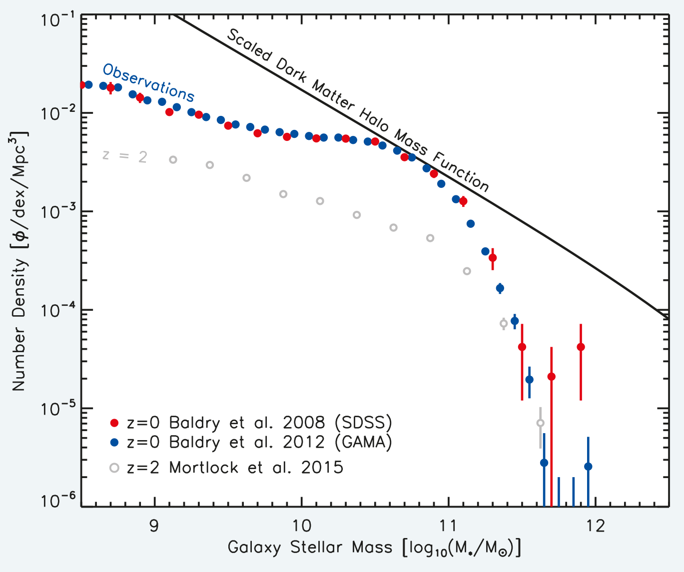
Star-forming galaxies in the first billion years

Rebecca Bowler describes how the hunt for the earliest bright star-forming galaxies has been boosted by ground-based surveys using the UKIRT and VISTA telescopes.

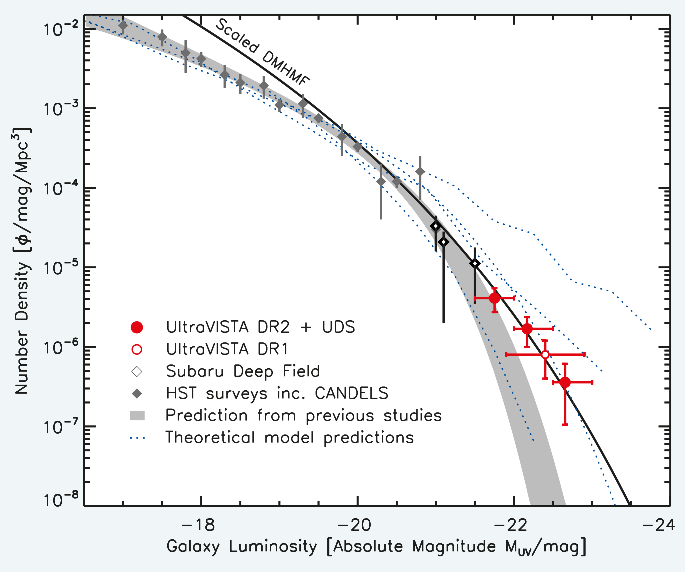
The discovery and characterization of galaxies in the first billion years is an exciting and rapidly moving field of astronomy. The observational frontier is constantly evolving and the past decade has seen the detection of hundreds of galaxies at redshifts exceeding 6. Until recently, however, samples have been limited to relatively faint galaxies found in extremely deep, but small area surveys such as the Hubble Ultra Deep Field (UDF). With new, wide-area surveys from ground-based telescopes it is now possible to probe larger cosmological volumes at high redshift and identify the first samples of early, highly star-forming galaxies. The prevalence and properties of these luminous galaxies provide a key test for the understanding of how rare, massive galaxies form at early times.

