8 M_\odot$ ) by a factor of about two at  $z$  of roughly 8 and by a factor of roughly five at  $z$  of roughly 9. Unless the low-mass samples are highly incomplete, the implication would be that most of the total stellar mass at  $z = 8$ –9 resides in the most massive galaxies. Although extreme, this is qualitatively consistent with the notion that the central regions of present-day massive elliptical galaxies host the oldest stars in the universe (together with globular clusters), and with the finding that by  $z$  of roughly 2 the stars in the central regions of massive galaxies already make up 10–20% of the total stellar mass density at that redshift<sup>30</sup>. A more fundamental issue is that these stellar mass densities are difficult to realize in a standard  $\Lambda$  cold dark matter cosmology, as pointed out by several

recent studies<sup>31,32</sup>. Our fiducial mass densities push against the limit set by the number of available baryons in the most massive dark matter halos.

The other extreme interpretation is that all the fiducial masses are larger than the true masses by factors of more than 10–100. We use standard techniques and multiple methods to estimate the masses. Under certain assumptions for the dust attenuation law and stellar population age sampling (favouring young ages with strong emission lines), low masses can be produced (Methods). This only occurs at specific redshifts ( $z = 5.6, 6.9, 7.7$  or about 10% of the redshift range of the sample) in which line-dominated and continuum-dominated models produce similar F410M–F444W colours. In addition, it is possible



**Fig. 4 | Cumulative stellar mass density, if the fiducial masses of the JWST-selected red galaxies are confirmed.** The solid symbols show the total mass density in two redshift bins,  $7 < z < 8.5$  and  $8.5 < z < 10$ , based on the three most massive galaxies in each bin. Uncertainties reflect Poisson statistics and cosmic variance. The dashed lines are derived from Schechter fits to UV-selected samples<sup>1</sup>. The JWST-selected galaxies would greatly exceed the mass densities of massive galaxies that were expected at these redshifts on the basis of previous studies. This indicates that these studies were highly incomplete or that the fiducial masses are overestimated by a large factor.

that techniques that have been calibrated with lower redshift objects<sup>17</sup> are not applicable. As an example, we do not include effects of exotic emission lines or bright active galactic nuclei (AGN)<sup>14</sup>. Part the sample is reported to be resolved in F200W (ref. 33) making a significant contribution from AGN less likely, but faint, red AGN are possible and would be highly interesting in their own right, even if they could lead to changes in the masses.

It is perhaps most likely that the situation is in between these extremes, with some of the red colours reflecting exotic effects or AGN and others reflecting high M/L ratios. Future JWST NIRSpec spectroscopy can be used to measure accurate redshifts as well as the precise contributions of emission lines and to the observed photometry. With deeper data, the stellar continuum emission can be detected directly for the brightest galaxies. Finally, dynamical masses are needed to test the hypothesis that our description of massive halo assembly in  $\Lambda$  cold dark matter is incomplete. It may be possible to measure the required kinematics with ALMA or from rotation curves with the NIRSpec integral field unit if the ionized gas is spatially extended<sup>30,31</sup>.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-05786-2>.

- Stefanon, M. et al. Galaxy stellar mass functions from  $z=10$  to  $z=6$  using the deepest Spitzer/infrared array camera data: no significant evolution in the stellar-to-halo mass ratio of galaxies in the first gigayear of cosmic time. *Astrophys. J.* **922**, 29 (2021).
- Brammer, G. & Maltari, J. *gbrammer/igrizli*: Release 2021. Zenodo (<https://zenodo.org/record/7767790>). ZCRGM-zMxY (2021).
- Scheerer, D. & de Barros, S. The impact of nebular emission on the ages of  $z=6$  galaxies. *Astron. Astrophys.* **502**, 423–426 (2009).
- Shim, H. et al.  $z=4$  H $\alpha$  emitters in the great observatories origins deep survey: tracing the dominant mode for growth of galaxies. *Astrophys. J.* **738**, 69 (2011).
- Labbé, I. et al. The spectral energy distributions of  $z=8$  galaxies from the IRAC ultra deep fields: emission lines, stellar masses, and specific star formation rates at 650 Myr. *Astrophys. J.* **777**, L19 (2013).
- Stark, D. P. et al. Keck spectroscopy of 3. *Astrophys. J.* **763**, 129 (2013).
- Smit, R. et al. Evidence for ubiquitous high-equivalent-width nebular emission in  $z=7$  galaxies: toward a clean measurement of the specific star-formation rate using a sample of bright, magnified galaxies. *Astrophys. J.* **784**, 58 (2014).
- Smit, R. et al. High-precision photometric redshifts from Spitzer/IRAC: extreme [3.6]–[4.5] colors identify galaxies in the redshift range  $z=6.6$ – $6.9$ . *Astrophys. J.* **801**, 122 (2015).
- Faist, A. L. et al. A coherent study of emission lines from broadband photometry: specific star formation rates and [O III]/H $\beta$  ratio at  $z < 6$ . *Astrophys. J.* **821**, 122 (2016).
- de Barros, S. et al. The GREATS H $\beta$ + [O III] luminosity function and galaxy properties at  $z=8$ : walking the way of JWST. *Mon. Not. R. Astron. Soc.* **489**, 2355–2366 (2019).
- Naidu, R. P. et al. Two remarkably luminous galaxy candidates at  $z=10$ – $12$  revealed by JWST. *Astrophys. J.* **940**, L14 (2022).
- Endsley, R. et al. A JWST/NIRCam study of key contributors to reionization: the star-forming and ionizing properties of UV-faint  $z=7$ – $8$  galaxies. Preprint at <https://arxiv.org/abs/2208.14999> (2022).
- Stefanon, M. et al. Blue rest-frame UV-optical colors in  $z=8$  galaxies from GREATs: very young stellar populations at 650 Myr of cosmic time. *Astrophys. J.* **927**, 48 (2022).
- Brammer, G. B., van Dokkum, P. G. & Coppi, P. EAZY: a fast, public photometric redshift code. *Astrophys. J.* **686**, 1503–1513 (2008).
- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G. & Byler, N. Deriving physical properties from broadband photometry with Prospector: description of the model and a demonstration of its accuracy using 129 galaxies in the local universe. *Astrophys. J.* **837**, 170 (2017).
- Smit, R. et al. Inferred H $\alpha$  flux as a star formation rate indicator at  $z=4$ – $5$ : implications for dust properties, burstiness, and the  $z=4$ – $8$  star formation rate functions. *Astrophys. J.* **833**, 254 (2016).
- Barro, G. et al. CANDELS: the progenitors of compact quiescent galaxies at  $z=2$ . *Astrophys. J.* **765**, 104 (2013).
- Roberts-Borsani, G. et al. Interpreting the Spitzer/IRAC colours of  $7 < z < 9$  galaxies: distinguishing between line emission and starlight using ALMA. *Astrophys. J.* **497**, 3440 (2020).
- Laporte, N. et al. Probing cosmic dawn: ages and star formation histories of candidate  $z > 9$  galaxies. *Mon. Not. R. Astron. Soc.* **505**, 3336–3346 (2021).
- Topping, M. W. et al. The ALMA REBELS survey: specific star formation rates in the reionization era. *Mon. Not. R. Astron. Soc.* **516**, 975–991 (2022).
- Whitler, L. et al. Star formation histories of UV-luminous galaxies at  $z \approx 6.8$ : implications for stellar mass assembly at early cosmic times. *Mon. Not. R. Astron. Soc.* **519**, 5859–5881 (2023).
- Hashimoto, T. et al. The onset of star formation 250 million years after the Big Bang. *Nature* **557**, 392 (2018).
- Roberts-Borsani, G. et al. Improving  $z=7$ – $11$  galaxy property estimates with JWST/NIRCam medium-band photometry. *Astrophys. J.* **910**, 86 (2021).
- Duncan, K. et al. The mass evolution of the first galaxies: stellar mass functions and star formation rates at  $z < 7$  in the CANDELS GOODS-South field. *Mon. Not. R. Astron. Soc.* **444**, 2960–2984 (2014).
- Davidson, J. et al. The COSMOS2015 galaxy stellar mass function: thirteen billion years of stellar mass assembly in ten snapshots. *Astron. Astrophys.* **605**, A70 (2017).
- Song, M. et al. The evolution of the galaxy stellar mass function at  $z=4$ – $8$ : a steepening low-mass end slope with increasing redshift. *Astrophys. J.* **825**, 5 (2016).
- Roberts-Borsani, G. W. et al.  $z \geq 7$  Galaxies with red spitzer/IRAC [3.6]–[4.5] colors in the full CANDELS data set: the brightest-known galaxies at  $z=7$ – $9$  and a probable spectroscopic confirmation at  $z=7.48$ . *Astrophys. J.* **823**, 143 (2016).
- van Dokkum, P. G. et al. Dense cores in galaxies out to  $z=2.5$  in SDSS, UltraVISTA, and the five 3D-HST/CANDELS fields. *Astrophys. J.* **791**, 45 (2014).
- Boylan-Kolchin, M. Stress testing  $\Lambda$ CDM with high-redshift galaxy candidates. Preprint at <https://arxiv.org/abs/2208.01611> (2022).
- Menci, N. et al. High-redshift galaxies from early JWST observations: constraints on dark energy models. *Astrophys. J.* **938**, L5 (2022).
- Finkelstein, S. et al. CEERS key paper I: an early look into the first 500 Myr of galaxy formation with JWST. Preprint at <https://arxiv.org/abs/2211.05792> (2022).
- Astropy Collaboration. The Astropy Project: sustaining and growing a community-oriented open-source project and the latest major release (v5.0) of the core package. *Astrophys. J.* **935**, 167 (2022).
- Treu, T. et al. The initial mass function of early-type galaxies. *Astrophys. J.* **709**, 1195 (2010).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

