

A population of ultraviolet-dim protoclusters detected in absorption

Andrew B. Newman¹, Gwen C. Rudie¹, Guillermo A. Blanc^{1,2}, Mahdi Qezlou^{1,3}, Simeon Bird³, Daniel D. Kelson¹, Victoria Pérez², Enrico Congiu², Brian C. Lemaux^{4,5}, Alan Dressler¹, and John S. Mulchaey¹

¹ Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, California 91101, USA

² Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile

³ Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, California 92521, USA

⁴ Department of Physics and Astronomy, University of California, Davis, One Shields Avenue, Davis, California 95616, USA

⁵ Gemini Observatory, NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, Hawai'i, 96720, USA

Galaxy protoclusters, which will eventually grow into the massive clusters we see in the local universe, are usually traced by locating overdensities of galaxies.¹ Large spectroscopic surveys of distant galaxies now exist, but their sensitivity depends mainly on a galaxy's star formation activity and dust content rather than its mass. Tracers of massive protoclusters that do not rely on their galaxy constituents are therefore needed. Here we report observations of Lyman- α absorption in the spectra of a dense grid of background galaxies,^{2,3} which we use to locate a substantial number of candidate protoclusters at redshifts 2.2-2.8 via their intergalactic gas. We find that the structures producing the most absorption, most of which were previously unknown, contain surprisingly few galaxies compared to the dark matter content of their analogs in cosmological simulations.^{4,5} Nearly all are expected to be protoclusters, and we infer that half of their expected galaxy members are missing from our survey because they are unusually dim at rest-frame ultraviolet wavelengths. We attribute this to an unexpectedly strong and early influence of the protocluster environment^{6,7} on the evolution of these galaxies that reduced their star formation or increased their dust content.

We have mapped Lyman- α (Ly α) absorption over 1.5 deg^2 and the redshift interval $z = 2.2\text{--}2.8$ via the Lyman- α Tomography IMACS Survey (LATIS),³ conducted at the Magellan Baade telescope at Las Campanas Observatory. The survey comprises deep spectra of a dense sample of Lyman-break galaxies and quasars. We reconstructed the three-dimensional (3D) distribution of Ly α absorption imprinted in these spectra by neutral hydrogen in the intergalactic medium (IGM), a technique known as Ly α or IGM tomography.^{2,8} This absorption traces the density field on large scales, and the resulting maps are expected to be particularly effective at locating protoclusters.⁹ LATIS spans three extragalactic deep fields and exceeds the volume of the pioneering Ly α tomography survey CLAMATO¹⁰ by a factor of 10, making it uniquely suited to the discovery of rare structures like protoclusters. The key survey products are 3D maps of the Ly α flux contrast $\delta_F = F/\langle F \rangle - 1$ smoothed to a resolution of $4 h^{-1}$ comoving Mpc (cMpc), where F is the continuum-normalized Ly α flux (Fig. 1), along with 2,241 secure redshifts of Lyman-break galaxies and quasars within the same volumes. These galaxies are selected based on either their photometric redshifts or optical colors³ and r -band magnitudes $r < 24.8$. Their median rest-frame ultraviolet (UV) luminosity density is $\approx L^*$. We refer to these as UV-selected galaxies, because the LATIS selection is typical of techniques based on the rest-UV (observed-frame optical) continuum emission, which traces unobscured star formation.

We identified a large sample of matter overdensities in the IGM maps and examined their galaxy contents. Following previous work,^{9,11} we identified these overdensities as local minima in the smoothed Ly α flux maps. We found 149 such minima having $\delta_F/\sigma_{\text{map}} < -2$, which we refer to as absorption peaks. (The map

standard deviation σ_{map} is dominated by IGM structure rather than observational noise, which is typically $1.8\times$ smaller.³) We interpret the nature of these absorption peaks using a suite of mock LATIS surveys performed in cosmological simulations. Each mock survey realization produces tomographic maps constructed from mock spectra that follow the exact spatial distribution of the LATIS sightlines and their noise properties. The mock maps accurately match the statistical properties of the observed ones. On average, stronger absorption peaks are associated with higher matter overdensities and with the progenitors of more massive $z = 0$ halos in the mock surveys (Methods), as previously reported.¹¹ Based on the statistical distribution in the mock surveys, among the 149 observed absorption peaks with $\delta_F/\sigma_{\text{map}} < -2$ we expect that 76 are progenitors of a galaxy cluster ($M_{z=0}/M_\odot > 10^{14}$) and 43 of a massive group ($10^{13.5} < M_{z=0}/M_\odot < 10^{14}$), on average. Among the 29 absorption peaks with $\delta_F/\sigma_{\text{map}} < -3$, we expect 23 protoclusters and 4 protogroups.

At the location of each absorption peak, we measured the overdensity δ_{gal} of LATIS galaxies within $8 h^{-1}$ cMpc. The connection between the galaxy overdensity and Ly α absorption is shown in Fig. 2. As the absorption increases (δ_F more negative), the galaxy overdensity initially increases as expected. However, the relationship is not monotonic: the strongest absorption peaks, which should correspond to the highest matter overdensities, show a statistically significant but surprisingly small galaxy overdensity ($\delta_{\text{gal}} \approx 1$), on average.

To further examine this unanticipated result, we calculated the overdensity of dark matter halos δ_{halo} in the mock surveys at the location of each simulated absorption peak. We considered halos with masses $M_{\text{vir}} > 10^{11.8} M_\odot$, which match the autocorrelation function of the LATIS galaxies (Methods). These halos were randomly sampled to match the space density of the LATIS galaxies. The expected trend between Ly α absorption and galaxy overdensity, assuming that $\delta_{\text{gal}} = \delta_{\text{halo}}$, is illustrated by the green bands in Fig. 2. The observed trend matches this expectation closely for the majority of absorption peaks having $\delta_F/\sigma_{\text{map}} \gtrsim -3.5$. But this agreement breaks down in the strongest absorption peaks, where the observed δ_{gal} becomes smaller than δ_{halo} in all but $p = 0.005\%$ of the mock surveys; such a significant difference occurs at any value of δ_F in only $p = 0.1\%$ of mock surveys. The observed trend in Fig. 2 is thus unlikely if UV-selected galaxies trace halos independently of the large-scale environment.

The unexpectedly low δ_{gal} is most evident in the eight strongest absorption peaks in LATIS with $\delta_F/\sigma_{\text{map}} < -3.8$. According to their IGM signature, 94% of these are protoclusters (Methods). Yet only the two with the highest δ_{gal} are definitely associated with previously known protoclusters.^{12–16} The remaining six systems have not been detected as galaxy overdensities, even though they lie in some of the most thoroughly observed extragalactic deep fields. Importantly, we find that these strong absorption peaks are not themselves unusual, only their

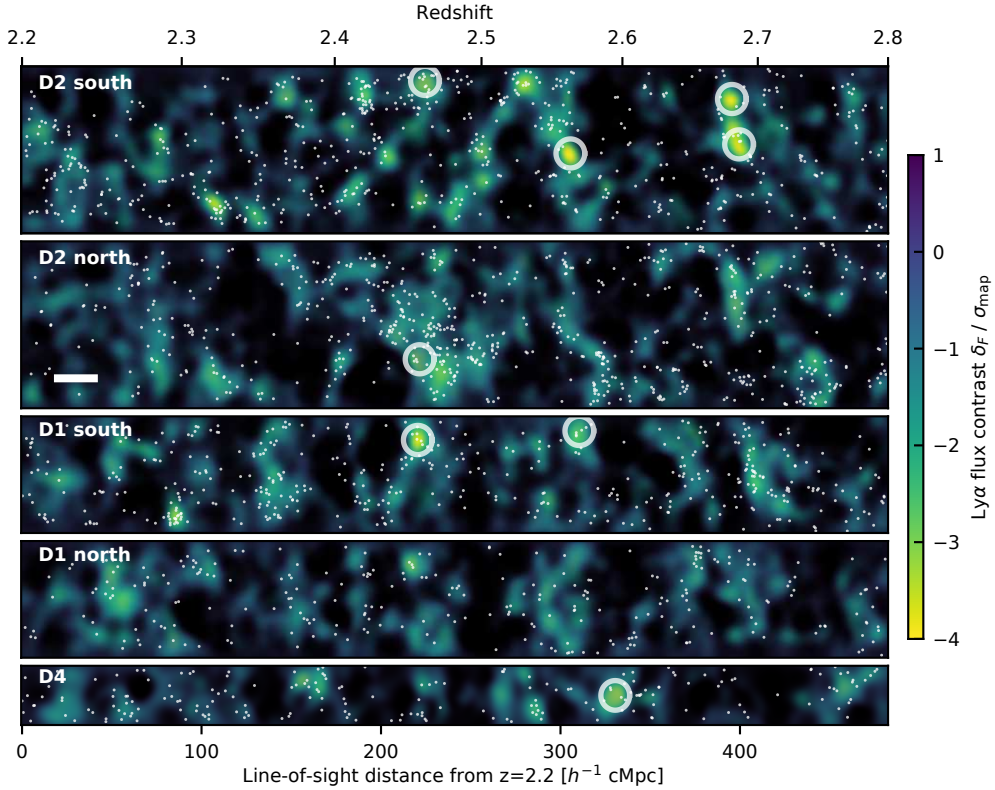


Fig. 1 | Maps of the intergalactic medium. Each panel shows a three-dimensional volume rendering viewed along the declination axis. Structures are rendered semi-transparent with color encoding the Ly α flux contrast. Points show the positions of LATIS galaxies. The maps are smoothed by a Gaussian kernel with $\sigma = 4 h^{-1} \text{ cMpc}$ to maximize sensitivity to protoclusters.⁹ The three survey fields are shown, with D1 and D2 (COSMOS) split into their northern and southern halves. The scale bar is $20 h^{-1} \text{ cMpc}$. The eight strongest absorption peaks are encircled.