The formation of early structures

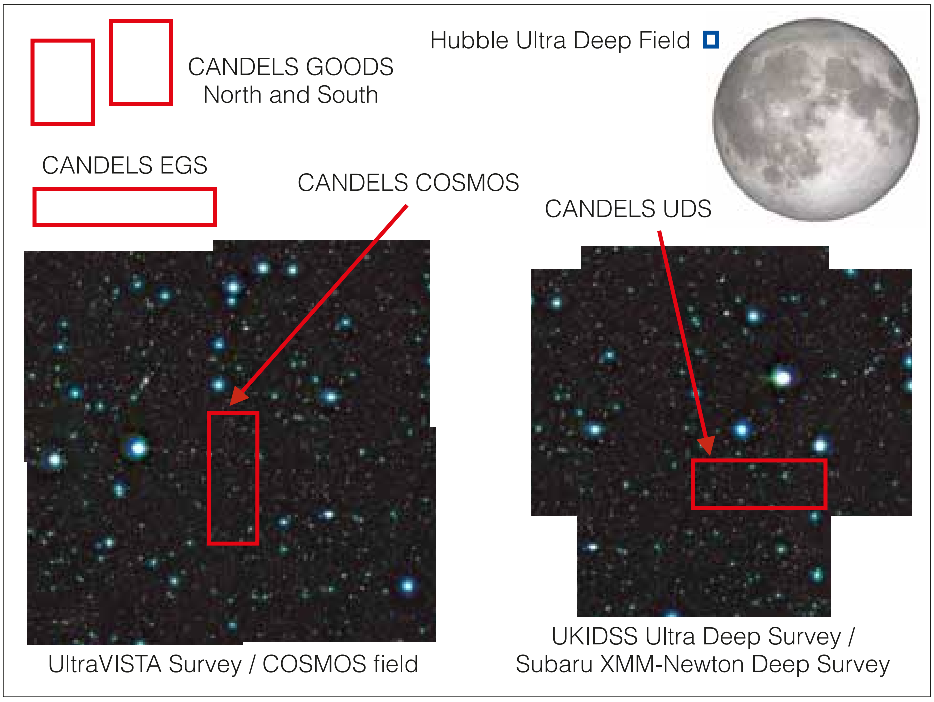
Detailed measurements of the cosmic microwave background from the Planck satellite and other facilities, coupled with the physics of gravitation, results in a tight prediction for the spectrum of mass (e.g. structures of dark and baryonic matter) and how this evolves with time. Because the abundance of dark matter dominates over baryonic matter by a factor of 5, the formation of gravitationally bound structures – "dark matter halos" – is frequently approximated by following the dark matter only, an approximation that appears to hold on large scales (Diemand & Moore 2011). Large gravitationally bound objects are very rare in the early universe, but their prevalence increases with time as smaller halos coalesce and merge, forming progressively more massive structures. As



1 The observed galaxy mass function at $z=0$ from the Sloan Digital Sky Survey (SDSS; Baldry *et al.* 2008) and the Galaxy and Mass Assembly Survey (GAMA; Baldry *et al.* 2012). Overplotted (solid line) is the scaled dark-matter halo-mass function, multiplied by a factor of ~0.05 in order to match the observed galaxy stellar-mass function around the knee of the curve. A clear deficit in the number of galaxies at the low- and high-mass end is visible, postulated to be due to strong feedback effects quenching star-formation in both low- and high-mass halos. The observed shape continues at least to $z \sim 2$, as demonstrated here using the results from Mortlock *et al.* (2015) shown in grey.



2 The galaxy luminosity function at $z=7$, adapted from Bowler *et al.* (2014), showing the number density of galaxies as a function of luminosity (or absolute magnitude, brighter galaxies are to the right). Results from predominantly HST surveys are taken from Bouwens *et al.* (2011) and McLure *et al.* (2013) and are shown in grey. Results based on using the SDF from Ouchi *et al.* (2009) are shown as black open diamonds. The open red circles show the preliminary results from the UltraVISTA survey (Bowler *et al.* 2012), which were later confirmed using deeper data from UltraVISTA and including the UDS (Bowler *et al.* 2014). The deviation of our results from the extrapolation of previous data, shown in grey, has important repercussions for galaxy formation models (dotted blue lines, see Bowler *et al.* 2015), suggesting that dust obscuration or feedback processes are not yet efficient. As in figure 1, the scaled dark-matter halo-mass function is shown as the solid black line.



3 A comparison of the sky area covered by several HST and ground-based fields used in the selection of high-redshift galaxies. The Hubble Ultra Deep Field, covering an area of 4.5 arcmin^2, is shown in the upper right as a small blue square. The five CANDELS fields, which are formed of multiple pointings of HST/WFC3, are shown in red. The ground-based UltraVISTA/COSMOS and UDS/SXDS fields are shown as the large irregular shapes, where the outline is caused by the overlap of different camera footprints used to produce the required optical and near-infrared data.