- Glazebrook, K. et al. A massive, quiescent galaxy at a redshift of 3.717. *Nature* **544**, 71–74 (2017).
- Riechers, D. A. et al. Rise of the Titans: gas excitation and feedback in a binary hyperluminous dusty starburst galaxy at  $z=6$ . *Astrophys. J.* **907**, 62 (2021).

### Observations, reduction and photometry

This article is based on the first imaging taken with the NIRCcam on JWST as part of the CEERS program (Principal Investigator, Finkelstein; Program Identifier, 1345). Four pointings have been obtained, covering roughly 38 arcmin<sup>2</sup> on the Extended Groth Strip HST legacy field and overlapping fully with the existing HST–Advanced Camera for Surveys (ACS) and WFC3 footprint. NIRCcam observations were taken in six broad-band filters, F115W, F200W, F150W, F227W, F356W and F444W, and one medium bandwidth filter F410M. The F410M medium band sits within the F444W filter and is a sensitive tracer of emission lines, enabling improved photometric redshifts and stellar mass estimates of high-redshift galaxies<sup>49</sup>.

Exposures produced by stage 2 of the JWST calibration pipeline (v.1.5.2) were downloaded from the MAST archive. The data reduction pipeline Grism redshift and line analysis software for space-based spectroscopy (Grizli<sup>50</sup>) was used to process, align and co-add the exposures. The pipeline mitigates various artefacts, such as ‘snow-balls’ and 1/f noise. To improve pixel-to-pixel variation, custom flat-field calibration images (<https://s3.amazonaws.com/grizli-v2/Nircam-Skyflats/flats.html>) were created from on-sky commissioning data (program COM-1063) that are the median of the source-masked and background-normalized exposures in each NIRCcam detector.

The pipeline then subtracts a large-scale sky background, aligns the images to stars from the Gaia DR3 catalogue and drizzles the images to a common pixel grid using *astrodizzle*. The mosaics are available online as part of the v.3 imaging data release (<https://s3.amazonaws.com/grizli-v2/jwstMosaics/v3/index.html>). Existing multi-wavelength ACS and WFC3 archival imaging from HST were also processed with Grizli. For the analysis in this paper all images are projected to a common 40 mas pixel grid. Remaining background structure in the NIRCcam mosaics is due to scattered light. The background is generally smooth on small scales and was effectively removed with a 5-inch median filter after masking bright sources.

We use standard astrophysics<sup>34</sup> and photutils<sup>36</sup> procedures to detect sources, create segmentation maps and perform photometry. The procedures are like those used in previous ground- and space-based imaging surveys. Briefly, we create an inverse variance weighted combined F227W + F356W + F444W image and detect sources after convolution with a Gaussian of 3 pixels full-width at half-maximum (0.12″) to enhance sensitivity for point sources. Point spread functions were matched to the F444W band using photutils procedures. Photometry was performed at the locations of detected sources in all filters using 0.32-inch-diameter circular apertures. The fluxes were corrected to total using the Kron autoscaling aperture measurement on the detection image. A second small correction was applied for light outside the aperture based on the encircled energy provided by the WebbPSF software. The final catalogue contains 42,729 sources and includes all available HST–ACS and JWST–NIRCcam filters (ten bands, spanning 0.43 to 4.4 μm). Photometry for HST–WFC3 bands was also derived, but only used for zeropoint testing as the HST–WFC3 images are several magnitudes shallower than NIRCcam.

### Photometric zero-points

The first JWST images were released with preflight zero-points for the NIRCcam filters. The preflight estimates do not match the in-flight performance, with errors up to roughly 20% in the long-wavelength bands. This analysis uses updated in-flight calibrations that were provided by the Space Telescope Science Institute on 29 July 2022 (jwst\_0942\_pmap) on the basis of observations of two standard stars. The calibrations improved the accuracy of the long-wavelength photometry but introduced errors in the short-wavelength bands, with variations up to 20% between detectors, as determined from comparisons to previous HST–WFC3 photometry and analyses of stars in the Large Magellanic Cloud and the globular cluster M92 (refs. 37,38).

We derived new zero-points for all short-wavelength and long-wavelength bands, for both NIRCcam modules, using two independent methods. The first method (‘GB’) uses zero-points that are based on standard stars observed by JWST in the B module and transferred to the A module using overlapping stars in the Large Magellanic Cloud. The second method (‘IL’) uses 5,000–10,000 galaxies at photometric redshifts  $0.1 < z < 5$  with a signal to noise ratio (SNR) greater than 15 from the CEERS parent catalogue and calculates the ratio between observed and EAZY model fluxes for each detector, module and photometric band. As the observed wavelengths sample different rest-frame parts of the SEDs of the galaxies, errors in the model fits can be separated from errors in the zero-points. More information on the methodology and the resulting zero-points are provided on github (<https://github.com/gbrammer/grizli/pull/107>).

The methods agree very well, with differences of  $3 \pm 3\%$  in all bands except F444W, where we find a difference of 8%. We use the GB values for all bands except F444W where we take the average of the GB and IL values (multiplicative corrections 1.064 for module A and 1.084 for module B). The photometry is listed in Extended Data Table 1. Using the fiducial zero-points, Extended Data Fig. 1 shows offsets with respect to EAZY model fluxes, split by detector, module and filter, showing only 0–3% residuals. A third independent method used colour-magnitude diagrams of stars in M92 (refs. 37,38) in F090W, F150W, F227W and F444W bands, with reported consistency with the GB values within the uncertainties. Our adopted zero-points agree with the most recent NIRCcam flux calibration (jwst\_0989\_pmap, October 2022) to within 4%. This paper adopts a 5% minimum systematic error (added in quadrature) for all photometric redshift and stellar population fits to account for calibration uncertainties. Finally, we compiled a sample of 450 galaxies with spectroscopic redshifts  $0.2 < z < 3.8$  from 3D-HST (ref. 39) and MOSDEF<sup>40</sup> to test photometric redshift performance, finding a normalized median absolute deviation of  $(z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}}) = 2.5\%$ .

### Sample selection

The JWST–NIRCcam imaging in this paper reaches 5σ depths from 28.5 to 29.5 AB, representing an order of magnitude increase in sensitivity and resolution beyond wavelengths of 2.0 μm and allowing us to select galaxies at rest-frame optical wavelengths to  $z$  of roughly 10. To enable straightforward model-independent reproduction of the sample we use a purely empirical selection of high-redshift galaxies based on NIRCcam photometry, rather than one on inferred photometric redshift or stellar mass. We select on a ‘double-break’ SED: no detection in the HST–ACS optical, blue in the NIRCcam short-wavelength filters and red in the NIRCcam long-wavelength filters, which is expected for sources at  $z \geq 7$  with Lyman break and with red UV-optical colours.

