Article

A population of ultraviolet-dim protoclusters detected in absorption

Andrew B. Newman1,2, Gwen C. Rudie1, Guillermo A. Blanc1-2, Mahdi Qezelzadeh1-2, Simeon Bird4, Daniel D. Kelson1, Victoria Pérez2, Enrico Congiu2, Brian C. Lemaux4,5, Alan Dressler1 & John S. Mulchaey1

Galaxy protoclusters, which will eventually grow into the massive clusters we see in the local Universe, are usually traced by locating overdensities of galaxies. 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, which we use to locate a substantial number of candidate protoclusters at redshifts 2.2 to 2.8 through their intergalactic gas. We find that the structures producing the most absorption, most of which were previously unknown, contain surprisingly few galaxies compared with the dark-matter content of their analogues in cosmological simulations. Nearly all of the structures 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 on the evolution of these galaxies that reduced their star formation or increased their dust content.

We have mapped Lyman-α (Lyα) absorption over 1.5 degree2 and the redshift interval z = 2.2–2.8 via the Lyman-α Tomography IMACS Survey (LATIS), conducted using the Inamori-Magellan Areal Camera and Spectrograph (IMACS) on the Magellan Baade telescope at the Las Campanas Observatory. The survey comprises deep spectra of a dense sample of Lyman-break galaxies and quasars. We reconstructed the three-dimensional distribution of Lyα absorption imprinted in these spectra by neutral hydrogen in the intergalactic medium (IGM), a technique known as Lyα or IGM tomography. This absorption traces the density field on large scales, and the resulting maps are expected to be particularly effective at locating protoclusters. LATIS spans three extragalactic deep fields and exceeds the volume of the pioneering COSMOS Lyα Mapping and Tomography Observations survey (CLAMATO) by a factor of ten, making it suited to the discovery of rare structures such as protoclusters. The key survey products are three-dimensional maps of the Lyα flux contrast $\delta_F = \langle F \rangle / \langle F \rangle - 1$, smoothed to a resolution of 4 h^-1 co-moving Mpc (cMpc; $h$ is the reduced Hubble constant), where $\langle F \rangle$ is the continuum-normalized Lyα flux (Fig. 1), along with 2,241 secure redshifts of Lyman-break galaxies and quasars within the same volumes. These galaxies are selected on the basis of either their photometric redshifts or their optical colours and r-band magnitudes $r < 24.8$. Their median rest-frame ultraviolet (UV) luminosity density is roughly 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, we identified these overdensities as local minima in the smoothed Lyα 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 $\sigma_{\text{map}}$ is dominated by IGM structure rather than observational noise, which is typically 1.8-times smaller.) We interpret the nature of these absorption peaks as resulting from 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. On the basis of 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 (descendant halo mass $M_{z=0}/M_\odot > 10^{14}$) 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 $\delta_{\text{peak}}$ of LATIS galaxies within 8 h^-1 cMpc. The connection between the galaxy overdensity and Lyα absorption is shown in Fig. 2. As the absorption increases (δF more negative), the galaxy overdensity initially increases as expected. However, the relationship is not monotonic; the

1Observatories of the Carnegie Institution for Science, Pasadena, CA, USA. 2Departamento de Astronomía, Universidad de Chile, Las Condes, Chile. 3Department of Physics and Astronomy, University of California, Riverside, Riverside, CA, USA. 4Department of Physics and Astronomy, University of California, Davis, Davis, CA, USA. 5Gemini Observatory, NSF’s NOIRLab, Hilo, HI, USA.
6e-mail: anewman@carnegiescience.edu
strongest absorption peaks, which should correspond to the highest matter overdensities, show a statistically significant but surprisingly small galaxy overdensity ($\delta_{\text{gal}} = 1$), on average.

To further examine this unexpected result, we calculated the overdensity of dark-matter halos $\delta_{\text{halo}}$ in the mock surveys at the location of each simulated absorption peak. We considered halos with virial 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 Lyα 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 with $\delta_{\text{gal}}/\delta_{\text{map}} \geq -3.5$. However, this agreement breaks down in the strongest absorption peaks, where the observed $\delta_{\text{gal}}$ becomes smaller than $\delta_{\text{halo}}$, in all but $p = 0.005\%$ of the mock surveys; such a significant difference occurs at any value of $\delta_{\text{gal}}$ 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 $\delta_{\text{gal}}$ is most evident in the eight strongest absorption peaks in LATIS with $\delta_{\text{gal}}/\delta_{\text{map}} < -3.8$. According to their IGM signature, 94\% of these are protoclusters (Methods). Yet only the two with the highest $\delta_{\text{gal}}$ are definitely associated with previously known protoclusters. 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 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 cMpc. These and other arguments (Methods) make it clear that contamination by high-column-density absorbers associated with galaxies, rather than megaparsec-scale overdensities, is not significant. Likewise, locally enhanced ionization (for example, from quasar radiation) could reduce Lyα absorption, but it cannot account for a low galaxy abundance within strong absorption peaks.