baryonic matter, the fuel for star formation, is gravitationally bound to the dark matter, the most massive (and arguably most luminous) galaxies should also be rare and indeed observations of both the mass and luminosity functions of galaxies do show such a trend (see figures 1 and 2). Nevertheless, the observed functions do not directly follow predictions based on a linear correspondence between stellar mass and dark-matter halo mass, as you would expect if all galaxies produced stars simply according to the abundance of fuel with a constant efficiency. Instead, a deficit in the number of low- and high-mass/luminosity galaxies is observed, resulting in a shallower slope at the low-mass end and a steeper (often parameterized as exponential) cut-off in the observed number densities of the most massive and luminous galaxies (e.g. Baldry *et al.* 2012). These deviations from the dark-matter-only predictions show that there must be additional baryonic physics having a strong impact on the proportion of the available fuel in each halo that can be converted into stars.

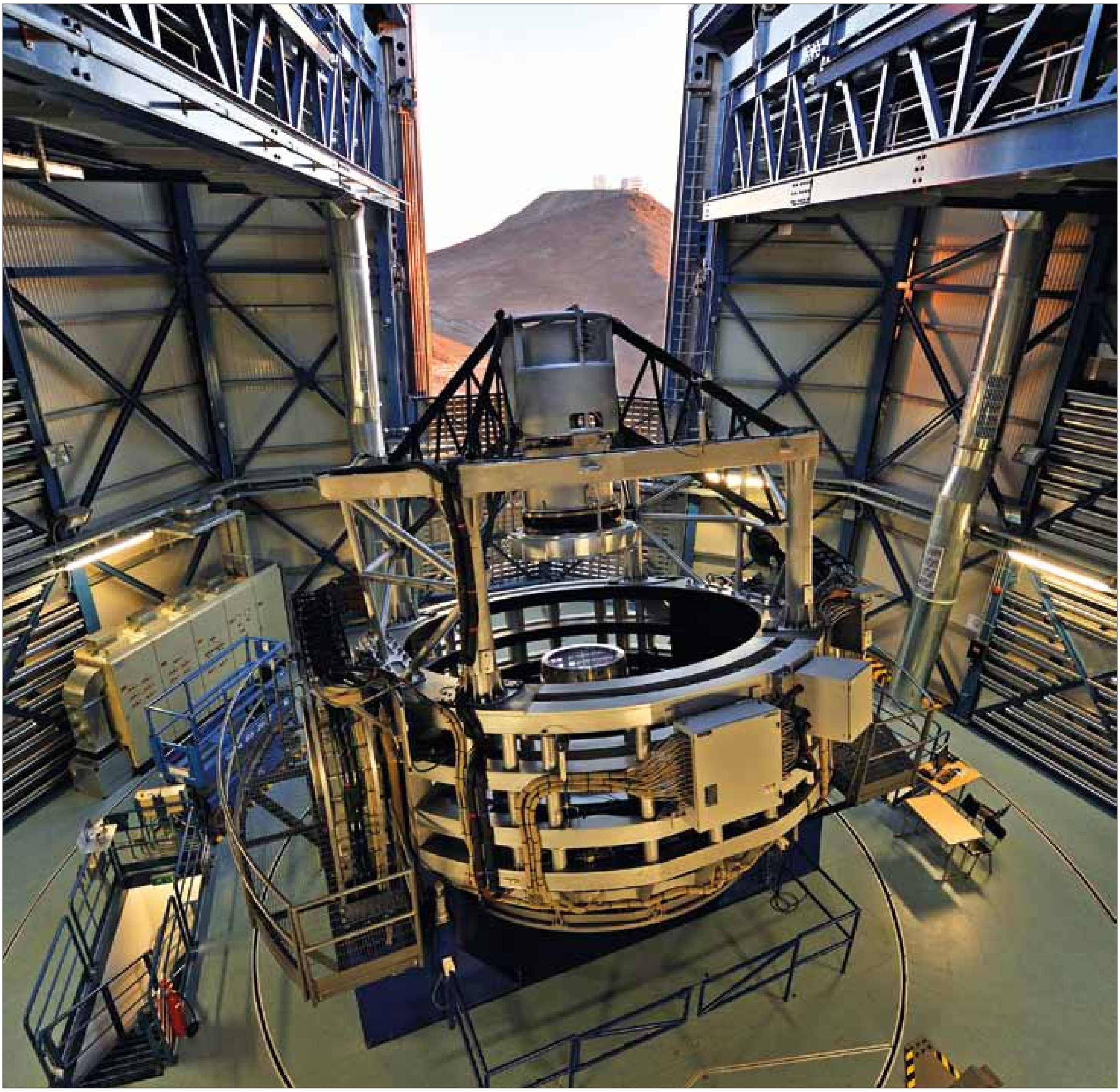
The creation of the first stars and galaxies has a dramatic effect on the chemical, thermal and dynamical structure of the baryonic matter, with various astrophysical processes feeding energy and metals into the surrounding medium (Silk & Mamon 2012). In low-mass galaxies, the energy produced in stellar processes – such as stellar winds and supernovae explosions – is sufficient to accelerate gas to velocities exceeding the escape velocity. It is postulated (although not yet possible to observe directly) that the formation of stars in low-mass galaxies is "quenched" or suppressed as a result of these stellar feedback processes, hence causing the deficit of observed objects with low masses and luminosities. Quenching of star formation in the most massive objects could be explained theoretically by the inability of hot virialized gas to cool sufficiently for star formation to take place, or by strong feedback from accretion onto a black hole (e.g. Finlator *et al.* 2011). Whereas stellar feedback happens rapidly (on timescales of <100 Myrs) the timescales for quenching star formation in the most massive halos are longer. For example, if black-hole feedback is the crucial quenching process, a time delay in the feedback effect would be expected because of the finite time required to build up a sufficiently massive central black hole, and therefore a relatively higher abundance of high-mass galaxies would be expected at preceding times. The importance of these processes can be tested by accurately measuring the shape and evolution in the mass and luminosity functions at high redshift.