galaxy contents. They occur in the LATIS maps with the expected frequency, and they arise from widespread absorption across most of the 10 sightlines that typically probe each structure within $4 h^{-1} \text{ cMpc}$. This and other arguments (Methods) make it clear that contamination by high-column density absorbers associated with galaxies, rather than Mpc-scale overdensities, is not significant. Likewise, locally enhanced ionization (e.g., from quasar radiation) could reduce Ly α absorption, but it cannot account for a low galaxy abundance within strong absorption peaks.

We conclude that most of the strongest Ly α absorption peaks are genuine protoclusters containing an enhanced population of dark matter halos that do not host UV-selected galaxies. By examining the radial profile of δ_{gal} (Fig. 3), we find that the dearth of such galaxies is most prominent toward the centers of the strongest absorption peaks, where the fraction of halos hosting UV-selected galaxies reaches about 50% of its value in other environments. The observed-frame optical and near-infrared colors and luminosities of those LATIS galaxies that are found within the strong absorption peaks are not significantly different from the overall population. This is consistent with some other studies showing little or no difference between Lyman-break galaxies in protoclusters and the field.^{17,18} At lower redshifts $z \lesssim 1$, the effect of the environment is mainly to modulate the mix of galaxy populations (e.g., star-forming and quiescent galaxies) rather than the properties of galaxies with one class.^{19,20} Analogously, one possibility is that the “missing” galaxies in the strongest LATIS absorption peaks are UV-dim because they are dusty or quiescent.

Indeed, previous studies have argued that dusty star-forming galaxies are remarkably abundant in many protoclusters²² and that optical- and UV-dark galaxies are the dominant type at high stellar masses

$M_* \gtrsim 10^{10.5} M_{\odot}$ and redshifts $z > 3$.^{23,24} There is also evidence for a boosted fraction of quenched^{25,26} or quenching²⁷ massive galaxies in a few protoclusters. However, these relatively massive and rare galaxies usually occur alongside a substantial overdensity of UV-selected galaxies. A few studies have identified intriguing cases where UV-selected galaxies appear less numerous than expected based on other tracers^{28–30}, possibly analogous to the UV-dim LATIS structures. The key advance of our study is to detect an ensemble of protoclusters independent of their galaxy content, statistically compare their observed members to the underlying halo population, and thereby infer the presence of galaxies not yet detected in spectroscopic surveys.

The mass distribution of halos hosting LATIS galaxies is expected to be only weakly sensitive to the environment (Methods). This suggests that the distinct properties of galaxies in the strong absorption peaks have a more subtle environmental origin. Quiescent galaxies are expected in semianalytic models to be more common in protoclusters, preferentially at low stellar masses, due to the higher fraction of satellite galaxies and the earlier formation of halos. But this enhancement is predicted to be modest, at most $\sim 20\%$.^{6,7} Our mock surveys based on the IllustrisTNG simulation (Methods) likewise show a very modest increment in the quiescent fraction (a few percentage points) within the strongest Ly α absorption peaks. Observations at lower redshifts suggest that the quiescent fraction of cluster galaxies reaches 50% by $z \sim 1.5$ ³¹ with substantial cluster-to-cluster scatter. Thus, if quiescent galaxies account for the missing population in the strongest LATIS absorption peaks, they must be more abundant than theoretically anticipated, and they likely quenched earlier than most cluster galaxies. That would be consistent with faster galaxy evolution occurring in the most evolved protoclusters, since the strongest Ly α absorption peaks are ex-

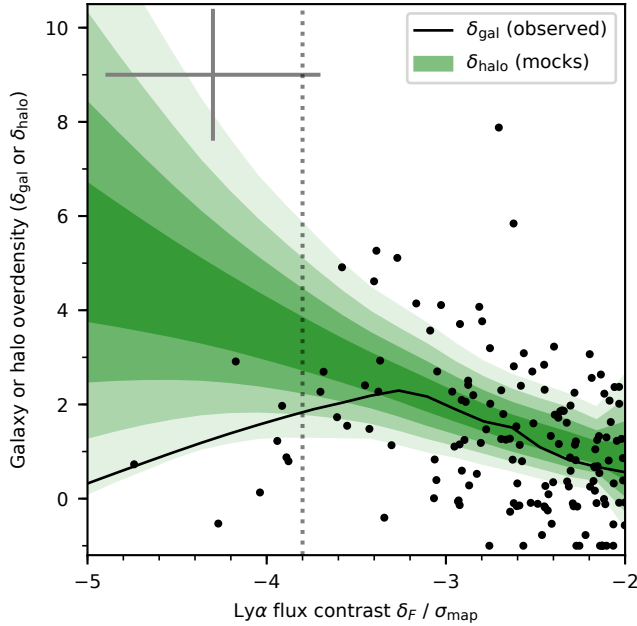


Fig. 2 | The galaxy overdensity δ_{gal} at the location of $\text{Ly}\alpha$ absorption peaks. Black points indicate individual observed absorption peaks, with a representative 1 s.e. error bar shown. More negative δ_F indicates more absorption. The black trendline²¹ is compared to the distribution of trendlines (green bands indicate percentiles equivalent to 1, 2, 3 and 4 s.d. in a normal distribution) in mock surveys that mimic LATIS within a cosmological simulation. In the simulations, we consider the overdensity of dark matter halos δ_{halo} with $M_{\text{vir}} > 10^{11.8} M_{\odot}$. Overdensities are measured within a cylinder that has a transverse radius and half-depth equal to $8 h^{-1} \text{ cMpc}$. In the strongest absorption peaks (left of vertical line), δ_{gal} is much lower than expected.

pected to trace the densest regions on $4 h^{-1} \text{ cMpc}$ scales.

It is also possible that the galaxies missing from the strongest LATIS absorption peaks are moderately UV-dim, rather than being extremely dusty or quiescent. Based on photometric redshift catalogs,³² we estimate that two-thirds of galaxies selected to match the stellar mass distribution of the LATIS sample are fainter than the rest-UV flux limit of our survey and similar ones. The spectral energy distributions of such galaxies indicate that most are only mildly less star-forming or more dusty than their UV-brighter counterparts. An overabundance of such galaxies might account for the low overdensities we observe within the strongest LATIS absorption peaks.

Further observations are needed to search for the implied missing galaxy population. This search is of great interest, but it will be difficult. Although photometric redshifts can, in principle, expand the range of galaxies beyond that accessible with rest-UV spectroscopy, our simulations show that they lack the precision needed to test the trend in Fig. 2 (Methods). Spectroscopic redshifts are thus necessary, but they are very sparse for UV-faint galaxies. Moreover, in order to substantially affect δ_{gal} , a “missing” population must be abundant and thus cannot have stellar masses far exceeding the LATIS galaxies’ median $M_* \approx 10^{9.8} M_{\odot}$.³² A spectroscopic redshift of a quiescent galaxy in this mass and redshift regime has never been obtained, demonstrating the challenge of assembling a complete inventory. Observations with the Atacama Large Millimeter Array (ALMA) will soon test whether sub-mm-selected galaxies reside in two of the strongest LATIS absorp-