We conclude that most of the strongest Lyα 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 $\delta_{\text{gal}}$ (Fig. 3), we find that the dearth of such galaxies is most prominent towards the centres 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 colours 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. At lower redshifts $z \leq 1$, the effect of the environment is mainly to modulate the mix of galaxy populations (for example, star-forming and quiescent galaxies) rather than the properties of galaxies with one class. 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 protoclusters and that optical and UV-dark galaxies are the dominant type at high stellar masses $M \geq 10^{10.5} M_\odot$ and redshifts $z > 3$. There is also evidence for a boosted fraction of quenched or quenching 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, 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 galaxy overdensity $\delta_{\text{gal}}$ at the location of Lyα absorption peaks.

Black points indicate the observed absorption peaks, colored with a representative 1-s.d. error bar shown. More negative $\delta_{\text{gal}}$ indicates a stronger absorption. The black trendline is compared with the distribution of absorbers (green bands indicate percentiles equivalent to 1 s.d., 2 s.d., 3 s.d. 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 $\delta_{\text{halo}}$ within $\sim 8 \, h^{-1} \text{cMpc}$ virial radius and half-depth equal to 8 $h^{-1} \text{cMpc}$. Overdensities are measured within a cylinder that has a transverse radius and half-depth equal to $8 \, h^{-1} \text{cMpc}$. In the strongest absorptions (left of vertical line), $\delta_{\text{gal}}$ is much lower than expected.

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 highly sparse for UV-faint galaxies. Moreover, to substantially affect $\delta_{\text{gal}}$, a ‘missing’ population must be abundant and thus cannot have stellar masses far exceeding the LATIS galaxies’ median $M^* = 10^{12} M_\odot$ (ref. 33). 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/submillimeter Array (ALMA) will soon test whether submillimetre-selected galaxies reside in two of the strongest LATIS absorption peaks. As our findings pertain to the average properties of absorption peaks, which individually have substantial uncertainties in $\delta_{\text{gal}}$, definitive follow-up observations must encompass an ensemble of such systems.

The only previous survey to systematically examine the galaxy contents of Lyα-absorption-selected protoclusters reported high galaxy overdensities in three systems (ref. 13). Although 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 a UV-dim population emerges only with a larger sample. Furthermore, both the Lyα absorption and the galaxy overdensity were measured quite differently in the earlier studies: the absorption was measured in 4 sightlines across 20 $h^{-1} \text{cMpc}$, an areal density 15-times lower than that of LATIS, and different tracers (Lyα and Hα emitters) were used to measure $\delta_{\text{gal}}$. The latter may prove to be an illuminating distinction.

Despite many searches, a significant population of massive protoclusters at $z = 2.5$ seems to have been missed because their galaxy populations are UV-dim. Studies of galaxy evolution in various $z = 2–3$ 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 Lyα absorption and the density field on scales of several co-moving megaparsecs, a cornerstone of modern cosmology that is thought to be well understood.

Online content
Any methods, additional references, Nature Research 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/s41586-022-04681-6.
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 (quality flag zflag = 3 or zflag = 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 Lyα absorption in transverse sightlines35. 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 $zR$ along the line of sight. Redshift was converted to line-of-sight distance assuming a pure Hubble flow, so like the $\delta_g$ maps, galaxy densities are measured in velocity space. In Fig. 2, we count galaxies within a $R = 8 h^{-1}\text{Mpc}$ cylinder, which optimizes the signal-to-noise ratio in $\delta_g^N$ based on the simulations outlined below. In Fig. 3, we count galaxies differentially in a series of nested cylinders, which are labelled by the mean radius $R$ 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 $\bar{\rho}$ was evaluated within each survey field and was found not to vary systematically with redshift within the map volume. The galaxy overdensity is then $\delta_g^N = N/\bar{\rho} - 1$, where $\bar{\rho}$ is the volume excluding unobserved portions (that is, those outside the survey footprint or close to bright stars). The space density of LATIS galaxies is $\bar{\rho} = 7 \times 10^{-4} h^{3}\text{Mpc}^{-3}$ corresponding to an average $\bar{\rho} = 2.4$ 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 a root mean square of 10%, when measured in apertures of size $R = 8 h^{-1}\text{Mpc} = 6.8'$ and accounting for the survey mask. These variations are five-times smaller than the typical uncertainty in $\delta_g^N$ and are not correlated with $\delta_g$, so they are safely neglected. Geometric constraints on slitmask design, which arise from the requirement that galaxies’ spectra do not overlap, induce only a negligible bias in $\delta_g$. 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 $\delta_g$ 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 $\pm 5\%$. Finally, although the target selection probability is dependent on $r$-band magnitude3, its distribution is consistent for galaxies located within absorption peaks binned according to $\delta_F$; thus, trends in target selection probability cannot affect trends with $\delta_F$ (Figs. 2, 3).