The following colour selection criteria were applied:

$$F150W - F227W < 0.7$$

$$F227W - F444W > 1.0$$

in addition to a non-detection requirement in HST–ACS imaging.

$$\text{SNR}(B_{435}, V_{606}, I_{814}) < 2$$

To ensure good SNR, we limit our sample to  $F444W < 27$  AB magnitude and  $F150W < 29$  AB magnitude and require  $\text{SNR}(F444W) > 8$ . We manually inspected selected sources and removed a small number of artefacts, such as hot pixels, diffraction spikes and sources affected by residual background issues or bright neighbours.

This selection complements the traditional drop-out colour selection techniques based on isolating the strong Lyman 1,216 Å break as it moves through the filters. Drop-out selection is not feasible here: the HST–ACS data are not deep enough to select drop-out galaxies to the same equivalent limits as the NIRCcam imaging. Screening for two breaks



has shown to be an effective redshift selection: a similar technique was used to successfully select bright galaxies at  $7 < z < 9$  from wide-field HST and Spitzer data<sup>28</sup>. A red F277W–F444W colour can be produced by large amounts of reddening by dust, evolved stellar populations with a Balmer Break<sup>24</sup>, strong optical emission lines<sup>50</sup> or a combination of these.

This selection produced a total of 13 sources, with a median SNR ratio in the F444W band of roughly 30 (see Fig. 2 and Extended Data Fig. 2). The resulting sample is dark at optical wavelengths (2 $\sigma$  upper limit of  $I_{814} > 30.4$  AB) and faint in F115W and F150W with median of roughly 28 AB magnitude, beyond the limits reached with HST–WFC3 except in small areas in the Hubble Ultra-Deep Field and the Frontier Fields. The absence of any flux in the ACS optical, the red  $I_{814} - F115W > 2.5$  and blue  $F115W - F150W$  of roughly 0.3 AB colours are consistent with a strong Lyman break moving beyond the ACS  $I_{814}$  band at redshifts  $z > 6$ . The NIRC2 F444W magnitudes are bright at around 26 AB, and the median  $F150W - F444W$  of roughly 2 AB colour is redder than any sample previously reported at  $z > 7$  (refs. 3, 18, 21, 29, 41).

### Fits to the photometry

Several methods are used to derive redshifts and stellar masses, all allowing extremely strong emission lines combined with a wide range of continuum slopes: (1) EAZY with extra templates that include strong emission lines, (2) Prospector with a strongly rising star formation history (SFH) prior that favours young ages, (3) Bagpipes to evaluate dependence on stellar population model assumptions and minimization algorithm. Finally, we also consider (4), a proposed template set for high-redshift galaxies with blue continua, strong emission lines and a non-standard IMF. Throughout, reported uncertainties are the 16th–84th percentiles of the probability distributions. A Salpeter<sup>42</sup> IMF is assumed throughout, for consistency with previous determinations of the high-redshift galaxy mass function<sup>32,43</sup> and constraints on the IMF in the centres of the likely descendants<sup>35,43–45</sup>. A summary of the results is presented in Extended Data Figs. 3 and 4.

**EAZY.** The main benefits of EAZY<sup>5</sup> are ease of use, speed and reproducibility. EAZY fits non-negative linear combinations of templates, with redshift and scaling of each template as free parameters. The allowed redshift range was 0–20 and no luminosity prior was applied.

The standard EAZY template set (tweak\_fpsp, QSF\_12\_v3) is optimized for lower redshift galaxies. High-redshift stellar populations tend to be younger, less dusty and have stronger emission lines. We create a more appropriate template set by removing the oldest and dustiest templates ( $A_V > 2.5$ ) from the standard set, keeping templates 1, 2, 7, 8, 9, 10 and 11, and adding two flexible stellar population synthesis (FPS) templates with strong emission lines. The first has a continuum that is roughly constant in  $F_\lambda$  with  $EW(H\beta+[OIII]) = 650$  Å, similar to NIRSpc-confirmed galaxies<sup>46</sup> at  $z = 7–8$ . The second has a red continuum that is constant in  $F_\lambda$  with  $EW(H\beta+[OIII]) = 1,100$  Å, comparable to line strengths inferred for bright Lyman-break galaxies at  $z = 7–9$  (ref. 29). Each template has an associated M/L ratio, so the template weights in the fit can be converted to a total stellar mass. We fit all galaxies in the catalogue with the default EAZY template set first and then refit all galaxies at  $z > 7$  using the new template set. The template set is available online with the photometric catalogue (<https://github.com/ivolabbe/red-massive-candidates>). The EAZY redshift distribution of the sample of 13 galaxies is  $7.3 < z < 9.4$ , with no low-redshift interlopers ( $z < 6$ ). EAZY masses range  $9.2 < \log(M_*/M_\odot) < 10.9$ .

**Prospector.** We perform a stellar population fit with more freedom than is possible in EAZY using the Prospector<sup>17,47</sup> framework, specifically the Prospector- $\alpha$  settings<sup>48</sup> and the mesa isochrones and stellar tracks stellar isochrones<sup>49,50</sup> from FPS<sup>51,52</sup>. This mode includes non-parametric star formation histories, with a continuity prior that disfavors large changes in the star formation rate between time bins<sup>53</sup>. It uses a two-component, age-dependent dust model, allows

full freedom for the gas-phase and stellar metallicity, and includes nebular emission in which the nebulae are self-consistently powered by the stellar ionizing continuum from the model<sup>54</sup>. The sampling was performed using the dynesty<sup>55</sup> nested sampling algorithm. We also adopt two new priors that disfavour high-mass solutions: first, a mass function prior on the stellar mass, adopting the observed  $z = 3$  mass function for  $z > 3$  solutions<sup>56</sup>, and second, a non-parametric SFH prior that favours rising SFHs in the early universe and falling SFHs in the late universe, following expectations from the cosmic star formation rate density. These are described in detail in ref. 57.

The masses from Prospector are consistent within the uncertainties with the EAZY masses, with a mean offset of  $\log(M_{\text{Prospector}}/M_{\text{EAZY}}) = 0.1$  for objects with greater than  $10^{10} M_\odot$ . The most massive objects as indicated by EAZY are also the most massive in the Prospector fits. Prospector also provides ages and star formation rates. The star formation rates are generally not very well constrained in the fits due to the lack of infrared coverage. The ages are also uncertain and depend strongly on the adopted prior. For a constant SFH prior Prospector finds typical ages of roughly 0.3 Gyr, with substantial Balmer breaks, whereas for strongly rising SFHs Prospector finds a median mass-weighted age of 34 Myr, with strong emission lines and large amounts of reddening ( $A_V \approx 1.5$ ). This is reminiscent of the age-dust degeneracy that is well known at lower redshift. The stellar masses do not vary significantly between these two priors. The red SEDs (Fig. 3) require high M/L ratios for a large range of the best-fit stellar population ages, as is well known from studies of nearby galaxies<sup>58</sup>.

**Bagpipes.** Fits with the Bayesian analysis of galaxies for physical inference and parameter estimation (Bagpipes<sup>39</sup>) software are also considered. Compared to Prospector, Bagpipes uses the Bruzual and Charlot stellar population models<sup>60</sup> and sampling algorithm Multinest<sup>61</sup>. Whereas Bagpipes does not cover new parameter space compared to Prospector, it allows us to evaluate how sensitive the masses are to the adopted stellar population model or fitting technique. Furthermore, Bagpipes is relatively fast, so we can use it to explore the effect of modeling assumptions to investigate the role of systematic uncertainties on the derived redshift and stellar mass. We focus on attenuation law, SFH, age sampling priors and SNR.

(1) Bagpipes\_csf\_salim: baseline model of constant SFH with redshift 0 to 20, age\_max from 1 Myr to 10 Gyr, metallicity between 0.01 and 2.5 Solar, ionization parameter  $-4 < \log(U) < -2$ , a Salim<sup>62</sup> attenuation  $0 < A_v < 4$  and adopting a linear prior in age and log prior in metallicity and ionization and uniform prior in redshift, age and  $A_v$ . The Salim law varies between a steep Small Magellanic Cloud (SMC)-like extinction law at low optical depth and a flat Calzetti-like dust law at large optical depth, in accordance with empirical studies<sup>62</sup> and theoretical expectations<sup>63</sup>. The Bagpipes masses and redshifts are similar on average to those of EAZY and Prospector, with a mean offset of  $\log(M_*/M_{\text{EAZY}}) = 0$  for the massive sample.