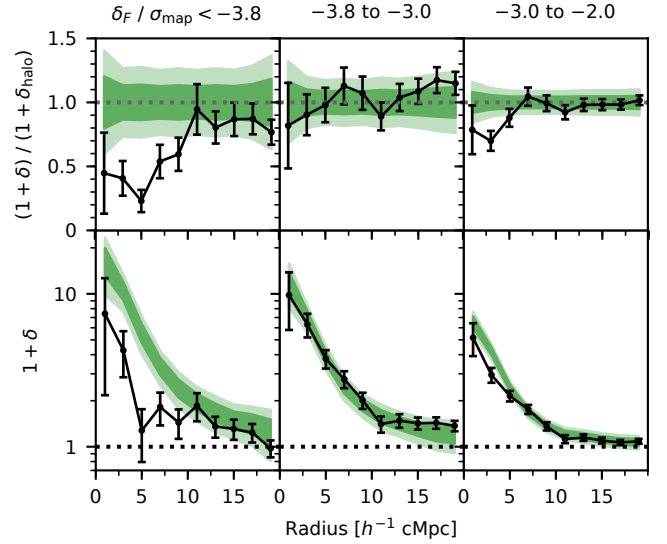


Fig. 3 | Average radial profiles of the galaxy or halo overdensity around $\text{Ly}\alpha$ absorption peaks. The observed $1 + \delta_{\text{gal}}$ (black points with 1 s.e. Poisson uncertainties) and simulated $1 + \delta_{\text{halo}}$ (green bands enclosing 68% and 95% of mock surveys with Poisson noise subtracted; Methods) are measured differentially in a series of nested cylinders, centered on absorption peaks grouped into three bins of δ_F . In the top row, both curves are divided by the expected $1 + \delta_{\text{halo}}$. The quantity $(1 + \delta_{\text{gal}})/(1 + \delta_{\text{halo}})$ is proportional to the fraction of halos that host UV-selected galaxies, which declines markedly toward the center of the strongest absorption peaks.

tion peaks in COSMOS. Since our findings pertain to the average properties of absorption peaks, which individually have substantial uncertainties in δ_F and δ_{gal} and are expected to span a range of overdensities and descendant masses (Methods), definitive follow-up observations must encompass an ensemble of such systems.

The only prior survey to systematically examine the galaxy contents of $\text{Ly}\alpha$ absorption-selected protoclusters reported high galaxy overdensities in three systems.^{33,34} While this may seem to contradict our findings, we emphasize that in most protoclusters, we do measure a galaxy overdensity in accord with simple expectations. Evidence for an UV-dim population only emerges with a larger sample. Further, both the $\text{Ly}\alpha$ absorption and the galaxy overdensity were measured quite differently in the earlier studies: the absorption was measured in 4 sightlines across $20 h^{-1} \text{ cMpc}$, a areal density $15\times$ lower than LATIS, and different tracers ($\text{Ly}\alpha$ and H α emitters) were used to measure δ_{gal} . The latter may prove to be an illuminating distinction.

Despite many searches, a significant population of massive protoclusters at $z \sim 2.5$ seems to have been missed because their galaxy populations are UV-dim. Studies of galaxy evolution in various $z \sim 2-3$ environments may have omitted the structures in which the environmental influence is strongest. A multiwavelength spectroscopic campaign is needed to conduct a galaxy census in the LATIS absorption peaks, measure the galaxy overdensity via multiple tracers, and understand the unusual galaxy properties. If a comprehensive search does not find substantial galaxy overdensities, it would suggest a surprising break in the connection between $\text{Ly}\alpha$ absorption and the density field on scales of several cMpc, a cornerstone of modern cosmology that is thought to be well understood.

1. Overzier, R. A. The realm of the galaxy protoclusters. A review. *Astron. Astrophys. Rev.* **24**, 14 (2016).
2. Lee, K.-G., Hennawi, J. F., White, M., Croft, R. A. C. & Ozbek, M. Observational Requirements for Ly α Forest Tomographic Mapping of Large-scale Structure at $z \sim 2$. *Astrophys. J.* **788**, 49 (2014).
3. Newman, A. B. *et al.* LATIS: The Ly α Tomography IMACS Survey. *Astrophys. J.* **891**, 147 (2020).
4. Klypin, A., Yepes, G., Gottl ber, S., Prada, F. & He , S. MultiDark simulations: the story of dark matter halo concentrations and density profiles. *Mon. Notices R. Astron. Soc.* **457**, 4340–4359 (2016).
5. Nelson, D. *et al.* The IllustrisTNG simulations: public data release. *Computational Astrophysics and Cosmology* **6**, 2 (2019).
6. Contini, E., De Lucia, G., Hatch, N., Borgani, S. & Kang, X. Semi-analytic model predictions of the galaxy population in protoclusters. *Mon. Notices R. Astron. Soc.* **456**, 1924–1935 (2016).
7. Muldrew, S. I., Hatch, N. A. & Cooke, E. A. Galaxy evolution in protoclusters. *Mon. Notices R. Astron. Soc.* **473**, 2335–2347 (2018).
8. Pichon, C., Vergely, J. L., Rollinde, E., Colombi, S. & Petitjean, P. Inversion of the Lyman α forest: three-dimensional investigation of the intergalactic medium. *Mon. Notices R. Astron. Soc.* **326**, 597–620 (2001).
9. Stark, C. W., White, M., Lee, K.-G. & Hennawi, J. F. Protocluster discovery in tomographic Ly α forest flux maps. *Mon. Notices R. Astron. Soc.* **453**, 311–327 (2015).
10. Lee, K.-G. *et al.* First Data Release of the COSMOS Ly α Mapping and Tomography Observations: 3D Ly α Forest Tomography at $2.05 < z < 2.55$. *Astrophys. J. Suppl. Ser.* **237**, 31 (2018).
11. Lee, K.-G. *et al.* Shadow of a Colossus: A $z = 2.44$ Galaxy Protocluster Detected in 3D Ly α Forest Tomographic Mapping of the COSMOS Field. *Astrophys. J.* **817**, 160 (2016).
12. Chiang, Y.-K., Overzier, R. & Gebhardt, K. Discovery of a Large Number of Candidate Protoclusters Traced by ~ 15 Mpc-scale Galaxy Overdensities in COSMOS. *Astrophys. J. Lett.* **782**, L3 (2014).
13. Chiang, Y.-K. *et al.* Surveying Galaxy Proto-clusters in Emission: A Large-scale Structure at $z = 2.44$ and the Outlook for HETDEX. *Astrophys. J.* **808**, 37 (2015).
14. Diener, C. *et al.* A Protocluster at $z = 2.45$. *Astrophys. J.* **802**, 31 (2015).
15. Casey, C. M. *et al.* A Massive, Distant Proto-cluster at $z = 2.47$ Caught in a Phase of Rapid Formation? *Astrophys. J. Lett.* **808**, L33 (2015).
16. Cucciati, O. *et al.* The progeny of a cosmic titan: a massive multi-component proto-supercluster in formation at $z = 2.45$ in VUDS. *Astron. Astrophys.* **619**, A49 (2018).
17. Cucciati, O. *et al.* Discovery of a rich proto-cluster at $z = 2.9$ and associated diffuse cold gas in the VIMOS Ultra-Deep Survey (VUDS). *Astron. Astrophys.* **570**, A16 (2014).
18. Lemaux, B. C. *et al.* VIMOS Ultra-Deep Survey (VUDS): Witnessing the assembly of a massive cluster at $z \sim 3.3$. *Astron. Astrophys.* **572**, A41 (2014).
19. Blanton, M. R. & Moustakas, J. Physical Properties and Environments of Nearby Galaxies. *Annu. Rev. Astron. Astrophys.* **47**, 159–210 (2009).
20. Muzzin, A. *et al.* The Gemini Cluster Astrophysics Spectroscopic Survey (GCLASS): The Role of Environment and Self-regulation in Galaxy Evolution at $z \sim 1$. *Astrophys. J.* **746**, 188 (2012).
21. Cappellari, M. *et al.* The ATLAS^{3D} project - XX. Mass-size and mass- σ distributions of early-type galaxies: bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *Mon. Notices R. Astron. Soc.* **432**, 1862–1893 (2013).
22. Casey, C. M. The Ubiquity of Coeval Starbursts in Massive Galaxy Cluster Progenitors. *Astrophys. J.* **824**, 36 (2016).
23. Wang, T. *et al.* A dominant population of optically invisible massive galaxies in the early Universe. *Nature* **572**, 211–214 (2019).
24. Shu, X. *et al.* A Census of Optically Dark Massive Galaxies in the Early Universe from Magnification by Lensing Galaxy Clusters. *Astrophys. J.* **926**, 155 (2022).
25. Wang, T. *et al.* Discovery of a Galaxy Cluster with a Violently Starbursting Core at $z = 2.506$. *Astrophys. J.* **828**, 56 (2016).
26. McConachie, I. *et al.* Spectroscopic Confirmation of a Protocluster at $z = 3.37$ with a High Fraction of Quiescent Galaxies. *arXiv:2109.07696* (2021).
27. Zavala, J. A. *et al.* On the Gas Content, Star Formation Efficiency, and Environmental Quenching of Massive Galaxies in Protoclusters at $z \approx 2.0$ – 2.5 . *Astrophys. J.* **887**, 183 (2019).
28. Chapman, S. C. *et al.* Do Submillimeter Galaxies Really Trace the Most Massive Dark-Matter Halos? Discovery of a High- z Cluster in a Highly Active Phase of Evolution. *Astrophys. J.* **691**, 560–568 (2009).
29. Hung, C.-L. *et al.* Large-scale Structure around a $z = 2.1$ Cluster. *Astrophys. J.* **826**, 130 (2016).
30. Shi, K. *et al.* How Do Galaxies Trace a Large-scale Structure? A Case Study around a Massive Protocluster at $Z = 3.13$. *Astrophys. J.* **879**, 9 (2019).
31. Nantais, J. B. *et al.* Evidence for strong evolution in galaxy environmental quenching efficiency between $z = 1.6$ and $z = 0.9$. *Mon. Notices R. Astron. Soc.* **465**, L104–L108 (2017).
32. Weaver, J. R. *et al.* COSMOS2020: A Panchromatic View of the Universe to z 10 from Two Complementary Catalogs. *Astrophys. J. Suppl. Ser.* **258**, 11 (2022).
33. Cai, Z. *et al.* Mapping the Most Massive Overdensities through Hydrogen (MAMMOTH). II. Discovery of the Extremely Massive Overdensity BOSS1441 at $z = 2.32$. *Astrophys. J.* **839**, 131 (2017).
34. Shi, D. D. *et al.* Spectroscopic Confirmation of Two Extremely Massive Protoclusters, BOSS1244 and BOSS1542, at $z = 2.24$. *Astrophys. J.* **915**, 32 (2021).

