“GiGa”: the Billion Galaxy HI Survey — Tracing Galaxy Assembly from Reionization to the Present


School of Earth & Space Exploration, Arizona State University, Box 871404, Tempe, AZ 85287-1404, USA; Email: Rogier.Windhorst@asu.edu

Abstract.

In this paper, we review the Billion Galaxy Survey that will be carried out at radio–optical wavelengths to micro–nanoJansky levels with the telescopes of the next decades. These are the Low-Frequency Array, the Square Kilometer Array and the Large Synoptic Survey Telescope as survey telescopes, and the Thirty Meter class Telescopes for high spectral resolution+AO, and the James Webb Space Telescope (JWST) for high spatial resolution near–mid IR follow-up. With these facilities, we will be addressing fundamental questions like how galaxies assemble with super-massive black-holes inside from the epoch of First Light until the present, how these objects started and finished the reionization of the universe, and how the processes of star-formation, stellar evolution, and metal enrichment of the IGM proceeded over cosmic time.

We also summarize the high-resolution science that has been done thus far on high redshift galaxies with the Hubble Space Telescope (HST). Faint galaxies have steadily decreasing sizes at fainter fluxes and higher redshifts, reflecting the hierarchical formation of galaxies over cosmic time. HST has imaged this process in great structural detail to $z < 6$. We show that ultradeep radio-optical surveys may slowly approach the natural confusion limit, where objects start to unavoidably overlap because of their own sizes, which only SKA can remedy with HI redshifts for individual sub-clumps. Finally, we summarize how the 6.5 meter James Webb Space Telescope (JWST) will measure first light, reionization, and galaxy assembly in the near–mid-IR.

Keywords: High resolution imaging — distant galaxies — galaxy assembly — reionization — first light — Square Kilometer Array

PACS: 98.52.Sw, 98.54.-h, 98.54.Gr, 98.54.Kt, 98.70.Dk

INTRODUCTION

For this review paper, I was asked to review the “Billion Galaxy Survey” that will be carried out at radio–optical wavelengths to micro–nanoJansky levels with the telescopes of the next decades. These facilities are, for instance, the Low-Frequency Array (LOFAR, Röttgering et al. 2005), the Square Kilometer Array (SKA; Schilizzi et al 2004), and the Large Synoptic Survey Telescope (LSST; Tyson 2007) — which will be used as survey telescopes at radio optical wavelengths — and the Thirty Meter Telescopes (TMT; e.g., Nelson et al. 2006) — including e.g., the Giant Segmented Mirror Telescope, Giant Magellan Telescope, plus the EU Extremely Large Telescope — which will be used for high spectral resolution+ adaptive optics follow-up, as well as the James Webb Space Telescope (Mather and Stockman, 2000), which will provide high spatial resolution near–mid-IR imaging and low-resolution spectroscopy. With a combination of these facilities available in the next decade, we will be addressing fundamental questions like:

- (1) How do HI clouds at $z > 6$ assemble over cosmic time into the giant spiral and elliptical galaxies seen today?
- (2) How and why did the (dwarf dominated) galaxy luminosity function (LF) and mass function evolve with epoch?
- (3) In the context of the galaxy formation–AGN paradigm, how did supermassive black hole (SMBH) growth keep up with the process of galaxy assembly?
- (4) How does the central accretion disk feed the SMBH, and how are radio jets and lobes in radio galaxies and quasars produced as a result?
- (5) How did AGN feedback control the bright-end evolution of the galaxy LF, and how did SN feedback shape the faint-end of the LF from $z > 6$ to $z=0$?