"To observe at higher redshifts requires wide-area imaging in the red and near-infrared"

The role of ground-based observations

While much of the recent progress in exploring the high-redshift universe has been driven by the Hubble Space Telescope, this instrument is limited by its small field of view (4.5 arcmin^2, or 0.0013 deg^2, for the Wide Field Camera 3; see figure 3). The result is samples of galaxies dominated by fainter and more numerous objects. To select the brightest galaxies, which can have number densities less than 1000 times that of the galaxies typically selected in the UDF, imaging surveys that cover a larger area on the sky are required. Optical and near-infrared cameras on the leading ground-based telescopes now typically have a field-of-view of between 0.5 and 1 deg^2 on the sky, allowing wider area, albeit shallower, surveys to be performed more efficiently than with the HST. Indeed, wide-area optical surveys from the Canada–France–Hawaii Telescope and Subaru Telescope have provided samples of bright high-redshift Lyman-break galaxies (galaxies where star formation is in progress) up to $z \sim 5$ (e.g. van der Burg *et al.* 2010, Yoshida *et al.* 2006). However, to push observations to even higher redshifts and into the first billion years of the universe, wide-area imaging in the red optical (i.e. the z-band) and, crucially, the near-infrared are required.

Absorption in the Earth's atmosphere, predominantly by the OH molecule, reduces the sensitivity of ground-based observations at the red-end of the spectrum and in the near-infrared, and also limits the photometric observations to certain well-defined filters corresponding to gaps in the absorption. In the past decade, the first sufficiently deep near-infrared data for the selection of $z > 6$ galaxies have been taken using the UKIRT and VISTA (figure 4) telescopes, forming the UKIRT Ultra Deep Survey (UDS; Lawrence *et al.* 2007) and the UltraVISTA survey (McCracken *et al.* 2012). The two ground-based fields form the top-most tier of a wedding cake of public near-infrared surveys, where the base of the cake consists of extremely wide-area surveys covering a substantial fraction of the sky to a shallow depth (LAS, VHS) with progressively smaller but deeper surveys forming the upper layers. The ultra-deep surveys have required a substantial investment in integration time, with the full UltraVISTA survey requiring ~1400 hours