Mock surveys

We created a suite of mock LATIS surveys within the 1-Gpc$^3$ MDPL2 simulation. The flux field at $z = 2.535$, near the LATIS midpoint, was calculated using the fluctuating Gunn–Peterson approximation4,8, which we previously found to accurately reproduce the distribution of fluctuations in the observed maps3. 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 (that is, 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 regularization49 and masked lines with high equivalent width6 to suppress damped absorbers, following the same methods applied to the observations. A mock tomographic map was created using the dachshund Wiener filtering code50 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 $M_{\text{vir}} > 10^{13.8} M_\odot$ (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 Fig. 2 shows that the mock surveys match the observed distributions of $\delta_g^N$ and $\delta_g$ with impressive accuracy.
The \( \delta_{\gamma} - \delta_{\text{gal-halo}} \) connection

To produce the distribution of the 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 \( \delta_{\gamma} \) as the LATIS data, to correctly model the uncertainty in the mean (\( \langle \delta_{\text{gal}} \rangle \)) conditional on \( \delta_{\gamma} \). For each such subset, we calculated a non-parametric trendline using locally weighted scatterplot smoothing (LOESS)\(^{35,37} \), a form of local linear regression. The same method was applied to the observations. The 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 \( \delta_{\gamma} \), owing to the declining number of absorption peaks. Varying the method used to define the trendlines (for example, varying the fraction of datapoints used in the local fitting between 30% and 70%, or using a second-order 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-\( \delta_{\gamma} \) data point is similarly inconsequential. The mock surveys based on MDPL2 (used in Figs. 2, 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 \( \delta_{\text{gal}} \) and \( \delta_{\text{halo}} \) for individual absorption peaks, in several bins of \( \delta_{\gamma} \), 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 \( \delta_{\text{gal}} \), so that the width of the bands indicates the additional sources of error (for example, bin-to-bin scatter from errors in \( \delta_{\gamma} \)) 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\(^{-1}\) Mpc (1 s.e.) uncertainty in the absorption peak position, which depresses \( \delta_{\text{gal}} \) and \( \delta_{\text{halo}} \) in the inner bins.

High-column density absorption lines

HCD lines, including Lyman-limit systems and damped Ly\( \alpha \) 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 surveys\(^{23,24} \) but it is expected to be much reduced in tomographic surveys such as LATIS, in which each map resolution element is probed by about ten sightlines. We filter lines with high equivalent width to mask DLAs, and the same procedure\(^{3} \) 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 \( \delta_{\gamma} = 0 \), 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. As the trendline distributions in the TNG- and MDPL2-based mocks are very similar (Extended Data Fig. 3a), particularly in the low-\( \delta_{\gamma} \), low-\( \delta_{\text{halo}} \), 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 (\( N_{\text{HI}} > 10^{17.2} \) 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, that is, 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, as the frequency of HCD lines in the TNG spectra is 30% ± 9% lower than observed\(^{13} \) (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 \( \delta_{\gamma} \) and \( \delta_{\text{halo}} \).

Connection to the matter overdensity at \( z = 2.5 \) and massive halos at \( z = 0 \)

The mock surveys allow us to connect the observed \( \delta_{\gamma} \) to the matter density contrast \( \delta_{\gamma} \) at the observed epoch, smoothed on a \( 4 \) h\(^{-1}\) Mpc scales like the flux limit. Extended Data Fig. 6 shows that the absorption peaks correspond to progressively larger matter overdensities as \( \delta_{\gamma} \) decreases, reaching 4.6σ overdensities, on average, when \( \delta_{\gamma}/\delta_{\gamma}\text{phot} < -4 \). We also evaluated the connection to massive halos at \( z = 0 \) by finding the most massive halo within 4 h\(^{-1}\) Mpc of each absorption peak and tracing the mass of its descendant, similar to earlier work\(^{2} \). Extended Data Fig. 6 shows that the distributions overlap substantially, but the strongest absorption peaks are always associated with protoclusters (\( M_{\text{vir}}(z=0)/M_{\text{vir}} > 10^{14} \)).