METHODS

Tomographic maps. We created tomographic maps using the spectra obtained as part of LATIS between Dec. 2017 and Apr. 2021. The survey and map making procedures have been fully described.³ The updated data set used here comprises observations of 11 LATIS “footprints,” which are distributed among the COSMOS and Canada–France–Hawaii Legacy Survey (CFHTLS) D1 and D4 fields. Only one planned footprint (D4M1, ref.³) has not been observed. In two footprints (D1M1 and D1M2), only half of the planned targets have been observed. We took this into account throughout our analysis, but we note that none of the strongest Ly α absorption peaks of main interest are located in these footprints. We used a total of 2,596 sightlines to construct the Ly α tomographic maps. The maps span a total area of 1.5 deg^2 and a volume, extending from $z = 2.2$ to 2.8 , of $3.7 \times 10^6 h^{-3} \text{ cMpc}^3$. The δ_F maps are smoothed by an isotropic gaussian kernel having $\sigma_{\text{ker}} = 4 h^{-1} \text{ cMpc}$. They are then normalized by their standard deviation σ_{map} , which is consistent among fields (ranging from 0.0457 to 0.0466) and with the mock surveys. Absorption peaks are identified as local minima. We exclude minima within σ_{ker} of the map boundary to suppress edge effects. We also iteratively remove minima located within $8 h^{-1} \text{ cMpc}$ of another minimum, beginning with the shallowest. Though not essential, this step ensures that the measurements are nearly statistically independent.

Galaxy density measurements. We measured galaxy overdensities using 2,241 redshifts of Lyman-break galaxies (LBGs) and quasars located within the tomographic map volume. Only high-confidence ($z_{\text{qual}} = 3$ or 4, ref.³) LATIS redshifts were used. The accuracy of the systemic redshifts of the LBGs was verified based on the symmetry of the average Ly α absorption in transverse sightlines.³⁵ At the location of each absorption peak, we counted the number N of galaxies within a cylinder that has a transverse radius R and extends $\pm R$ along the line of sight. Redshift was converted to line-of-sight distance assuming a pure Hubble flow, so like the δ_F maps, galaxy densities are measured in velocity space. In Fig. 2 we count galaxies within a $R = 8 h^{-1} \text{ cMpc}$ cylinder, which optimizes the signal-to-noise ratio in δ_{gal} based on the simulations outlined below. In Fig. 3 we count galaxies differentially in a series of nested cylinders, which are labeled by the mean radius $\bar{R}$ of the inner and outer cylinders. The half-length of the cylinder along the line of sight is always equal to the radius. The mean galaxy space density $\bar{N}$ was evaluated within each survey field and was found not to vary systematically with redshift within the map volume. The galaxy overdensity is then $\delta_{\text{gal}} = N/(\bar{N}V) - 1$, where V is the search volume excluding unobserved portions (i.e., those outside the survey footprint or close to bright stars). The space density of LATIS galaxies is $\bar{N} = 7 \times 10^{-4} h^3 \text{ cMpc}^{-3}$ corresponding to an average $\bar{N}V = 2.4$ galaxies in randomly placed cylinders.

This simple procedure is sufficient because of the uniformity of LATIS over the relevant angular scales. The fraction of candidates that were observed and assigned a confident redshift varies within each field with an rms of 10%, when measured in apertures of size $R = 8 h^{-1} \text{ cMpc} \sim 6.8'$ and accounting for the survey mask. These variations are $5\times$ smaller than the typical uncertainty in δ_{gal} and are not correlated with δ_{gal} , so they are safely neglected. Geometric constraints on slitmask design, which arise from the requirement that galaxies’ spectra not overlap, induce only a negligible bias in δ_{gal} . The fraction of target pairs that were observed is constant for angular separations larger than 12 arcsec, much smaller than any aperture considered in this paper. To verify the insensitivity of δ_{gal} to slit density limits, we selected halos from a large N -body simulation (see below), added a randomly distributed population of foreground interlopers, and placed the combined “targets” onto two large pseudo-masks (equivalent to 18 deg^2) respecting mask design constraints. We find that the recovered halo space densities, for the range of apertures and overdensities shown in Fig. 3, are biased by $\lesssim 5\%$. Finally, although the target selection probability is dependent on r -band magnitude,³ its distribution is consistent for galaxies located within absorption peaks binned according to δ_F ; thus, trends in target selection probability cannot affect trends with δ_F (Figs. 2 and 3).

Galaxy clustering. We measured the autocorrelation function $w_p(R)$ of LATIS galaxies at $2 < z < 3$, projected within a velocity window corresponding to $\pm 20 h^{-1} \text{ cMpc}$, using the Landy-Szalay estimator.³⁶ The results are shown for the COSMOS and CFHTLS-D1 fields in Extended Data Fig. 1. (The D4 field is too small to usefully estimate w_p .) Random

pairs were generated respecting the survey mask and redshift distribution. The integral constraint, which accounts for the survey geometry, was evaluated following standard methods.³⁷ We then computed $w_p(R)$ for dark matter subhalos in the MultiDark *Planck* 2 (MDPL2) N -body simulation⁴ using the snapshot at $z = 2.535$. We considered halos with virial masses $M_{\text{vir}} > M_{\text{min}}$ for several values of M_{min} . We fit the LATIS measurements to the halo clustering over the range $1 < R/(h^{-1} \text{ cMpc}) < 7$ and derive $M_{\text{min}}/M_{\odot} = 10^{11.89 \pm 0.03}$ and $10^{11.70 \pm 0.05}$ for the COSMOS and D1 fields, respectively. Since the field-to-field difference is larger than the formal uncertainties, we average the two fields and consider halos with $M_{\text{vir}} > 10^{11.8} M_{\odot}$ by default, consistent with ref.³⁸.

We tested the sensitivity of our results to the galaxy-halo connection. First, we repeated the analysis perturbing the halo mass limit by the field-to-field variation of ± 0.1 dex and found that the small p -values derived from Fig. 2 changed by less than a factor of 2. Second, we considered a model in which the probability of a halo hosting a LATIS-selected galaxy varies smoothly with mass. This probability followed the error function commonly used in halo occupation distribution models, which is 0.5 at a mass M_{min} and transitions from 0 to 1 at a rate controlled by $\sigma_{\log M}$ (see eqn. 6 of ref.³⁹). We set $\sigma_{\log M} = 0.6$ dex to match the highest values inferred from LBG clustering^{39,40} and then tuned M_{min} to match our clustering data. In this model, LATIS galaxies occupy halos with a wide mass range ($\log M_{\text{halo}}/M_{\odot}$ has a median of 12.0 and a s.d. of 0.45) compared to our fiducial mass-limited model. We find that the comparison between the observed and mock trendlines in Fig. 2 is unaffected. The reason is that the center of the δ_{halo} trendline distribution is fixed by the clustering data while its width is dominated by observational errors, not scatter in the halo bias.