4 The Visible and Infrared Survey Telescope for Astronomy (VISTA) is a 4.1 m telescope sited at the Paranal Observatory in Chile. Run by ESO, it is the largest operational survey telescope providing imaging in the near-infrared, using its single camera, VIRCAM, to survey the southern sky from 1500 m high in the Atacama Desert. VIRCAM is sensitive from 0.9–2.3 μm and consists of 16 separate detectors in a 4×4 grid, which cover a field-of-view of 1.65 deg^2 if images are combined. The UltraVISTA survey (McCracken *et al.* 2013) forms the deepest of six ongoing public surveys that have been allocated 75% of the telescope's operation time, and will eventually have comparable depth to the HST CANDELS survey but over $\gtrsim$3 times the area. The UltraVISTA data overlaps with imaging from previous programmes, in particular in the optical and far-infrared taken as part of the Cosmic Evolution Survey (COSMOS; Scoville *et al.* 2007). (G Hüdepohl/ESO)

to produce images of comparable depth to the CANDELS survey from HST (Grogin *et al.* 2011), covering $\gtrsim$10 times the area. The resulting data, when combined with the pre-existing wealth of multiwavelength imaging available in the fields, extending from X-ray to radio wavelengths, has revolutionized not only the study of bright high-redshift galaxies, but also opened up new parameter space at lower redshifts (such as probing the masses of galaxies at $z \sim 2$, e.g. Mortlock *et al.* 2015).

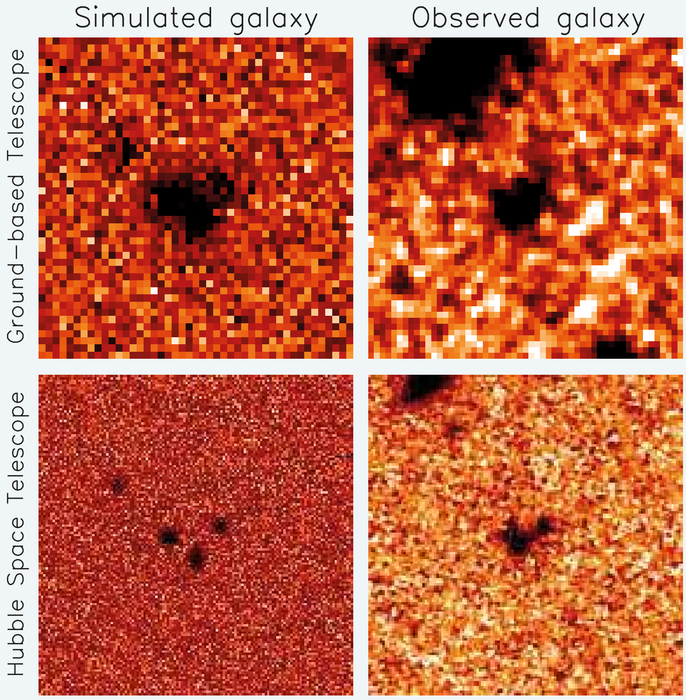
The first study to use the wide-area near-infrared imaging at $z > 5$ was performed using early data from the UK Infrared Telescope Infrared Deep Sky Survey (UKIDSS) UDS in the Subaru XXM-Newton Deep Survey field (SXDS; McLure *et al.* 2006, 2009). Although not deep enough to provide continuum detections for the majority of galaxies, the early near-infrared data proved essential for cleaning the samples of contaminant galaxies at lower redshift (where the Balmer break mimics the Lyman break), by excluding very red galaxies with spectral energy distributions (SEDs) that continue to rise through the near-infrared bands. A key result of the study, and later work by Willott *et al.* (2013) using data from the CFHT Legacy Survey, was that the bright-end of the luminosity function at $z \sim 6$ appeared to drop off steeply, as seen at lower redshifts, following an exponential decline defined by the Schechter function. The deviation observed at $z \sim 6$ from the underlying dark-matter distribution suggested that the onset of any high-mass quenching processes must have taken place earlier, further strengthening the motivation for comparable studies at even higher redshifts.

Although the near-infrared data were an

1 Unveiling galaxy sizes and morphologies with the Hubble Space Telescope

Despite the success of the UltraVISTA and UDS surveys in finding the brightest galaxies at $z \sim 7$, one disadvantage of ground-based telescopes becomes apparent when trying to study the sizes and morphologies of distant galaxies. Ground-based seeing (defined as full-width at half-maximum of a point source in the final image) is typically 0.8 arcsec in the optical and near-infrared, overwhelming the true size of the galaxy which is typically ~0.1 arcsec (half-light radius). In contrast, HST/WFC3 can provide close to diffraction-limited images with a resolution of <0.2 arcsec. Measurements of the sizes of $z > 6$ galaxies in HST surveys fields (e.g. the UDF and CANDELS images) have shown them to be compact, with half-light radii <1 kpc. The data have also shown some evidence for a size–luminosity relationship as found at lower redshift. However, the existence of such a relationship, which would suggest an early onset of the processes required to form the relation such as feedback and/or mergers, is controversial (Jiang *et al.* 2013) due to the limited luminosity range probed by HST-based samples (Ono *et al.* 2013).