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\( \alpha \) absorption\(^{21} \), LATIS galaxies span about 1.3 dex in stellar mass\(^{22,23} \), which is thought to correspond to only about 0.6 dex in halo mass\(^{25} \). We find that the shape of the halo mass distribution, measured in our mock surveys within subvolumes defined by \( \delta_{\gamma} \), varies only slightly over this narrow mass range: even considering all halos with \( M_{\text{vir}} > 10^{14} M_{\odot} \), the median \( M_{\text{halo}} \) 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 catalogue\(^{26} \). By comparison with the LATIS spectroscopic redshifts, we found that the photometric redshift \( z_{\text{phot}} \) estimates have a random uncertainty of \( \sigma_{\text{phot}}(1+z) = 0.03 \) and a redshift-dependent bias reaching up to \( \Delta z/(1+z) = 0.03 \). We corrected for this bias in our comparisons. We calculated \( \delta_{\gamma} \) within a redshift interval of \( \pm 0.2 \) centred on each LATIS absorption peak, following an earlier \( z_{\text{phot}} \)-based protocluster search\(^{27} \), and within a transverse separation of 8 h\(^{-1}\) Mpc to match Fig. 2. We simulated the same observation in our mock surveys by considering halos with \( M_{\text{vir}} > 10^{14} M_{\odot} \), which match the space density of the \( z_{\text{phot}} \) sample, and perturbing their \( z \) coordinates...
of the catalogue that overlaps the LATIS maps, and it has the highest Hyperion16, of which components have previously been identified using 8 strongest LATIS absorption peaks (analysis, but the large \( z \sigma \) trendline in Fig. 2 is not notably affected. We conclude that spectroscopic trend in Fig. 2 is not inconsistent with the photometric redshift analysis, but the large \( z \sigma \) uncertainties prevent a strong test.

Counterparts of the strongest absorption peaks in the literature

In Extended Data Table 1, we provide the coordinates and properties of the 8 strongest LATIS absorption peaks (\( \delta F/\delta \text{map} < -3.8 \)) discussed in this paper. Among these, LATIS-D2-3 is part of a large proto-supercollider known as Hyperion16, of which components have previously been identified using galaxy surveys13,15 and Lyα tomography14. LATIS-D2-1 was previously identified (ID 20) in a catalogue of galaxy overdensities identified using photometric redshifts16; it is the highest such overdensity in the portion of the catalogue that overlaps the LATIS maps, and it has the highest \( \delta \text{map} \) listed in Extended Data Table 1. LATIS-D2-4 has a potential match (ID 25) in the same catalogue, 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

As the flux density of the Lyα 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 modelling 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). As 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 Lyα 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 Lyα absorption increases modestly in strength by 5%. This suggests that there may be a residual additive ‘pedestal’ around 5% of the mean Lyα 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 \( \delta \text{F} \). This should minimally affect our analysis, as \( \delta \text{F} \) is ultimately normalized by \( \sigma \text{map} \) 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 \( 6 \times 10^{-5} \) electrons per second per pixel, 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 and the fiducial maps, we find that the difference \( \delta \text{F}/\delta \text{map} \) 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

The 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 astropy44,65 and dachshund6. Other supporting analysis code is available from the corresponding author upon request.

Additional information

Correspondence and requests for materials should be addressed to Andrew B. Newman. Reprints and permissions information is available at http://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 $M_{\text{min}}$ indicated in the legend. To reduce the range, the dimensionless $w_p(R)/R$ is scaled by $(R/h \text{ cMpc})^{1.5}$. Error bars show $1 \sigma$ 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 $\delta_F$, evaluated at each $(1 \text{h}^{-1} \text{cMpc})^3$ map voxel, in the LATIS maps (black curves) and the MDPL2 mock surveys (red bands enclose 68% and 95% of realizations). The dashed curve show 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 $\delta_F$ and $\delta_{\text{gal}}$ 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 Lyα absorption peaks. Grey histograms show the distributions of $\delta_{\text{gal}}$ observed around Lyα absorption peaks within an 8 $h^{-1}$ cMpc aperture following Fig. 2. Green histograms show the distributions of $\delta_{\text{halo}}$ in the MDPL2 mock surveys, normalized to match the integrals of the corresponding gray histograms. The absorption peaks are split into four bins of $\delta_F$, 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 (\(\delta_{\text{F}}/\sigma_{\text{max}} < -3.8\)) within 4 \(h^{-1}\) cMpc is shown and compared to the TNG mock surveys. Along each sightline, \(\delta_{\text{F}}\) 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 $\delta_m$, 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 $\delta_m/\sigma_{\text{mean}}$ 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 redshifts. 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 $\delta_F$. Bands indicate percentiles equivalent to 1, 2, and 3 s.d. in a normal distribution.
### Extended Data Table 1 | Coordinates and properties of the 8 strongest LATIS absorption peaks having $\delta F/\sigma_{\text{map}} < -3.8$

| Name      | R.A. (deg) | Decl. (deg) | Redshift | $\delta F/\sigma_{\text{map}}$ | $\delta_{\text{gal}}$ |
|-----------|------------|-------------|----------|-------------------------------|------------------------|
| 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                   |

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 $\delta_{\text{gal}}$. 