(2) Bagpipes\_rising\_salim: this model is not intended to search for best fit in a wide parameter space but only in a restricted space to increase the emission-line contribution to the redder filter, F444W, and decrease the stellar masses. The model is restricted to rising star formation rates at high redshift (delayed  $\tau > 0.5$  Gyr) and redshifts to  $z < 9.0$  to force the H $\beta$ + $[OIII]$  complex to fall within the F444W filter. The fits show strong emission lines, low ages (median roughly 30 Myr) and high dust content (median  $A_V \approx 1.7$ ). Even with these restrictions, the mean stellar mass agrees well with the baseline (mean  $\log(M_*/M_\odot) = -0.1$  for objects with greater than  $10^{10} M_\odot$ ).

(3) Bagpipes\_csf\_salim\_logage: like the model in (1) but with a logarithmic age prior, which is heavily weighted towards very young ages. For the five reddest, most massive galaxies in (1) the results are unchanged, whereas six other galaxies are now placed at significantly lower masses (inconsistent with model (1), given the uncertainties), including 14,924 (from  $\log(M_*/M_\odot) = 10.1$  to 8.7). The  $P(z)$  of these

lower mass solutions is narrow and clustered in narrow spikes at  $z = 5.6, 6.9, 7.7$ , where the F410M filter cannot distinguish between strong lines and continuum SEDs (Extended Data Figs. 5 and 6).

(4) *Bagpipes\_csf\_salim\_logage\_snr10*: to test if the fit in (3) is driven by the high SNR in long-wavelength filters (which put all the weight in the fits there), we impose an error floor of 10% on the photometry that roughly balances the SNR across all NIRCcam bands. As JWST is still in early days of calibration, some limit on SNR is prudent. The SNR-limited fits result in high-mass solutions for 11 out of 13 galaxies. Notably, the uncertainties on the stellar mass do not encompass the low-mass solution from (3) indicating that detailed assumptions on the treatment of SNR can introduce systematic changes.

(5) *Bagpipes\_csf\_smc\_logage*: SMC extinction is often used in modelling high-redshift galaxies<sup>14</sup>. Our *Bagpipes* modelling use Salim-type dust that includes the SMC-like extinction at low optical depth, but it is useful to evaluate fits that are restricted to a steep extinction law in combination with a logarithmic age prior favouring young ages. The results are different from any of the modelling above: ten out of 13 galaxies show very low stellar masses (in the range  $10^8 M_\odot - 10^9 M_\odot$ ) in combination with extremely young ages (1–5 Myr). Another notable aspect is that these fits do not match the blue part of the SED well (NIRCcam short-wavelength F115W, F150W, F200W) and the fits seem driven by the high SNR in the NIRCcam long-wavelength filters (Extended Data Fig. 3). Most fits have significantly worse  $\chi^2$  than the high-mass fits (EAZY, Prospector, *Bagpipes* (1)–(4)).

In conclusion, the derived masses depend on assumed attenuation law, parameterization of ages and treatment of photometric uncertainties. Together, these aspects can produce lower redshifts and lower masses by up to factors of 100 in ways that are not reflected by the random uncertainties. Therefore, different assumptions can change the stellar masses and redshifts systematically and the uncertainties are probably underestimated.

Although neither high nor low-mass models can be excluded with the data available now, there are two features that would suggest the ultra-young, low-mass solutions are less plausible. First, whereas 1–5 Myr ages are formally allowed, the galaxy would not be causally connected:  $10^3 M_\odot$  of star formation would have started spontaneously on timescales less than a dynamical time (although dynamical times are uncertain until velocity dispersions and corresponding sizes are measured). In addition, the probability of catching most galaxies at that precise moment is low, given the roughly 200 Myr search window at  $z = 7-9$ . It would suggest there are more than 40 older and more massive galaxies for every galaxy in our sample.

Second, the  $P(z)$  of the low-mass solutions are extremely narrow and concentrated at nearly discrete redshifts  $z = 5.6, 6.9$  and  $7.7$  (for example,  $38094\ z = 6.93 \pm 0.01$ ). Here strong H $\alpha$  and H $\beta$ +OIII transition between the overlapping F356W, F410M and F444W filter edges (Extended Data Fig. 5). A single line can contribute to several bands (for example, [OIII]5007 at  $z = 6.9$ ), with great flexibility due to the rapidly varying transmission at the filter edges. The result is that line and continuum-dominated models are degenerate due to undersampling of the SED and resulting aliasing, but only at specific redshifts.

Although finding one 5 Myr galaxy exactly in this narrow window could be luck, we find that ten out of 13 galaxies can only be fit with low-mass, ultra-young models at these discrete redshifts  $z = 5.6, 6.9$  and  $7.7$ . Such an age and  $P(z)$  distribution for the sample, at precisely the redshifts where this fortuitous overlap between filters occurs (roughly less than 8% of the redshift range between  $z = 5-9$ ), is not implausible. To rule out that the spiked nature of the  $P(z)$  is the result of our double-break selection, we perform simple simulations. We take random draws from the posteriors of line-dominated model E, redshift the models to a uniform distribution between four and ten, perturb with the observational errors and apply our double-break selection criterion to the simulated photometry (Extended Data Fig. 6). This suggests that even if the sample were line dominated with ages less than 5 Myr, the redshift

distribution should be different (not spiked) suggesting that these fits suffer from aliasing. By contrast,  $P(z)$  of high-mass model B is broadly self-consistent with the selection function based on the model B fits.

The likely reason that this effect primarily occurs with an SMC extinction law is because of the strong wavelength dependence (steep in the far ultraviolet, flatter in optical). For the sample in this paper, fits with SMC have difficulty reproducing the overall (rest-optical) red SED shape. This can be clearly seen in Extended Data Fig. 3, where the SMC-based fits have strongly ‘curved’ continua, which are generally too steep in the rest-UV and too flat in the rest-optical (F356W, F410W, F444W bands), requiring strong emission lines at specific redshifts to produce the red colours.

**FSPS-hot model.** For completeness we also consider recently proposed ‘fspd-hot’ models<sup>64</sup>, which consist of templates with blue continua, strong emission lines and with a modified extremely bottom-light IMF that produces lower masses. Such an IMF is proposed to be appropriate for the extreme conditions that might be expected in high-redshift galaxies. For ten out of 13 galaxies (including all massive, greater than  $10^{10} M_\odot$  sources), the fspd-hot template set provides poorer fits to the photometry than the fspd-wulturecorn set (median  $\Delta\chi^2 = 31$ ), due to the lack of red templates. The fspd-hot set places nine out of 13 galaxies in a narrow redshift range  $z = 7.7$  with very small uncertainties  $\sigma(z) = 0.05$ , reminiscent of the spiked distribution found earlier for *Bagpipes* model (5). The blue template set can only produce red colours if strong emission lines are placed at specific redshifts. Because the fits are poor overall and no further insight is gained, we do not consider these masses further to avoid confusion due to adopting vastly different IMFs. The extremely bottom-light IMF, with suppression of (invisible) low-mass stars, is untestable with photometric data.

## Fiducial redshifts and stellar masses

The results of all methods are shown graphically in Extended Data Fig. 4. Most methods explored produce good fits and consistent masses and redshifts. Rather than favour one method over the others we derive fiducial masses and redshifts for each object by taking the median values of the EAZY (1), Prospector (2), the five *Bagpipes* fits (3–7) results from each galaxy. As discussed in the main text, the consistency between various methods may largely indicate a consistency in underlying assumptions. Different assumptions can change the stellar masses and redshifts systematically in ways that are not reflected by the random uncertainties. The fiducial redshifts and masses are listed in Extended Data Table 2.

Furthermore, we do not consider contributions from exotic emission-line species nor include AGN templates in the fits<sup>14</sup>. All objects in this paper should be considered candidate massive galaxies, to be confirmed with spectroscopy.

## Lensing

A potential concern is that the fluxes (and therefore the masses) of some or all the galaxies are boosted by gravitational lensing. No galaxy is close to the expected Einstein radius of another object. The bright galaxy that is 1.2° to the southwest of 38094 has  $z_{\text{grism}} \approx 1.15$  and  $M_* \approx 10.63$  (object number 28717 in the 3D-HST AEGIS catalogue<sup>23</sup>), and an Einstein radius (roughly 0.4") that is 0.3× the distance to 38094. If we assume that the mass profile of the lensing galaxy is an isothermal sphere, then the magnification is  $1/(1 - \theta_s/\theta_E)$  where  $\theta_s$  is the separation from the foreground source and  $\theta_E$  is the Einstein radius. This would indicate a relatively modest −0.15 dex correction to the stellar mass. We apply this correction when calculating densities in Fig. 4.