These are some of critical science drivers for radio and optical telescopes of the next decade. Before we give a preview of possible answers to these questions, we will first briefly consider how the Hubble Space Telescope (HST) has revolutionized the topic of galaxy assembly in the last decade. One of the remarkable discoveries by HST was that the numerous faint blue galaxies are in majority late-type (Abraham et al. 1996, Glazebrook et al. 1995, Driver et al. 1995) and small (Odewahn et al. 1996, Pascalello et al. 1996) star-forming objects. These are the building blocks of the giant galaxies seen today. By measuring their distribution over rest-frame type versus redshift, HST has shown that galaxies of all Hubble types formed over a wide range of cosmic time, but with a notable transition around redshifts...
Towards higher redshifts, despite the over time to form the larger/massive galaxies that we see today (e.g., Navarro, Frenk, & White 1996). Subgalactic units rapidly merged from the end of reionization to grow bigger units at lower redshifts (Pascale et al. 1996). Merger products start to settle as galaxies with giant bulges or large disks around redshifts z=1 (Lilly et al. 1998, 2007). These evolved mostly passively since then, resulting in giant galaxies today, possibly because the epoch-dependent merger rate was tempered at z<1 by the extra expansion induced by Λ (Cohen et al. 2003, Ryan et al. 2007). To avoid caveats from the morphological K-correction (Giavalisco et al. 1996, Windhorst et al. 2002), galaxy structural classification needs to be done at rest-frame wavelengths longwards of the Balmer break at high redshifts (Taylor-Mager et al. 2007). JWST will make such studies possible with 0′′1–0′′2 FWHM resolution at observed near-IR wavelengths (1–5 μm), corresponding to the rest-frame optical–near-IR at the median redshift of faint galaxies (zmed≈1–2; Mobasher et al. 2007).

RADIO & OPTICAL SIZES AT THE FAINTEST FLUXES: NATURAL CONFUSION?

The HST/ACS GOODS survey (Ferguson et al. 2004) showed that the median sizes of faint galaxies decline steadily towards higher redshifts, despite the Θ–z relation that minimizes at z=1.65 in WMAP ΛCDM cosmology. While surface brightness (SB) and other selection effects in these studies are significant, this work suggests evidence for intrinsic size evolution of faint galaxies, where galaxy half-light radii rhl evolve with redshift as: rhl(z) = rhl(0)(1+z)^−s with s ≈ 1. This reflects the hierarchical formation of galaxies, where sub-galactic clumps and smaller galaxies merge over time to form the larger/massive galaxies that we see today (e.g., Navarro, Frenk, & White 1996).
Fig. 2 Median 1.4 GHz radio source size vs. $S_{1.4}$ flux density, as summarized by Windhorst et al. (1993, 2003). The dotted line indicates the median extrapolated to the rest-frame UV–optical disk sizes seen by HST at $AB < 30$ mag, or expected for JWST at $AB < 32$ mag. The pink line indicates the natural confusion limit (see Fig. 1), above which radio sources unavoidably start to overlap in ultradeep surveys, requiring HI information to disentangle them.

The HST/ACS Hubble UltraDeep Field (HUDF; Beckwith et al. (2006) showed that high redshift galaxies are intrinsically very small, with typical sizes of $r_{HL} \approx 0.12$ or 0.7–0.9 kpc at $z \approx 4–6$. A combination of ground-based and HST surveys shows that the apparent galaxy sizes decline steadily from the RC3 to the HUDF limits (Fig. 1 here; Odewahn et al. 1996; Cohen et al. 2003, Windhorst et al. 2006). At the bright end, this is due to the survey SB-limits, which have a slope of +5 mag/dex in Fig. 1. At the faint end, ironically, this appears not only to be due to SB-selection effects (cosmological (1+z)$^4$ SB-dimming), since for $B_J \gtrsim 23$ mag the samples do not bunch up against the survey SB-limits. Instead it occurs because: (a) their hierarchical formation and size evolution; (b) at $J_{AB} > 26$ mag, one samples the faint end of the luminosity function (LF) at $z_{med} \approx 2–3$, resulting in intrinsically smaller galaxies (see e.g., Fig. 4b; Yan & Windhorst 2004); and (c) the increasing inability to properly deblend faint galaxies at fainter fluxes.

This leads ultradeep surveys to slowly approach the “natural” confusion limit, where a fraction of the objects unavoidably overlaps with neighbors due to their finite object size (Fig. 1), rather than the finite instrumental resolution, which causes the instrumental confusion limit. Most galaxies at $J_{AB} \approx 28$ mag are likely unresolved point sources at $r_{HL} \approx 0.1$ FWHM, as suggested by hierarchical size simulations in Fig. 1 (Kawata et al. 2002). This is, amongst others, why such in the UV–optical–near-IR objects are best imaged from space, which provides the best point-source and SB-sensitivity. The fact that many faint objects remain unresolved at the HST diffraction limit effectively reduces the (1+z)$^4$ SB-dimming to a (1+z)$^2$ flux-dimming (with potentially an intermediate case for partially resolved objects, or linear objects that are resolved in only one direction), mitigating the incompleteness of faint galaxy samples. The trick in ultradeep surveys is therefore to show that this argument has not become circular, and that larger galaxies at high
redshift are not missed. Other aspects that compound these issues are size-overestimation due to object confusion, size-bias due to the sky-background and due to image noise.