To investigate the sizes and morphologies of the brightest $z \sim 7$ galaxies, we have been using the high-resolution imaging capabilities of Hubble to follow up the $z \sim 7$ galaxies selected from the ground-based UltraVISTA and UDS imaging. The programme will eventually produce high-resolution images of 20 of the brightest $z \sim 7$ galaxies presented in Bowler *et al.* (2014), allowing the first measurements of the sizes, morphologies and merger fractions of these extreme galaxies. The data will also shed light on the importance of mergers and interactions for galaxies at the bright end of the luminosity function at $z \sim 7$, as previous studies at $z \sim 6$–7 have suggested that brighter galaxies appear more disturbed than fainter objects (Jiang *et al.* 2013). Preliminary results suggest that the brighter galaxies are indeed significantly larger than fainter objects, with several galaxies showing multiple components suggestive of interactions or mergers (see figure 5). A detailed analysis of any biases in the measurement of sizes and morphologies for the bright sample are ongoing, but the initial results strongly suggest that interactions may play a crucial role in the evolution of the brightest galaxies in the first billion years.



5 Images of bright galaxies in the first billion years from a galaxy formation simulation (left column, Dayal *et al.* 2013) and from observations (right column), where the real galaxy shown in the right-hand images was selected in the UltraVISTA data (Bowler *et al.* 2012, 2014) and was imaged with the HST as part of our follow-up campaign as described in the text. The upper and lower images show the same galaxy as observed from ground-based telescopes, where the atmosphere blurs the image, and observed with the Hubble Space Telescope respectively. The improved resolution obtainable from space reveals that these bright galaxies are formed of multiple components, potentially in a violent merging system. Ongoing Hubble Space Telescope follow-up imaging of 20 galaxies from the Bowler *et al.* (2014) sample will allow the morphologies and sizes of these extreme objects to be uncovered for the first time.

important part of the selection of galaxies at $z \sim 6$, the real breakthrough provided by the availability of deep near-infrared surveys was the ability to detect the first robust samples of very bright $z = 7$ galaxies. These objects are z-band or optical-dropout galaxies and hence were undetectable in the previously available ground-based images. The UltraVISTA and UDS surveys both provide data in the near-infrared J, H and K filters from 1.1–2.5 μm, which sample the rest-frame UV continuum light from $z = 7$ galaxies and, crucially, deep imaging is also now available in the Y-band filter at ~1 μm. The Y-band filter nestles against the red optical z-band filter, providing a sensitive measure of the position of the Lyman break and allowing the clean separation of contaminant populations such as red low-redshift galaxies and cool galactic brown dwarfs, which although red, cannot reproduce the steep Lyman break when sampled by the z, Y and J filters.

Previously, the best wide-area imaging for the selection of $z \sim 7$ galaxies had come from the Subaru Deep Field (Ouchi *et al.* 2009, see figure 2), a 0.5 $\deg^2$ field imaged in the Y-band using the red-sensitive CCDs of the Subaru Telescope. While the results had insufficient dynamic range to look at the shape of the luminosity function in detail, and suffered from an uncertain contamination rate due to the lack of deep near-infrared imaging at $\lambda > 1\,\mu m$, they did, however, provide an improved prediction of the number of $z \sim 7$ galaxies that should be expected over the UltraVISTA and UDS fields. Despite the unprecedented area and depth of the new surveys, the best-fitting Schechter function fit (with an exponential decline to bright magnitudes, shown in figure 2) to the previous imaging predicted at most one galaxy at $z \sim 7$ brighter than an absolute magnitude (M_{UV}) of –22.0 in the full 1.5 $\deg^2$ of imaging.