## Volume

Stellar mass densities for galaxies with  $M_* > 10^{10} M_\odot$  are calculated by grouping the galaxies in two broad redshift bins ( $7 < z < 8.5$  and  $8.5 < z < 10$ ). At  $z$  of roughly 8.5, the Lyman Break moves through the

F115W filter, allowing galaxies to be separated into the two bins. The cosmic volume is estimated by integrating between the redshift limits over 38 arcmin<sup>2</sup>, making no corrections for contamination or incompleteness. The key result is driven by the most massive galaxies. Any incompleteness would increase the derived stellar mass densities, whereas contamination would decrease it. Cosmic variance is about 30%, calculated using a web calculator<sup>5,65</sup>. The error bars on the densities are the quadratic sum of the Poisson uncertainty and cosmic variance, with the Poisson error dominant. The volume estimate is obviously simplistic, but the colour selection function (Extended Data Fig. 6) suggests that most of the sample should lie between  $7 < z < 10$ . A more refined treatment does not seem warranted given that the main (orders of magnitude) uncertainty in our study is the interpretation of the red colours of the galaxies.

## Data availability

The HST data are available in the MAST (<http://archive.stsci.edu>), under program ID 1345. Photometry, EAZY template set, fiducial redshifts and stellar masses of the sources presented here are available at <https://github.com/ivolabbe/red-massive-candidates>.

## Code availability

Publicly available codes and standard data reduction tools in the Python environments were used: Grizli<sup>4</sup>, EAZY<sup>5</sup>, astropy<sup>63</sup>, photutils<sup>64</sup> and Prospector<sup>17,36,37</sup>.

36. Bradley, L. et al. astropy/photutils: 1.5.0. Zenodo <https://doi.org/10.5281/zenodo.6825092> (2022).
37. Boyer, M. L. et al. The JWST resolved stellar populations early release science program I: NIRC2 flux calibration. Preprint at <https://arxiv.org/abs/2209.03348> (2022).
38. Nardiello, D. et al. Photometry and astrometry with JWST-1. NIRC2 point spread functions and the first JWST colour-magnitude diagrams of a globular cluster. Preprint at <https://arxiv.org/abs/2209.06547> (2022).
39. Skelton, R. E. et al. 3D-HST WFC3-selected photometric catalogs in the five CANDELS/3D-HST fields: photometry, photometric redshifts, and stellar masses. *Astrophys. J. Suppl. Ser.* **214**, 24 (2014).
40. Kriek, M. et al. The MOSFIRE Deep Evolution Field (MOSDEF) survey: rest-frame optical spectroscopy for ~1500 H- selected galaxies at  $1.37 < z < 3.8$ . *Astrophys. J. Suppl. Ser.* **218**, 15 (2015).
41. Zitron, A. et al. Lyman- $\alpha$  emission from a luminous  $z \sim 8.68$  galaxy: implications for galaxies as tracers of cosmic reionization. *Astrophys. J.* **810**, L12 (2015).
42. Salpeter, E. The luminosity function and stellar evolution. *Astrophys. J.* **121**, 161 (1955).
43. Cappellari, M. et al. Systematic variation of the stellar initial mass function in early-type galaxies. *Nature* **544**, 485–488 (2012).
44. Conroy, C. & van Dokkum, P. The stellar initial mass function in early-type galaxies from absorption line spectroscopy. II. Results. *Astrophys. J.* **760**, 71 (2012).
45. van Dokkum, P. et al. The stellar initial mass function in early-type galaxies from absorption line spectroscopy. III. Radial gradients. *Astrophys. J.* **841**, 68 (2017).
46. Schaefer, D. et al. First look with JWST spectroscopy: resemblance among  $z \sim 8$  galaxies and local analogs. *Astron. Astrophys.* **665**, L4 (2022).
47. Johnson, B. D., Leja, J., Conroy, C. & Speagle, J. S. Stellar population inference with prospector. *Astrophys. J.* **254**, 22 (2021).
48. Leja, J. et al. An older, more quiescent universe from panchromatic SED fitting of the 3D-HST survey. *Astrophys. J.* **877**, 140 (2019).

49. Choi, J. et al. MESA isochrones and stellar tracks (MIST). I. Solar-scaled models. *Astrophys. J.* **823**, 102 (2016).
50. Dotter, A. MESA isochrones and stellar tracks (MIST) 0: methods for the construction of stellar isochrones. *Astrophys. J. Suppl. Ser.* **222**, 8 (2016).
51. Conroy, C., Gunn, J. E. & White, M. The propagation of uncertainties in stellar population synthesis modeling. I. The relevance of uncertain aspects of stellar evolution and the initial mass function to the derived physical properties of galaxies. *Astrophys. J.* **699**, 486 (2009).
52. Conroy, C. & Gunn, J. E. The propagation of uncertainties in stellar population synthesis modeling. III. Model calibration, comparison, and evaluation. *Astrophys. J.* **712**, 833 (2010).
53. Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C. & Speagle, J. S. How to measure galaxy star formation histories. II. Nonparametric models. *Astrophys. J.* **876**, 3 (2019).
54. Byler, N., Dalcanton, J. J., Conroy, C. & Johnson, B. D. Nebular continuum and line emission in stellar population synthesis models. *Astrophys. J.* **840**, 44 (2017).
55. Speagle, J. DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Mon. Not. R. Astron. Soc.* **493**, 3132–3158 (2020).
56. Leja, J. et al. A new census of the O.2. *Astrophys. J.* **893**, 111 (2020).
57. Wang, B. et al. Inferring more from less: Prospector as a photometric redshift engine in the era of JWST. *Astrophys. J.* **944**, L58 (2023).
58. Bell, E. & de Jong, R. Stellar mass-to-light ratios and the Tully-Fisher relation. *Astrophys. J.* **550**, 1 (2001).
59. Carnall, A. C. et al. Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms. *Mon. Not. R. Astron. Soc.* **480**, 4379–4401 (2018).
60. Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).
61. Feroz, F. et al. Importance nested sampling and the MultiNest algorithm. *Open J. of Astrophys.* **2**, 10 (2019).
62. Salim, S., Boquien, M. & Lee, J. C. Dust attenuation curves in the local universe: demographics and new laws for star-forming galaxies and high-redshift analogs. *Astrophys. J.* **859**, 11 (2018).
63. Chevillard, J. et al. Insights into the content and spatial distribution of dust from the integrated spectral properties of galaxies. *Mon. Not. R. Astron. Soc.* **432**, D61 (2013).
64. Steinhart, C. L. et al. Templates for fitting photometry of ultra-high-redshift galaxies. Preprint at <https://arxiv.org/abs/2208.07879> (2022).
65. Trenti, M. & Stiavelli, M. Cosmic variance and its effect on the luminosity function determination in deep high- $z$  surveys. *Astrophys. J.* **676**, 767–780 (2008).

**Acknowledgements** We are grateful to the CEERS team for providing these exquisite public JWST data so early in the mission. We thank M. Boylan-Kolchin for helpful discussions on the theoretical context of this work. Cloud-based data processing and file storage for this work is provided by the AWS Cloud Credits for Research program. The Cosmic Dawn Center is funded by the Danish National Research Foundation. K.W. wishes to acknowledge funding from Alfred P. Sloan Foundation grant no. FG-2019-12514. M.S. acknowledges project no. PID2019-109592GB-I00/AEI/10.13039/501100011033 from the Spanish Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación.

**Author contributions** I.L. performed the photometry, devised the selection method and led the analysis. P.v.D. drafted the main text. I.L. wrote the Methods section and produced the figures. G.B. developed the image processing pipeline and created the image mosaics. E.N. and R.B. identified the first double-break galaxy, prompting the systematic search for these objects. J.L., B.W., K.A.S. and E.M. ran the Prospector analysis. All authors contributed to the manuscript and aided the analysis and interpretation.