Fig. 2 shows the median angular size of 1.4 GHz radio sources vs. $S_{1.4}$ flux from 100 Jansky to 30 $\mu$Jy (e.g., Windhorst et al. 2003). The SKA radio source sizes at 10–100 nJy are estimated from the HST galaxy size distribution to AB=30 mag in Fig. 1, where the object density reaches $\geq 10^6$ objects/deg$^2$ (see Fig. 3a). Fig. 1–2 clearly show that SKA must anticipate the small $\leq 0'$3 radio sizes of faint star-forming galaxies at all redshifts $z \geq 1$. As in Fig. 1, the purple line in Fig. 2 indicates the natural confusion limit due to intrinsic radio source sizes. Above this line, radio sources unavoidably start to overlap statistically, and would significantly impact catalog reliability and completeness, and — unless other measures are taken (such as disentangling the sources with very deep HI-imaging and redshifts, see below) — eventually fundamentally limit our ability to survey the sky to arbitrarily deep limits. The HI disk sizes of nearby dwarf galaxies are similar to their synchrotron radio continuum-sizes and their optical sizes (Deeg et al. 1997). Hence, Fig. 2 suggests that SKA needs to have baselines at $\sim 0'$3 FWHM to match the HI and radio continuum sizes that we expect from the galaxy disk sizes in the deepest HST images (Fig. 1).

THE BILLION GALAXY SURVEY AT RADIO AND OPTICAL WAVELENGTHS

To produce a billion galaxy survey, let us now consider what such a survey requires. Fig. 3a shows the HUDF galaxy counts to AB $\leq 30$ mag ($\leq 4$ nJy at 1 $\mu$m), and Fig. 3b shows the normalized differential 1.4 GHz radio source counts expected for SKA to $S_{1.4} \approx 1–10$ nJy. From the HUDF galaxy counts, we expect $\geq 2 \times 10^6$ galaxies/deg$^2$ to AB=31.5 mag (1 nJy at $\approx 1$ $\mu$m), which is the deepest level currently reachable with HST, and which will be routinely obtained with JWST (see below). A similar flux level of 1–10 nJy at GHz frequencies is the main goal for the SKA. Hopkins et al. (2000) show a simulated 12-hour SKA 1.4 GHz image with 1 deg$^2$ FOV that has FWHM $\leq 0'$1 and reaches a flux level of 100 nJy (5-$\sigma$). Similarly, a 1200 hr SKA integration would reach $\sim 10$ nJy. Fig. 3a shows that the normalized differential SKA source counts at nJy levels will be dominated by blue, mostly point-like or disk-shaped radio sources. These will likely primarily reside in star-forming galaxies at redshifts $z \approx 1–2$, where the cosmological volume maximizes in WMAP cosmology, while a minority will be — possibly more extended — radio sources that reside in earlier-type galaxies which host weak AGN (see the models in Fig. 3a).

To obtain a billion galaxies at either radio or optical wavelengths, one must thus at least survey $\sim 10^3$ deg$^2$, or a significant fraction of the high-latitude sky. Since neither HST, nor JWST, nor the TMT’s will have a large enough FOV to do so, this can only be done with SKA at radio wavelengths, and with the LSST at optical wavelengths, provided that the LSST will survey 2$\pi$ ster or $2 \times 10^5$ deg$^{-2}$ to its anticipated all-sky flux limit of AB$\approx 27$ mag. Fig. 1–3 show that at the radio flux level of $S_{1.4} \approx 1–10$ nJy, one expects 1 object in every box 2$''5 \times 2$"5. This means that one must always carry out SKA-surveys at nJy levels with sufficient spatial resolution to avoid object confusion, and that SKA must have coverage at significantly long baselines that yield FWHM $\leq 0'$3. For HST and JWST, this number is FWHM $\leq 0'$08, as discussed in Fig. 1. For SKA, this means practically that one must always obtain many HI line channels, so that one can disentangle overlapping continuum sources in redshift space, and find all the enclosed HI. Nature simply does not leave us an alternative.