Mock surveys. We created a suite of mock LATIS surveys within the 1 Gpc³ MDPL2 simulation. The flux field at $z = 2.535$, near the LATIS midpoint, was calculated using the fluctuating Gunn–Peterson approximation,^{41,42} which we previously found to accurately reproduce the distribution of fluctuations in the observed maps.³ For each mock survey realization and each of the three LATIS survey fields, we sampled the simulated flux field using the exact configuration of the observed sightlines (i.e., the relative positions of each spectral pixel in x , y , and z). The mock spectra were smoothed and resampled to match the observations. Random noise and continuum errors were injected following the noise properties of each sightline. We applied mean flux regularization⁴³ and masked lines with high equivalent width³ to suppress damped absorbers, following the same methods applied to the observations. A mock tomographic map was created using the *dachshund* Wiener filtering code⁹ using the same parameters applied to the observations. The process was repeated over 100 mock surveys covering nearly disjoint subvolumes of the simulation. To calculate the halo overdensity at a simulated absorption peak, we selected a random subset of subhalos in the search aperture having $M_{\text{vir}} > 10^{11.8} M_{\odot}$ (see “Galaxy clustering”), with the selection probability set to match the global space density of halos and LATIS galaxies. We propagated the survey mask into the simulated volumes and excluded halos in “unobserved” regions. Extended Data Figure 2 shows that the mock surveys match the observed distributions of δ_F and δ_{gal} with impressive accuracy.

We additionally created a separate suite of mock surveys using the *IllustrisTNG300-1* simulation.⁵ Although its $(205 h^{-1} \text{ cMpc})^3$ volume is only about twice that of LATIS, and thus too small to capture variance in the density field itself, these mocks are useful to test the relevance of physics not present in the collisionless MDPL2 simulation, including hydrodynamics, feedback, and the formation of high-column density (HCD) lines from dense gas. We used the *fake_spectra* code^{44,45} to create 168,000 high-resolution synthetic spectra on a regular grid in the $z = 2.58$ snapshot. Self-shielding of dense gas was modeled using results from radiative transfer calculations⁴⁶; other details may be found in ref.⁴⁷. From the TNG spectra, we created 100 mock surveys. In each realization, we selected random sightlines with a density equal to LATIS, averaged over the redshift range of the maps (750 deg^{-2}), and we assigned each sightline a continuum-to-noise ratio drawn from the observed distribution. Mock observations were then synthesized from the spectra as described above. One physical effect not present in TNG is enhanced photo-ionization near quasars. This is expected to have a mild effect on the protocluster signal in most cases.⁴⁸ More importantly, it can only reduce the observed absorption, and thus cannot account

for our main result, the low galaxy density observed around the strongest absorption peaks. The same comment applies to any “pre-heating” of the intracluster medium⁴⁹ beyond the feedback implemented in TNG.

The $\delta_F - \delta_{\text{gal}}/\text{halo}$ connection. To produce the distribution of trendlines in Fig. 2, we sampled random subsets of absorption peaks from the full suite of MDPL2 mock surveys. The subsets were constrained to have the same number of peaks in several bins of δ_F as the LATIS data, in order to correctly model the uncertainty in the mean $\langle \delta_{\text{gal}} \rangle$ conditional on δ_F . For each such subset, we calculated a non-parametric trendline using locally weighted scatterplot smoothing (LOESS²¹), a form of local linear regression. The same method was applied to the observations. Shaded regions in Fig. 2 show percentiles of the suite of trendlines. The percentile curves are non-monotonic because the uncertainty increases rapidly at low δ_F due to the declining number of absorption peaks. Varying the method used to define the trendlines (e.g., varying the fraction of datapoints used in the local fitting between 30%-70%, or using a second-order global polynomial fit instead of LOESS) affects the trendline shape but has a minor effect (0.1 s.d.) on the location of the observed trendline within the mock distribution. Excluding the lowest- δ_F data point is similarly inconsequential. The mock surveys based on MDPL2 (used in Figs. 2 and 3) and TNG produce very similar trendline distributions (Extended Data Fig. 3a), despite substantial differences in the physics and the mock observation synthesis. We thus find that a simple treatment of the IGM physics is sufficient for our purposes.

The distributions of δ_{gal} and δ_{halo} for individual absorption peaks, in several bins of δ_F , are shown in Extended Data Fig. 4. Generalizing the result in Fig. 2, we find that the observed and mock survey distributions, not only their means, are consistent in all but the strongest absorption peaks. To construct the $\delta_{\text{halo}}(R)$ profiles in Fig. 3 from the mock surveys, we subtracted the Poisson noise in quadrature from the standard deviation of δ_{halo} so that the width of the bands indicates the additional sources of error (e.g., bin-to-bin scatter from errors in δ_F) that are hard to estimate otherwise. The standard Poisson noise is instead shown as error bars on the observations. The mock surveys naturally incorporate the $4 h^{-1} \text{ cMpc}$ (1 s.e.) uncertainty in the absorption peak position, which depresses δ_{gal} and δ_{halo} in the inner bins.

High-column density absorption lines. HCD lines, including Lyman-limit systems and damped Ly α systems (DLAs) that occur near galaxies, produce absorption that could wrongly be interpreted as a large-scale matter overdensity if it were assumed to arise from the diffuse IGM. Contamination by HCD lines is the predominant concern for single-sightline surveys⁵⁰ but is expected to be much reduced in tomographic surveys like LATIS, in which each map resolution element is probed by ~ 10 sightlines.

We filter lines with high equivalent width to mask DLAs, and the same procedure³ is applied to the observations and all mock surveys. We particularly examined the 77 sightlines near the strongest absorption peaks (Extended Data Fig. 5) and found that 88% have $\delta_F < 0$, indicating that the peaks are produced by widespread IGM absorption, consistent with expectations from the mock surveys with no evidence of contamination by DLAs. More subtle effects were evaluated probabilistically using TNG. Since the trendline distributions in the TNG- and MDPL2-based mocks are very similar (Extended Data Figure 3a), particularly in the low- δ_F , low- δ_{halo} region of main interest in this paper, HCD lines seem to negligibly affect the analysis in Fig. 2. To further probe this insensitivity, we precisely isolated the effect of HCD lines by creating a matched suite of TNG mock surveys, which are identical to the fiducial TNG mocks except that all HCD lines ($N_{\text{HI}} > 10^{17.2} \text{ cm}^{-2}$) are explicitly masked. The differential effect of HCD lines on individual absorption peaks is shown in Extended Data Fig. 3b. As expected, the distribution has a tail to negative shifts, i.e., enhanced absorption produced by HCD lines. We consider the influence of these shifts on the trendline distribution in Extended Data Fig. 3c, which shows that excluding HCD lines produces a shift of just 0.06 s.d. Although this may be a slight underestimate, since the frequency of HCD lines in the TNG spectra is $30 \pm 9\%$ lower than observed⁵¹ (depending on the column densities considered), the effect of HCD lines would be very minor even if doubled. We conclude that the significance of the discrepancy in Fig. 2 is virtually insensitive to the presence of HCD lines, because the map perturbations they produce are subdominant to other sources of noise in δ_F and δ_{gal} .

Connection to the matter overdensity at $z \sim 2.5$ and massive halos at

$z = 0$. The mock surveys allow us to connect the observed δ_F to the density contrast $\delta_m = \rho/\langle\rho\rangle - 1$ at the observed epoch, smoothed on $\sigma = 4 h^{-1}$ cMpc scales like the flux field. The left panel of Extended Data Fig. 6 shows that the absorption peaks correspond to progressively larger matter overdensities as δ_F decreases, reaching 4.6σ overdensities, on average, when $\delta_F/\sigma_{\text{map}} < -4$. We also evaluated the connection to massive halos at $z = 0$ by finding the most massive halo within $4 h^{-1}$ cMpc of each absorption peak and tracing the mass of its descendant, similar to earlier work.⁹ The right panel of Fig. 6 shows that the distributions overlap substantially, but the strongest absorption peaks are almost always associated with proto-clusters ($M_{\text{vir}}(z=0)/M_\odot > 10^{14}$).

Environmental trends in the halo mass distribution. The halo mass distribution is expected to be tilted to higher masses in dense environments traced by strong Ly α absorption.⁵² LATIS galaxies span about 1.3 dex in stellar mass³², which is thought to correspond to only ~ 0.6 dex in halo mass⁵³. We find that the shape of the halo mass distribution, measured in our mock surveys within subvolumes defined by δ_F , varies only slightly over this narrow mass range: even considering all halos with $M_{\text{vir}} > 10^{11.8} M_\odot$, the median M_{vir} shifts by < 0.1 dex within the strongest absorption peaks. We conclude that higher halo masses alone are unlikely to account for a major shift in the properties of galaxies within absorption peaks.