**Competing interests** The authors declare no competing interests.

## Additional information

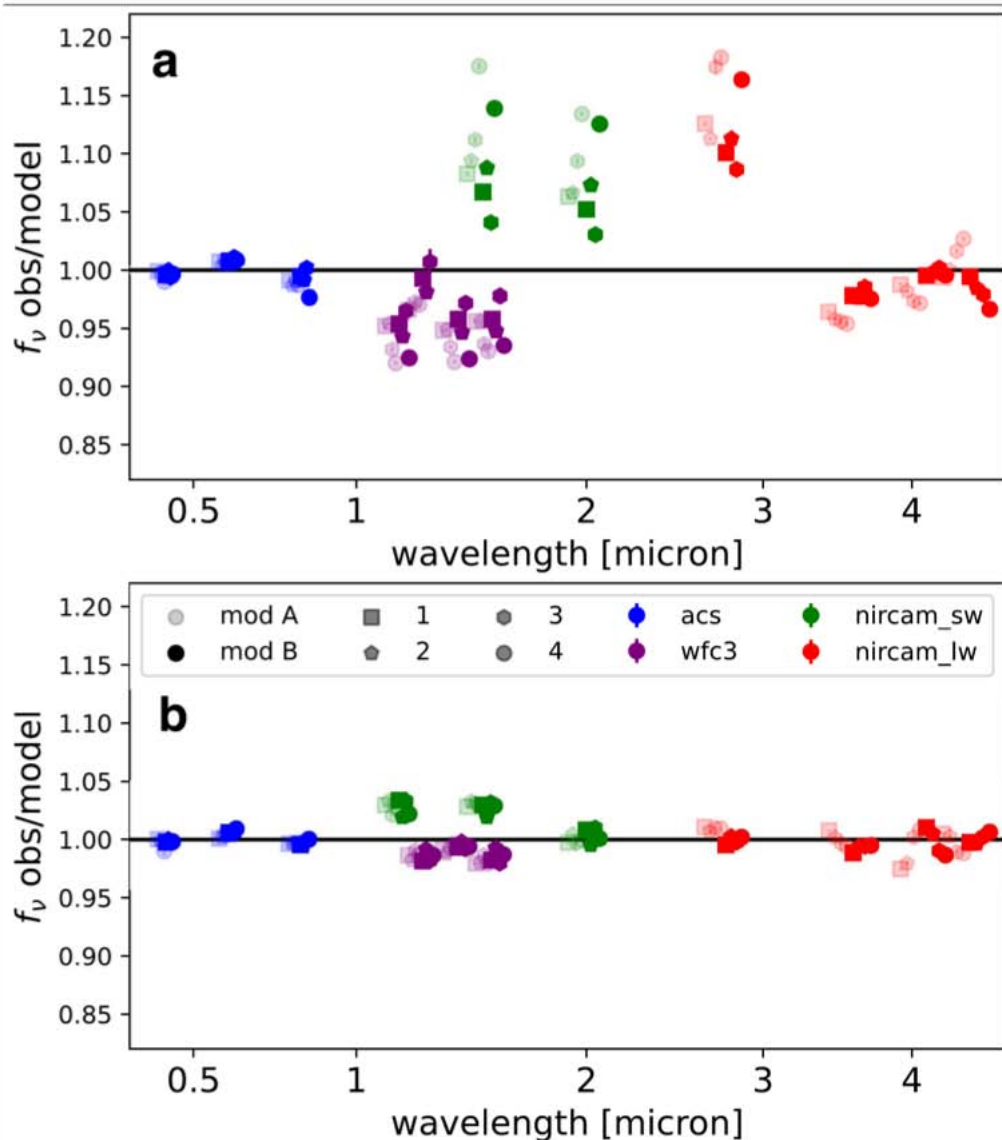
**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-05786-2>.

**Correspondence** and requests for materials should be addressed to Ivo Labbé.

**Peer review information** Nature thanks James Dunlop and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

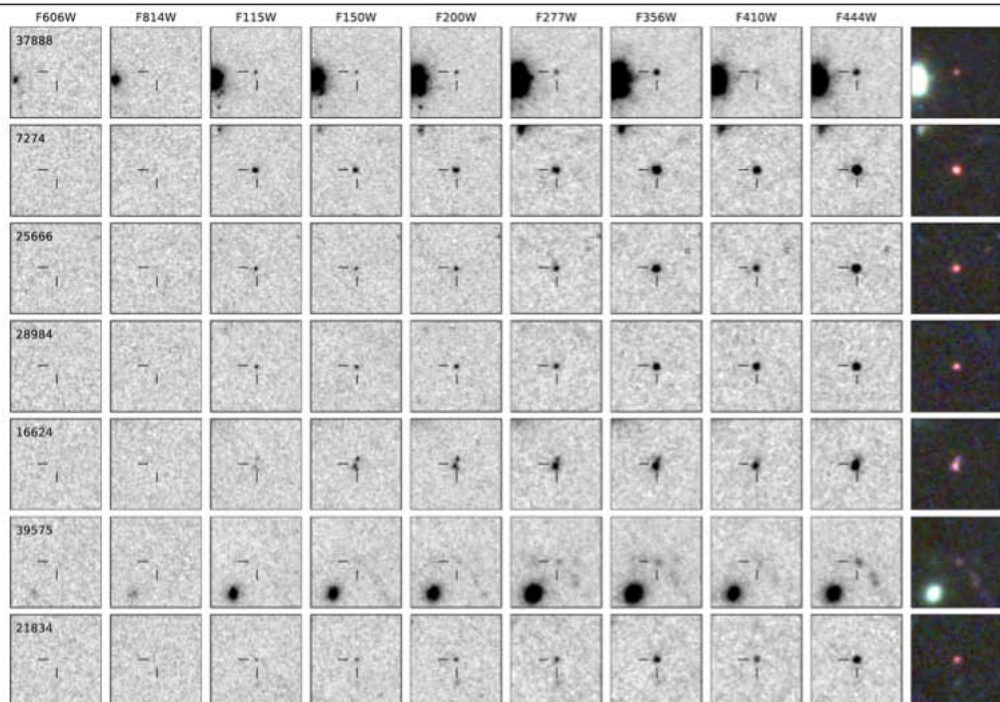
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.





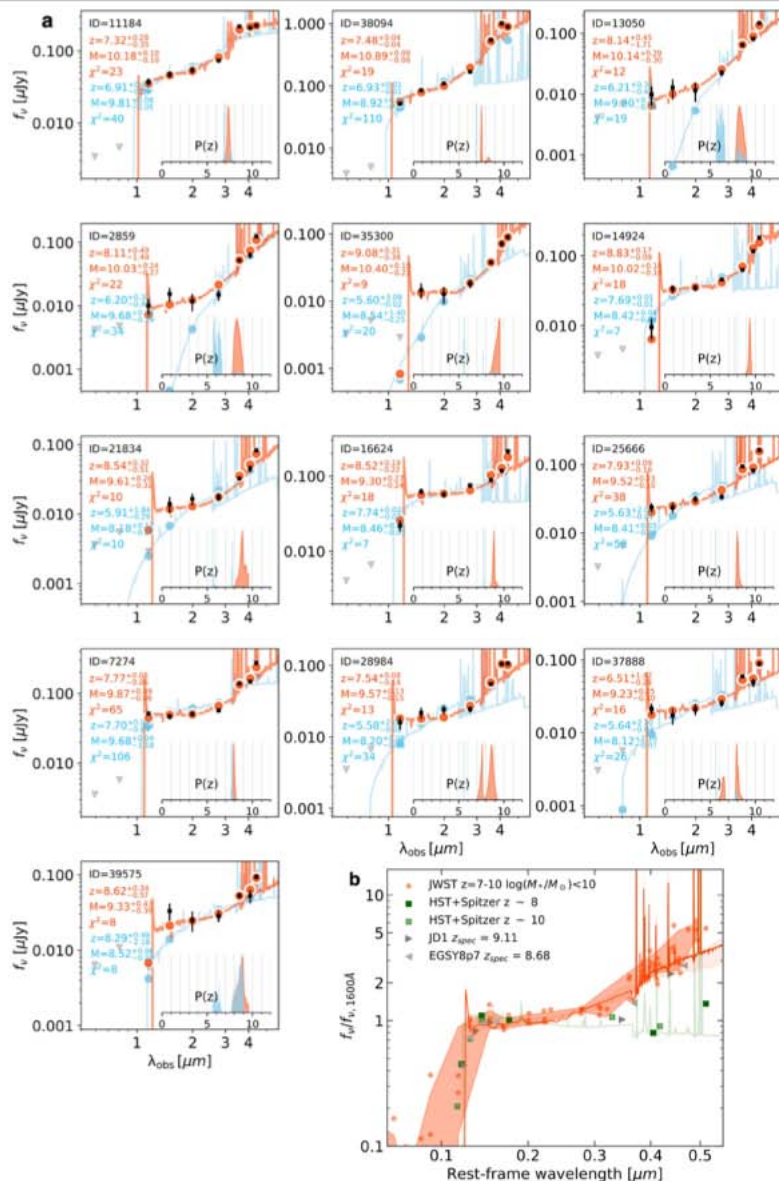
**Extended Data Fig. 1 | Systematic offsets in photometry as a function of wavelength.** The offsets are estimated by the ratio of the observed fluxes to the EAZY best-fit model fluxes for 5,000–10,000 sources at  $0.1 < z < 5$  in the CEERS field. The offsets are calculated separately for each detector (1–4), module (A/B), and filter. Symbols are slightly spread out in wavelength for