As with many large imaging programmes, the UltraVISTA and UDS data have been released gradually over the duration of the full survey, which can be several years. Surprisingly, even in the relatively shallow first data release (DR1) of the UltraVISTA survey, several $z \sim 7$ galaxies were present in the imaging. In fact, careful analysis of the data resulted in the discovery of a total of ten candidate $z \sim 7$ galaxies, with four robust candidates and several more tentative objects over 1 $\deg^2$ of imaging (Bowler *et al.* 2012, figure 2, open red circle). Although tantalizing, the luminosity function results were only an estimate, based on a small number of objects and without an in-depth analysis of any potential biases that could manifest in such a result, for example scattering or gravitational lensing effects. Only with the deeper second data release from UltraVISTA, and the inclusion of the independent UDS field, could a more robust measurement be made of the number density of these bright galaxies. The results confirmed an excess of galaxies above the previous best-fitting function (filled red circles), and whereas the commonly assumed Schechter function fit (shown as a black line in figure 2) would have predicted only a

2 Cosmic variance

Cosmic variance describes the field-to-field differences in derived number counts arising from the large-scale structure in the universe where, for example, the survey field could be imaging a galaxy cluster or alternatively a region of low density such as a void or filament. Multiple sight-lines on the sky are therefore essential to determine the genuine average space density of galaxies, with wider area surveys being less susceptible than small fields such as the UDF. The effects of cosmic variance can be clearly seen in the derived number counts of galaxies in the five independent CANDELS fields, which show 20% variance above the statistical errors (Bouwens *et al.* 2015). At $z\sim6$, the UltraVISTA field contains almost twice the number density of galaxies than the UDS, indicating that multiple >1 deg^2 fields are essential to accurately constrain the average space density of galaxies at high redshifts. Previous (e.g. CANDELS) and upcoming surveys have been designed to shed light on and mitigate the effect of cosmic variance, for example the continuing VISTA VIDEO survey (Jarvis *et al.* 2013) will image two additional degree-scale fields, providing a valuable comparison to the results obtained to-date in the UltraVISTA and UDS datasets.

handful of galaxies in the two fields, a total of 34 objects were detected. In particular, several galaxies (which were those found in the first epoch of imaging) were unusually bright and have become subject to follow-up campaigns with other facilities to determine their properties in more detail (see box 1). Instead of a steep exponential decline in the number densities of the brightest galaxies as observed at $z\leq6$, the observations instead showed a more gradual decline in the luminosity function, a decline which closely followed the underlying dark-matter halo distribution. The highly star-forming galaxies detected in UltraVISTA and UDS appeared not to have been quenched as similarly bright objects at lower redshift have been, hence suggesting that the quenching mechanism, be it accretion onto a supermassive black hole or some other process, was not yet efficient 800 Myrs after the Big Bang.

The bright end of the luminosity function

Studying the shape and evolution of the luminosity function and any relationship it has with the underlying distribution of dark matter holds the key to understanding the build-up of galaxies in the early universe. A new determination of the luminosity function at $z\sim6$ (Bowler *et al.* 2015) further highlights the epoch between $z=7$ and $z=5$ (corresponding to ~200 Myrs) as a period of rapid evolution in the observed number counts of galaxies. Not surprisingly, however, there is currently no simple interpretation of the observed evolution in the luminosity function of galaxies, and uncertainties remain in the bright end of the observed luminosity function as a result of newly discovered strong cosmic variance affects (see box 2). The role of black-hole feedback or other quenching mechanisms such as inefficient gas cooling have been shown in hydrodynamical and semi-analytic simulations to strongly affect the predicted number of bright galaxies; however, the observational constraints required to distinguish between different scenarios are lacking. Furthermore, the relationship between the baryonic and dark-matter mass of a galaxy and the observed galaxy luminosity are likely to be complicated, because observations of high-redshift galaxies tend to be limited to the rest-frame UV wavelengths which are sensitive to only recent star formation. The luminosity functions derived from rest-frame UV light are also very sensitive to dust within the galaxy, which strongly attenuates the light from young stars. Recent theoretical work has suggested that dust obscuration may in fact be a critical process in shaping the observed bright end of the rest-frame UV luminosity function (e.g. Cai *et al.* 2014).