Photometric redshift maps. We evaluated the overdensity of near-infrared-selected galaxies ($K_s < 25$ AB) around the LATIS absorption peaks located in the COSMOS field using a photometric redshift catalog.³² By comparison to the LATIS spectroscopic redshifts, we found that the z_{phot} estimates have a random uncertainty of $\sigma_{z_{\text{phot}}}/(1+z) = 0.03$ and a redshift-dependent bias reaching up to $\Delta z/(1+z) = 0.03$. We corrected for this bias in our comparisons. We calculated δ_{gal} within a redshift interval of $\pm\sigma_{z_{\text{phot}}}/2$ centered on each LATIS absorption peak, following an earlier z_{phot} -based protocluster search¹², and within a transverse separation of $8 h^{-1}$ cMpc to match Fig. 2. We simulated the same observation in our mock surveys by considering halos with $M_{\text{vir}} > 10^{11.4} M_\odot$, which match the space density of the z_{phot} sample, and perturbing their z coordinates according to the z_{phot} uncertainty. Extended Data Figure 7 shows that the overdensities are greatly reduced from the spectroscopic δ_{gal} (Fig. 2) due to the large smearing along the line of sight. Similar to Fig. 2, the observed $\delta_F - \delta_{\text{gal}}$ trendline agrees with the mean mock survey at $\delta_F/\sigma_{\text{map}} > -3$ and falls to lower δ_{gal} at more extreme δ_F ; unlike Fig. 2, this discrepancy is not significant ($\sim 1.5\sigma$). We conclude that spectroscopic trend in Fig. 2 is not inconsistent with the photometric redshift analysis, but the large z_{phot} uncertainties prevent a strong test.

Counterparts of the strongest absorption peaks in the literature. In Extended Data Table 8 we provide the coordinates and properties of the 8 strongest LATIS absorption peaks ($\delta_F/\sigma_{\text{map}} < -3.8$) discussed in this paper. Among these, LATIS1-D2-3 is part of a large proto-supercluster known as Hyperion,¹⁶ of which components have previously been identified using galaxy surveys^{13–15} and Ly α tomography.¹¹ LATIS1-D2-1 was previously identified (ID 20) in a catalog of galaxy overdensities identified using photometric redshifts;¹² it is the highest such overdensity in the portion of the catalog that overlaps the LATIS maps, and it has the highest δ_{gal} listed in Extended Data Table 1. LATIS1-D2-4 has a potential match (ID 25) in the same catalog, but we do not see a corresponding structure in our photometric redshift maps, which we calibrated against a large spectroscopic library, and conclude that this match is uncertain. As far as we are aware, the other structures have not been previously discussed.

Tests of sky subtraction. Since the flux density of the Ly α forest is typically $30\times$ smaller than the sky background, small biases in sky subtraction could affect the galaxy spectra. Sky subtraction was performed on a per-exposure basis using two-pass, iterative b-spline modeling of the two-dimensional spectra. We evaluated biases by stacking the two-dimensional spectra and comparing the target (“on-source”) to a “blank” part of the slit (“off-source”), which was taken as the region 1.6–2.4 arcsec distant from the target (the slit length is 6 arcsec). Since we are particularly interested in the sky subtraction quality near the absorption peaks, we considered the sightlines with a transverse distance within $8 h^{-1}$ cMpc of an absorption peak. (We found consistent results when considering only the strongest absorption peaks.) These spectra were shifted into the rest frame of the absorption peak, continuum normalized, and averaged. The Ly α absorption is clearly

evident on source, as expected. We also find some residual light off source; after subtracting it from the on-source spectrum and renormalizing the continuum, the Ly α absorption increases modestly in strength by 5%. This suggests that there may be a residual additive “pedestal” around 5% of the mean Ly α forest flux, or equivalently 0.2% of the sky background. If the pedestal comprises residual sky background, we note that its spectrum is essentially smooth at the resolution of the maps. An additive uncertainty of 5% translates to an equal multiplicative uncertainty in δ_F . This should minimally affect our analysis, since δ_F is ultimately normalized by σ_{map} and so is insensitive to a constant multiplicative rescaling. To verify this we propagated the effect of an additive pedestal through our analysis, assuming that it can be represented as a constant count rate of 6×10^{-5} electron s $^{-1}$ pixel $^{-1}$, which reproduces the fractional estimate above. We subtracted this rate from all two-dimensional spectra, re-extracted the one-dimensional spectra, and re-created the maps. Comparing the strength of absorption peaks between the modified the fiducial maps, we find that the difference $\Delta\delta_F/\sigma_{\text{map}}$ has a mean of 0, as anticipated, and a standard deviation of 0.2, which is $3\times$ smaller than the random noise and so has a minimal effect. In particular, the trendline in Fig. 2 is not notably affected. We conclude that our results are robust to small systematic uncertainties in the sky subtraction.

Data availability. Data supporting the findings of this study are available from the corresponding author upon request.

Code availability. We have made use of public codes including *astropy*^{54,55} and *dachshund*.⁹ Other supporting analysis code is available from the corresponding author upon request.

Methods References

35. Rakic, O., Schaye, J., Steidel, C. C. & Rudie, G. C. Calibrating galaxy redshifts using absorption by the surrounding intergalactic medium. *Mon. Notices R. Astron. Soc.* **414**, 3265–3271 (2011).
36. Landy, S. D. & Szalay, A. S. Bias and Variance of Angular Correlation Functions. *Astrophys. J.* **412**, 64 (1993).
37. Roche, N. D., Almaini, O., Dunlop, J., Ivison, R. J. & Willott, C. J. The clustering, number counts and morphology of extremely red ($R - K > 5$) galaxies to $K \leq 21$. *Mon. Notices R. Astron. Soc.* **337**, 1282–1298 (2002).
38. Trainor, R. F. & Steidel, C. C. The Halo Masses and Galaxy Environments of Hyperluminous QSOs at $z \sim 2.7$ in the Keck Baryonic Structure Survey. *Astrophys. J.* **752**, 39 (2012).
39. Durkalec, A. *et al.* The VIMOS Ultra Deep Survey. Luminosity and stellar mass dependence of galaxy clustering at $z \sim 3$. *Astron. Astrophys.* **612**, A42 (2018).
40. Ishikawa, S. *et al.* The Galaxy-Halo Connection in High-redshift Universe: Details and Evolution of Stellar-to-halo Mass Ratios of Lyman Break Galaxies on CFHTLS Deep Fields. *Astrophys. J.* **841**, 8 (2017).
41. Gunn, J. E. & Peterson, B. A. On the Density of Neutral Hydrogen in Intergalactic Space. *Astrophys. J.* **142**, 1633–1636 (1965).
42. Weinberg, D. H., Hernquist, L., Katz, N., Croft, R. & Miralda-Escudé, J. Hubble Flow Broadening of the Ly α Forest and its Implications. In Petitjean, P. & Charlot, S. (eds.) *Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System*, 133 (1997).
43. Lee, K.-G., Suzuki, N. & Spergel, D. N. Mean-flux-regulated Principal Component Analysis Continuum Fitting of Sloan Digital Sky Survey Ly α Forest Spectra. *Astron. J.* **143**, 51 (2012).
44. Bird, S. *et al.* Reproducing the kinematics of damped Lyman α systems. *Mon. Notices R. Astron. Soc.* **447**, 1834–1846 (2015).
45. Bird, S. FSFE: Fake Spectra Flux Extractor. *Astrophysics Source Code Library*, ascl:1710.012 (2017).
46. Rahmati, A., Pawlik, A. H., Raičević, M. & Schaye, J. On the evolution of the H I column density distribution in cosmological simulations. 2427–2445 (2013).
47. Qezlou, M., Newman, A. B., Rudie, G. C. & Bird, S. Characterizing protoclusters and protogroups at $z \sim 2.5$ using Lyman- α Tomography. arXiv:2112.03930 (2021).
48. Miller, J. S. A., Bolton, J. S. & Hatch, N. A. Searching for the shadows of giants - II. The effect of local ionization on the Ly α absorption signatures of protoclusters at redshift $z \sim 2.4$. *Mon. Notices R. Astron. Soc.* **506**, 6001–6013 (2021).
49. Kooistra, R., Inoue, S., Lee, K.-G., Cen, R. & Yoshida, N. Detecting preheating in proto-clusters with Lyman- α Forest Tomography. arXiv:2109.09954 (2021).
50. Cai, Z. *et al.* Mapping the Most Massive Overdensity Through Hydrogen (MAM-MOTH) I: Methodology. *Astrophys. J.* **833**, 135 (2016).
51. Zafar, T. *et al.* The ESO UVES advanced data products quasar sample. II. Cosmological evolution of the neutral gas mass density. *Astron. Astrophys.* **556**, A141 (2013).
52. Lee, K.-G. & White, M. Revealing the $z \sim 2.5$ Cosmic Web with 3D Ly α Forest Tomography: a Deformation Tensor Approach. *Astrophys. J.* **831**, 181 (2016).
53. Behroozi, P. S., Wechsler, R. H. & Conroy, C. The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z = 0 - 8$. *Astrophys. J.* **770**, 57 (2013).
54. Astropy Collaboration *et al.* Astropy: A community Python package for astronomy. *Astron. Astrophys.* **558**, A33 (2013).
55. Astropy Collaboration *et al.* The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *Astron. J.* **156**, 123 (2018).