clarity. **a.** The first in-flight NIRCам flux calibration update of 29 July 2022 (jwst\_0942.pmap) introduced significant offsets in NIRCам short-wavelength zeropoints. **b.** After adopting our fiducial zeropoints, residual offsets are  $< 3\%$  across all bands. This paper adopts a 5% minimum systematic error for all photometric redshift and stellar population fits.



**Extended Data Fig. 2 | Images of the seven galaxies with apparent lowest mass.** The galaxies satisfy the color-color selection and have fiducial masses  $\log(M_*/M_\odot) < 10$ . The layout and panels of the figure are identical to Fig. 2 in the

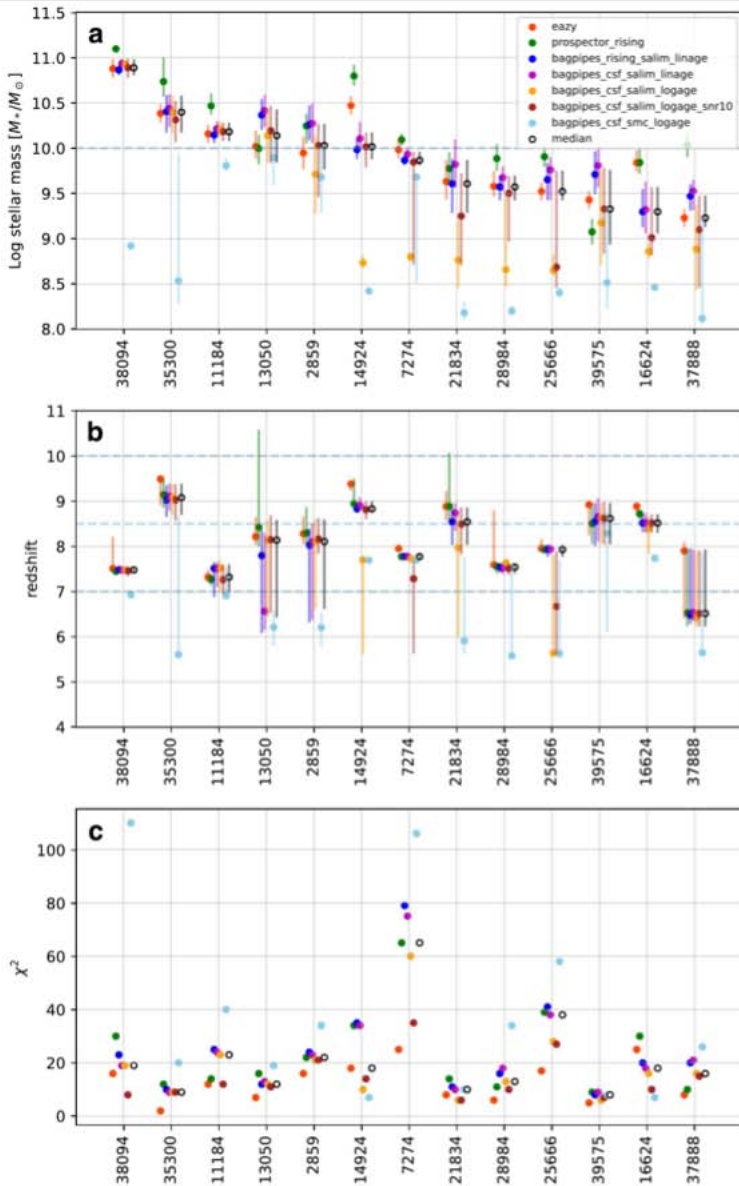
main text. Each cutout has a size of  $2.4'' \times 2.4''$ . The filters range from the  $0.6 \mu\text{m}$  F606W filter of HST/ACS to the  $4.4 \mu\text{m}$  F444W JWST/NIRCam filter.



**Extended Data Fig. 3 | Spectral energy distributions of all 13 galaxies that satisfy the color-color selection. a.** The layout of the figure is identical to Fig. 3a in the main text. In addition, an alternative model fit (model E, see Methods) is shown that produces low stellar masses (blue), but generally

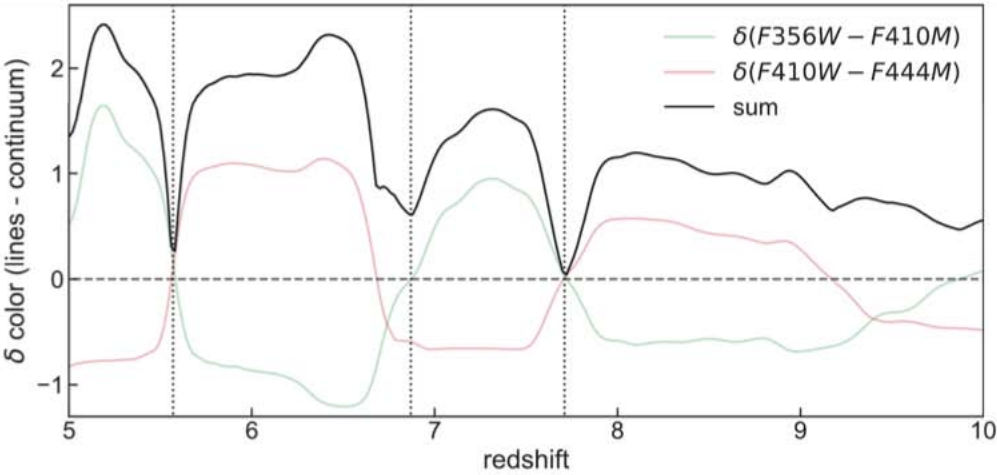
requires extremely young ages ( $< 5$  Myr) at specific narrow redshift intervals. **b.** The panel at the lower right shows the averaged rest-frame SED of the seven galaxies with fiducial  $\log(M^*/M_\odot) < 10$ , compared to previously-found galaxies at similar redshifts (see Fig. 3).