"With ALMA, dust obscuration can now be measured directly for the first time"

While typical galaxies at $z>6$ have been assumed to be relatively dust-free, based on their observed blue colours (e.g. Dunlop *et al.* 2013), at lower redshift there exists a strong colour–luminosity relationship that implies that the brightest galaxies are redder than their fainter counterparts. This effect has been attributed to the build up of dust in the most luminous and massive galaxies. At high redshift, a colour–luminosity relation has been observed to at least $z=5$ (e.g. Rogers *et al.* 2014), with tentative evidence that it extends to higher redshifts. If the most luminous galaxies at high redshift are indeed the most dusty, the observed evolution at the bright end of the luminosity function between $z=7$ and 5 could be accounted for by the gradual build-up of dust in these galaxies, with the onset of such obscuration dictated by the dust-production timescales of asymptotic giant branch stars, for example. With the commissioning of the Atacama Large Millimetre Array (ALMA), dust obscuration can now be measured directly for the first time. Recent detections of dust emission from $z>6$ galaxies hint at an exciting future direction in the field, which has been dominated so far by measurements in the rest-frame UV.

Looking to the future, detailed measurements of the rest-frame optical spectra for $z>5$ galaxies will eventually be possible with the James Webb Space Telescope, giving an insight into the physical properties of galaxies at high redshift. The near-infrared sensitivity of JWST will, crucially, allow an accurate determination of galaxy masses at high redshift, thus providing a direct comparison between the stellar and dark-matter masses. On the ground, the addition of independent fields and deeper data from the continuing UltraVISTA survey will provide an increasingly detailed analysis of the shape and evolution of the rest-frame UV luminosity functions in the next decade, and will ultimately be surpassed by future surveys from the Euclid satellite and the Large Synoptic Survey Telescope. For example, the proposed Euclid deep survey will cover 40 deg^2 to the depth of UltraVISTA, providing samples of thousands of galaxies of similar luminosities to those discovered in the UltraVISTA survey, with HST-like resolution. ●

AUTHOR
Rebecca Bowler is a postdoctoral researcher at the Institute for Astronomy, University of Edinburgh, UK.

REFERENCES
Baldry I et al. 2008 *Mon. Not. Roy. Astron. Soc.* **388** 945
Baldry I K et al. 2012 *Mon. Not. Roy. Astron. Soc.* **421** 621
Bouwens R J et al. 2011 *Astrophys. J.* **737** 90
Bouwens R J et al. 2015 *Astrophys. J.* **803** 34
Bowler R A A et al. 2012 *Mon. Not. Roy. Astron. Soc.* **426** 2772
Bowler R A A et al. 2014 *Mon. Not. Roy. Astron. Soc.* **440** 2810
Bowler R A A et al. 2015 in press arXiv:1411.2976
Cai Z-Y et al. 2014 *Astrophys. J.* **785** 65
Dayal P et al. 2013 *Mon. Not. Roy. Astron. Soc.* **434** 1486
Diemand J & Moore B 2011 *Adv. Sci. Letts* **4(2)** 297
Dunlop J S et al. 2013 *Mon. Not. Roy. Astron. Soc.* **432** 3520
Finlator K et al. 2011 *Mon. Not. Roy. Astron. Soc.* **410** 1703
Grogin N A et al. 2011 *Astrophys. J. Supp.* **197** 35
Jiang L et al. 2013 *Astrophys. J.* **773** 153
Lawrence A et al. 2007 *Mon. Not. Roy. Astron. Soc.* **379** 1599
McCracken H J et al. 2012 *Astron. & Astrophys.* **544** A156
McCracken H J et al. 2013 *The Messenger* **154** 29
McLure R J et al. 2006 *Mon. Not. Roy. Astron. Soc.* **372** 357
McLure R J et al. 2009 *Mon. Not. Roy. Astron. Soc.* **395** 2196
McLure R J et al. 2013 *Mon. Not. Roy. Astron. Soc.* **432** 2696
Mortlock A et al. 2015 *Mon. Not. Roy. Astron. Soc.* **447** 2
Ono Y et al. 2013 *Astrophys. J.* **777** 155
Ouchi M et al. 2009 *Astrophys. J.* **706** 1136
Rogers A B et al. 2014 *Mon. Not. Roy. Astron. Soc.* **440** 3714
Scoville N et al. 2007 *Astrophys. J. Supp.* **172** 1
Silk J & Mamon G A 2012 *Research in Astron. Astrophys.* **12** (8) 917
van der Burg R F J et al. 2010 *Astron. & Astrophys.* **523** A74
Willott C J et al. 2013 *Astrophys. J.* **770** 8
Yoshida M et al. 2006 *Astrophys. J.* **653** 988