Acknowledgments This paper includes data gathered with the 6.5-meter Magellan Telescopes located at Las Campanas Observatory, Chile. We gratefully acknowledge the support of the Observatory staff. A.B.N., S.B., and B.C.L. acknowledge support from the National Science Foundation under Grant Nos. 2108014, 2107821, and 1908422, respectively. E.C. acknowledges support from ANID project Basal AFB-170002. This work is based in part on observations obtained with MegaPrime / MegaCam, a joint project of CFHT and CEA/IRFU, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l’Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. Also based in part on data products produced at Terapix available at the Canadian Astronomy Data Centre as part of the Canada–France–Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

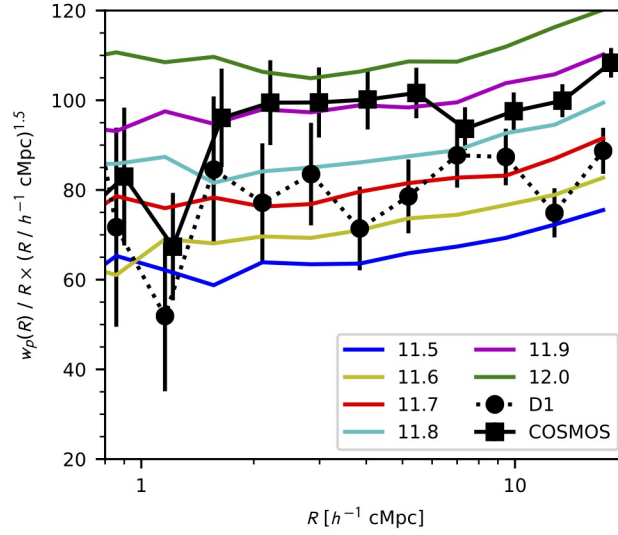
Author contributions A.B.N., G.C.R., and G.A.B. designed LATIS, obtained telescope access, and, along with V.P. and E.C., conducted the observations. D.D.K. created the data reduction software. A.B.N., M.Q., and S.B. created the mock surveys. A.B.N. processed the observations, created the Ly α and galaxy density maps, and drafted the manuscript. All authors contributed to the interpretation and manuscript preparation.

Competing interests The authors declare no competing interests.

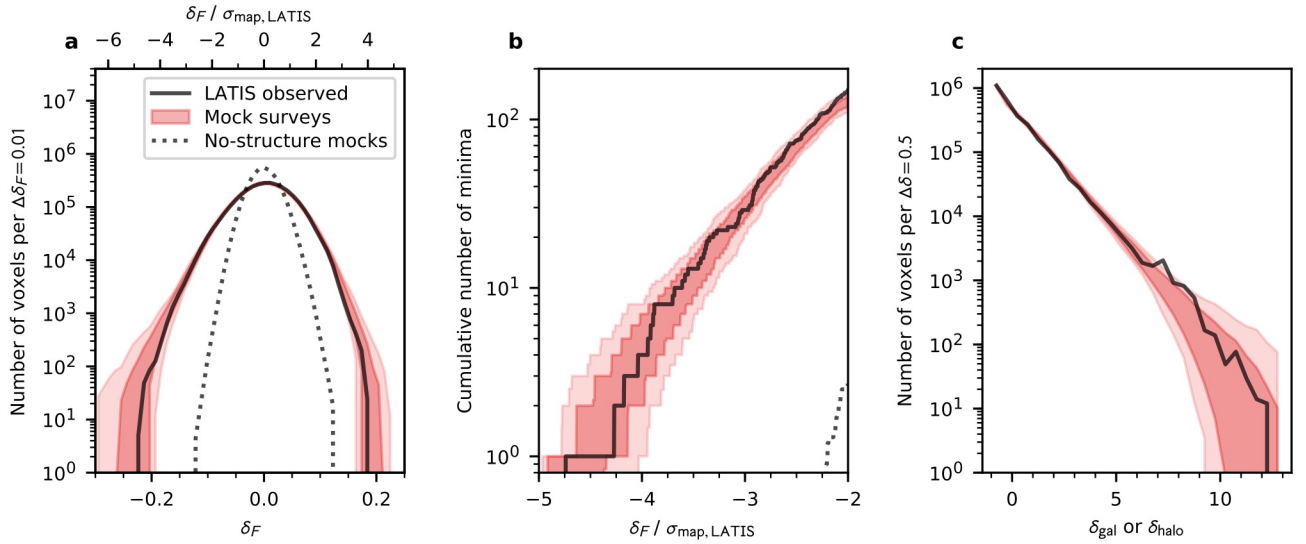
Additional Information

Correspondence and requests for materials should be addressed to Andrew Newman (anewman@carnegiescience.edu).

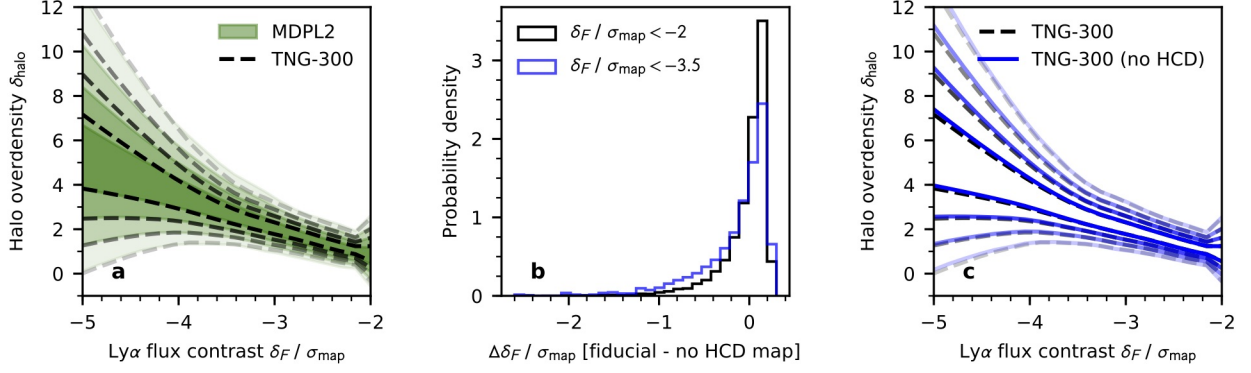
Reprints and permissions information is available at www.nature.com/reprints.



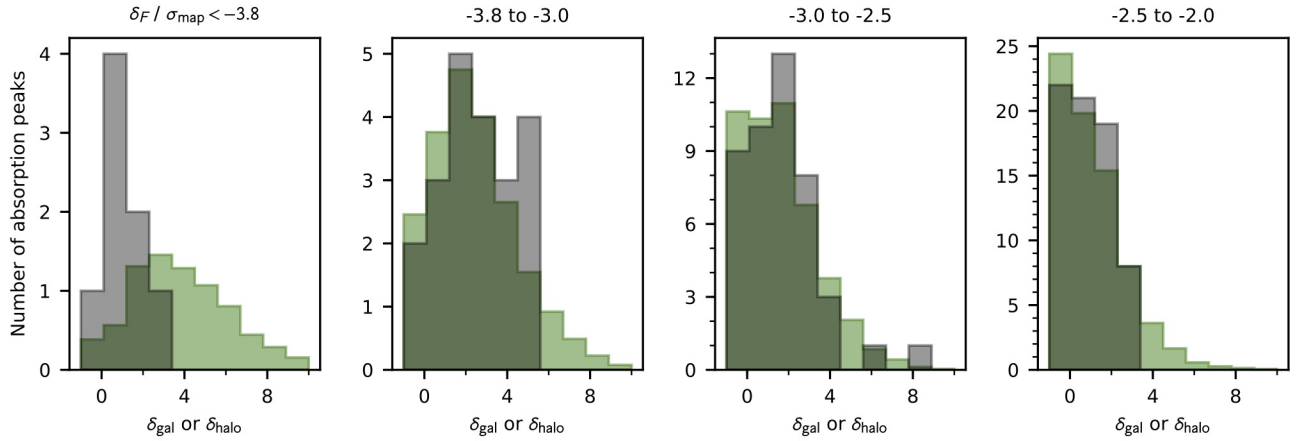
Extended Data Fig. 1 | The projected correlation function $w_p(R)$ of LATIS galaxies. We consider galaxies at $2 < z < 3$, evaluating w_p separately in the COSMOS and CFHTLS-D1 fields, and compare to the correlation function of dark matter halos with $M_{\text{vir}} > M_{\text{min}}$ for several values of M_{min} indicated in the legend. To reduce the range, the dimensionless $w_p(R)/R$ is scaled by $(R/h^{-1}\text{cMpc})^{1.5}$. Error bars show 1 s.e. Poisson uncertainties.