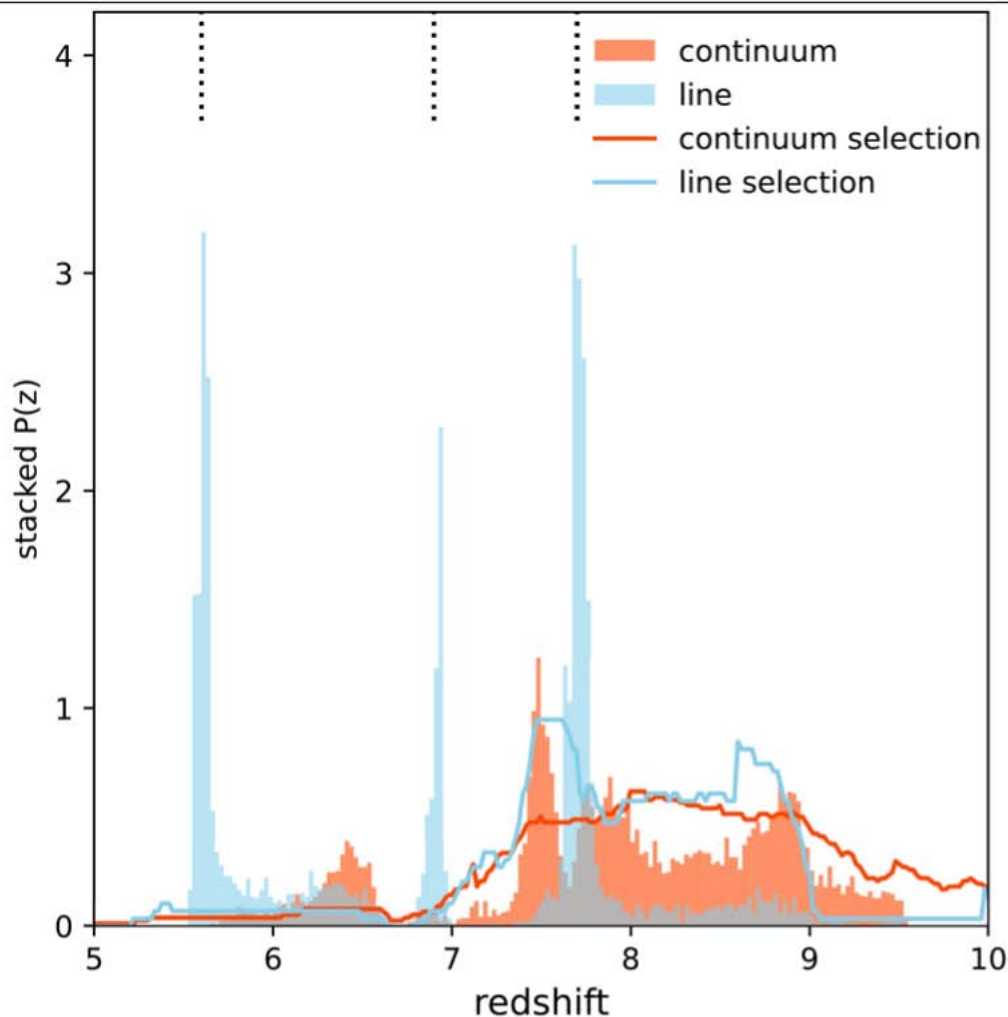
**Extended Data Fig. 4 | Results of the stellar population fitting.** Masses (a), redshifts (b), and the chi-squared fit quality (c) of the 13 galaxies that satisfy the color-color selection. For each galaxy seven different measurements are

shown, as well as the median of the seven that is adopted as the fiducial value (see Methods section). These medians are listed in Extended Data Table 2.



**Extended Data Fig. 5 | Color difference between emission line and continuum-dominated models.** The line-dominated model is a 5 Myr old constant SFH with nebular emission lines. The continuum dominated model is a 50 Myr old CSF without emission lines. Two colors differences involving the line-sensitive F410M filter are shown: F356W-F410M (green) and F410M-F444W

(red) and the sum of their absolute values. When H $\alpha$  and H $\beta$ + [OIII] move through the filters with redshift, the emission line sensitive medium-band F410M filter produces a strong signature, except at  $z = 5.6, 6.9, 7.7$ , where the lines transition between filters. Here continuum and line-dominated SEDs produce similar colors due to undersampling of the SED by the filters.



**Extended Data Fig. 6 | Stacked redshift probability distribution of all 13 galaxies in the sample.** The  $P(z)$  were derived using Bagpipes (as described in Methods). Redshifts of a high mass solution are shown in red (model B: Salim dust attenuation law, rising SFH, linear age prior, continuum dominated) and a low mass solution are shown in blue (model E: SMC dust, logarithmic age prior, emission line dominated). Other high mass fits (e.g., Prospector, EAZY) and low mass fits produce similar  $P(z)$ . Solid curves show expected selection function

under the assumption of continuum (red) or line-dominated models (blue). The high-mass continuum-dominated  $P(z)$  broadly traces the expected selection functions. The low-mass line-dominated  $P(z)$  is not expected for selection of a line-dominated model. The  $P(z)$  is concentrated at narrow redshifts around  $z = 5.6, 6.9, 7.7$  (black dotted lines) where the line-sensitive F410M cannot distinguish between continuum and strong lines due to aliasing.



Extended Data Table 1 | HST/ACS and JWST/NIRCam Photometry of the double break sample

id	f435w	f606w	f814w	f115w	f150w	f200w	f277w	f356w	f410m	f444w
2859	3±4	-2±4	4±5	10±3	15±4	12±4	15±3	52±2	63±6	125±3
7274	-12±6	1±4	6±6	52±3	47±4	51±3	57±3	138±2	147±5	273±3
11184	-10±6	-4±3	-2±5	37±4	47±4	54±6	74±3	219±2	209±6	225±3
13050	5±8	-2±4	7±7	10±3	13±5	12±4	23±3	65±2	82±6	148±4
14924	—	-3±4	1±5	10±3	34±4	35±3	46±2	63±2	117±5	183±2
16624	—	1±4	-3±7	22±4	63±5	57±4	75±3	89±2	117±7	212±3
21834	3±4	-1±4	2±6	4±3	14±4	17±3	18±2	33±2	45±5	83±3
25666	-5±7	2±3	10±7	24±3	24±4	31±3	34±3	94±2	82±6	163±3
28984	-3±7	2±4	-1±7	16±3	22±4	24±3	24±2	55±2	105±5	107±3
35300	—	-4±3	4±5	1±3	15±4	13±4	18±2	38±2	72±7	90±3
37888	1±5	-4±3	2±6	21±4	17±4	21±4	26±3	59±2	49±6	89±3
38094	2±4	2±4	-6±5	52±3	86±4	110±3	169±3	546±3	1003±8	893±4
39575	3±8	4±6	-6±11	0±6	33±8	25±8	28±4	53±4	53±11	94±6

Units are nJy. A fixed 5% uncertainty is added in quadrature to the photometric uncertainties account for calibration errors before fitting with EAZY, Prospector, and Bagpipes.

Extended Data Table 2 | Fiducial redshifts and stellar masses of the double break sample

id	ra	dec	redshift	stellar mass $\log(M_*/M_\odot)$
2859	214.840534	52.817942	8.11(+0.49, - 1.49)(+0.75, - 2.30)	10.03(+0.24, - 0.27)(+0.46, - 0.75)
7274	214.806671	52.837802	7.77(+0.05, - 0.06)(+0.27, - 2.15)	9.87(+0.09, - 0.06)(+0.30, - 1.36)
11184	214.892475	52.856892	7.32(+0.28, - 0.35)(+0.38, - 0.46)	10.18(+0.10, - 0.10)(+0.42, - 0.43)
13050	214.809155	52.868481	8.14(+0.45, - 1.71)(+2.45, - 2.33)	10.14(+0.29, - 0.30)(+0.45, - 0.54)
14924	214.876150	52.880833	8.83(+0.17, - 0.09)(+0.67, - 3.22)	10.02(+0.16, - 0.14)(+0.90, - 1.63)
16624	214.844772	52.892108	8.52(+0.19, - 0.22)(+0.46, - 0.80)	9.30(+0.27, - 0.24)(+0.72, - 0.87)
21834	214.902227	52.939370	8.54(+0.32, - 0.51)(+1.52, - 2.92)	9.61(+0.26, - 0.32)(+0.49, - 1.50)
25666	214.956837	52.973153	7.93(+0.09, - 0.16)(+0.23, - 2.32)	9.52(+0.23, - 0.10)(+0.52, - 1.17)
28984	215.002843	53.007594	7.54(+0.08, - 0.14)(+1.25, - 1.98)	9.57(+0.13, - 0.15)(+0.47, - 1.42)
35300	214.830662	52.887777	9.08(+0.31, - 0.38)(+0.40, - 3.50)	10.40(+0.19, - 0.23)(+0.60, - 2.11)
37888	214.912510	52.949435	6.51(+1.42, - 0.28)(+1.58, - 0.90)	9.23(+0.25, - 0.10)(+0.92, - 1.17)
38094	214.983019	52.955999	7.48(+0.04, - 0.04)(+0.74, - 0.56)	10.89(+0.09, - 0.08)(+0.22, - 1.99)
39575	215.005400	52.996706	8.62(+0.34, - 0.57)(+0.45, - 2.51)	9.33(+0.43, - 0.39)(+0.69, - 1.11)

The adopted redshift and stellar mass are the medians of redshifts and masses computed with 7 different methods (EAZY, Prospector, and Bagpipes (5 variations, including dust, SFH, age prior, and SNR limit), see Methods. A Salpeter IMF is assumed. Two uncertainties are listed ( $\pm(\text{ran}) \pm (\text{sys})$ ) with random uncertainties (ran) corresponding to the median 16th and 84th percentile of the combined posterior distributions, and systematic uncertainties (sys) corresponding to the extremes of all model fits.