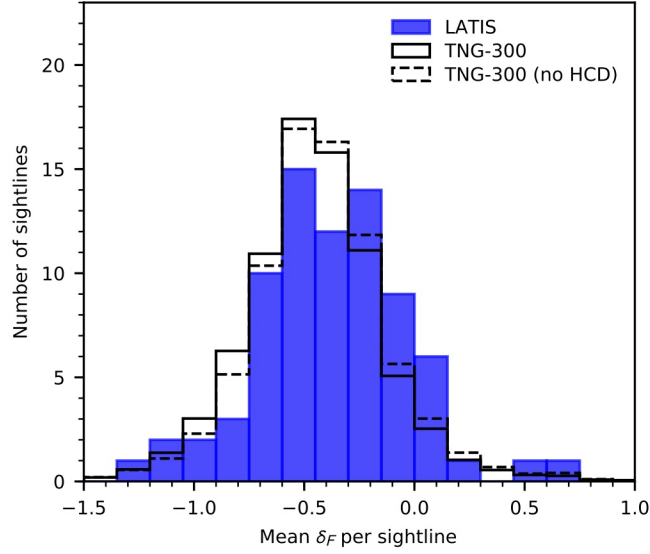
Extended Data Fig. 2 | Comparison of the statistical properties of the LATIS maps and the mock surveys. **a**, The distributions of the flux contrast δ_F , evaluated at each $(1 h^{-1} \text{cMpc})^3$ map voxel, in the LATIS maps (black curves) and the MDPL2 mock surveys (red bands enclose 68% and 95% of realizations). The dashed curve shows the expected distribution from observational noise alone, as estimated from mock surveys of structureless volumes, i.e., with $\delta_F = 0$ everywhere. **b**, The cumulative number of absorption peaks in the LATIS maps and the mock surveys. **c**, The distributions of the galaxy and halo overdensities, evaluated at each map voxel in the observed and mock maps. Altogether the mock surveys accurately match the observed distributions of δ_F and δ_{gal} individually.



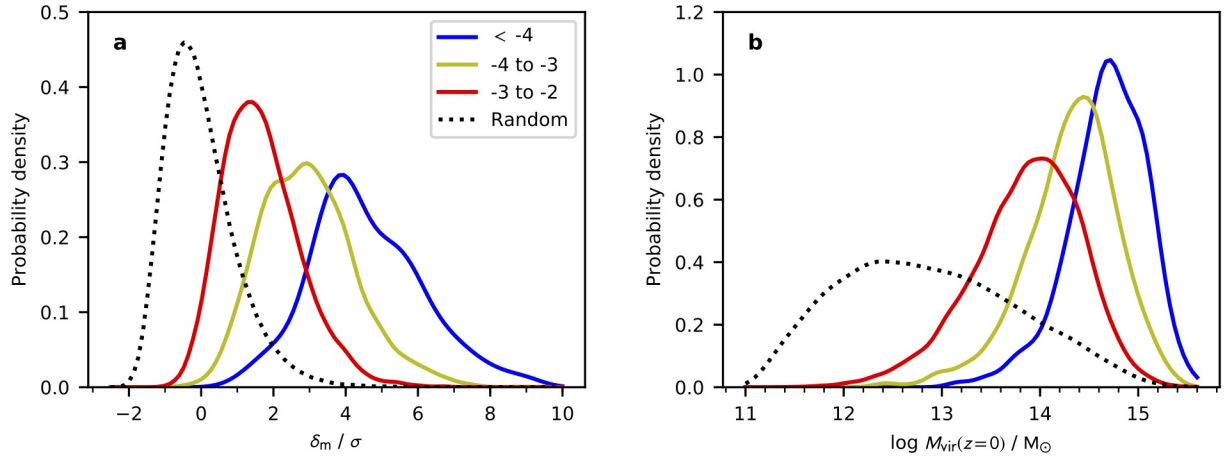
Extended Data Fig. 3 | Evaluation of the robustness of the $\delta_F - \delta_{\text{halo}}$ connection. **a**, Comparison of the trendline distributions, following Fig. 2, derived from mock surveys based on the MDPL2 (dark matter only) and IllustrisTNG300 (including hydrodynamics, galaxy formation, and HCD lines) simulations. For each simulation, curves show percentiles equivalent to 1, 2, 3, and 4 s.d. in a normal distribution. **b**, The distribution of differences in $\delta_F / \sigma_{\text{map}}$ between absorption peaks in the fiducial TNG mock surveys and the corresponding peaks in mocks that explicitly exclude HCD lines ($N > 10^{17.2} \text{ cm}^{-2}$). **c**, Comparison of the trendline distributions, following panel **a**, derived from the fiducial TNG mock surveys and those that exclude HCD lines.



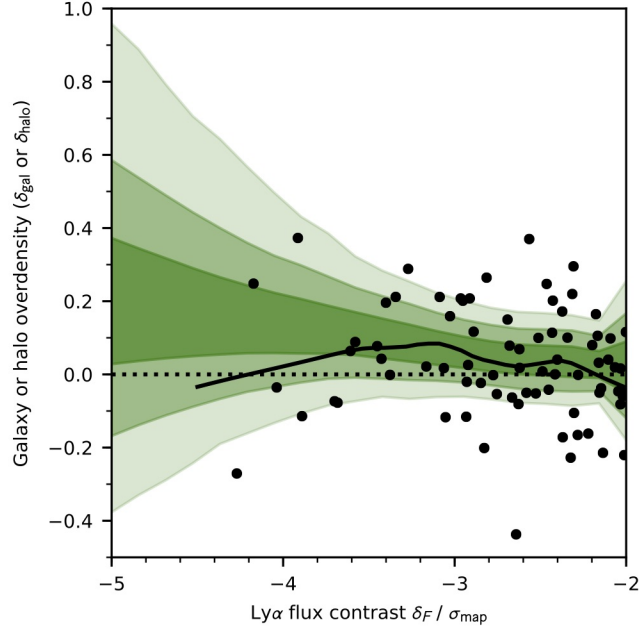
Extended Data Fig. 4 | Distributions of the galaxy or halo overdensity around Ly α absorption peaks. Gray histograms show δ_{gal} observed around Ly α absorption peaks within an $8 h^{-1} \text{ cMpc}$ aperture following Fig. 2. Green histograms show the distributions of δ_{halo} in the MDPL2 mock surveys, normalized to match the integrals of the corresponding gray histograms. The absorption peaks are split into four bins of δ_F , highlighting the discrepancy in the strongest absorption peaks.



Extended Data Fig. 5 | Evidence of widespread absorption near the strongest absorption peaks. The distribution of absorption in the 77 sightlines probing the 8 strongest absorption peaks ($\delta_F/\sigma_{\text{map}} < -3.8$) within $4 h^{-1}$ cMpc is shown and compared to the TNG mock surveys. Along each sightline, δ_F is averaged within $|\Delta z| < 4 h^{-1}$ cMpc of the absorption peak. The TNG distributions are scaled to match the integral of the LATIS histogram.



Extended Data Fig. 6 | The connection of absorption peaks with the matter overdensity and descendant halos. **a**, The distribution of the matter density contrast δ_m , smoothed with a Gaussian kernel having $\sigma = 4 h^{-1}$ cMpc and expressed in units of its standard deviation, at the location of absorption peaks in the mock surveys. Bins of $\delta_F/\sigma_{\text{map}}$ are indicated in the legend. The dotted curve shows the global distribution. **b**, The analogous distributions of the $z = 0$ descendant halo mass $M_{\text{vir}}(z = 0)$.



Extended Data Fig. 7 | The $\delta_F - \delta_{\text{gal}}$ trend evaluated using photometric redshifts. For each LATIS absorption peak in the COSMOS field, a black point indicates the galaxy overdensity estimated using photometric redshifts³². A trendline (black curve) analogous to Fig. 2 is overlaid on the distribution of trendlines derived from mock surveys (green), demonstrating the large uncertainties at low δ_F . Bands indicate percentiles equivalent to 1, 2, and 3 s.d. in a normal distribution.

Name	R.A. (deg)	Decl. (deg)	Redshift	$\delta_F / \sigma_{\text{map}}$	δ_{gal}
LATIS1-D2-0	150.103	2.175	2.562	-4.27	-0.53
LATIS1-D2-1	149.532	1.988	2.460	-4.17	2.91
LATIS1-D2-2	149.691	2.175	2.679	-4.04	0.13
LATIS1-D2-3	150.347	2.274	2.457	-3.91	1.97
LATIS1-D2-4	150.033	2.175	2.685	-3.89	0.88
LATIS1-D1-0	36.155	-4.534	2.568	-3.94	1.22
LATIS1-D1-1	36.220	-4.353	2.455	-3.88	0.79
LATIS1-D4-0	334.102	-17.731	2.594	-4.74	0.73

Extended Data Fig. 8 | Coordinates and properties of the 8 strongest LATIS absorbers having $\delta_F / \sigma_{\text{map}} < -3.8$. Using the mock surveys, we estimate 1σ errors of 2.8 arcmin in sky position (per coordinate), 0.002 in redshift, 0.6 in $\delta_F / \sigma_{\text{map}}$, and 0.9 in δ_{gal} .