In summary, with a dedicated Billion Galaxy Survey, SKA will measure how galaxies of all types assembled their HI and turned it into stars over a wide range of cosmic time: from $z \approx 6$ to $z=0$. The SKA HI and radio continuum sizes of $10^7–10^9$ $M_\odot$ starforming objects are small enough that with $\leq 0'$3 FWHM resolution, SKA will: (1) properly separate these objects from their neighbors without major source confusion; and (2) not resolve them in majority, somewhat mitigating the effects from cosmological SB-dimming. Further details on the future of HI surveys and their other main science goals are given by Drs. M. Haynes, J. Lazio, and S. Meyers (this Volume).

FIRST LIGHT, REIONIZATION & GALAXY ASSEMBLY WITH JWST

The James Webb Space Telescope (JWST) is designed as a deployable 6.5 meter segmented IR telescope for imaging and spectroscopy from 0.6 $\mu$m to 28 $\mu$m. After its planned 2013 launch (Mather & Stockman 2000), JWST will be automatically deployed and inserted into an L2 halo orbit. It has a nested array of sun-shields to keep its temperature at $\leq 40$ K, allowing faint imaging to AB $\approx 31.5$ mag ($\approx$1 nJy) and spectroscopy to AB $\approx 29$ mag in the near–mid-IR.

First Light: The WMAP polarization results imply that the Dark Ages which started at recombination ($z=1089$) lasted until the First Light objects started shining at $z \approx 20$, and that the universe was first reionized at redshifts at
z \approx 9-13, if reionization happened instantaneous, or as early as z \approx 11-20, if the reionization process was drawn out (Dunkley et al. 2008, Spergel et al. 2007). The epoch of First Light is thought to have started with Population III stars of 200-300 M_\odot at z > 10-20 (Bromm et al. 2003). Groupings of Pop III stars and possibly their extremely luminous supernovae should be visible to JWST at z \approx 10-20 (Gardner et al. 2006). This is why JWST needs NIRCam at 0.6-5 \mu m and MIRI at 5-28 \mu m. The First Light epoch and its embedded Pop III reionizing sources may have heated the IGM enough that it could not cool and form the IMF of the first Pop II stars until z < 8-10 (Cen 2003). The IMF of Pop II stars may have formed in dwarf galaxies with masses of 10^6-10^9 M_\odot with a gradual onset between z \approx 9 and z \approx 6. The reionization history may have been more complex and/or heterogeneous, with some Pop II stars forming in sites of sufficient density immediately following their Pop III predecessors at z > 10.

HST/ACS can detect objects at z > 6.5, but its discovery space A_\parallel \times D_\parallel \log(l) cannot trace the entire reionization epoch. HST/NICMOS similarly is limited to z \approx 8 and provides limited statistics. HST/WFC3 can explore the redshift range z > 7-8 with a wider FOV than NICMOS. Fig. 4b shows that with proper survey strategy (area and depth), JWST can trace the LF throughout the entire reionization epoch, starting with the first star-forming objects in the First Light epoch at z \approx 20, to the first star-forming dwarf galaxies at the end of the reionization epoch at z \approx 6. To observe the LF of First Light star-clusters and subsequent dwarf-galaxy formation may require JWST to survey GOODS-sized areas to AB \approx 31.5 mag (\approx 1 nJy at 10-\sigma; See Fig. 4b), using 7 filters for reliable photometric redshifts, since objects with AB > 29 mag will be too faint for spectroscopy.

Reionization: The HUDF data showed that the LF of z \approx 6 objects is potentially very steep (Bouwens et al. 2006, Yan & Windhorst 2004b), with a faint-end Schechter slope \alpha \approx 1.8-1.9 after correcting for sample incompleteness (Fig. 4a). Deep HST/ACS grism spectra confirmed that 85-93% of HUDF i-band dropouts to z_{AB} \approx 27 mag are at z \approx 6 (Malhotra et al. 2005). The steep faint-end slope of the z \approx 6 LF implies that many galaxies may have collectively provided enough UV-photons to complete reionization at z \approx 6 (Yan & Windhorst 2004a). This assumes that the Lyman continuum escape fraction at z \approx 6 is as large as observed for Lyman Break Galaxies at z \approx 3 (Steidel et al. 1999), which is reasonable — although not proven — given the expected lower dust content in dwarf galaxies at z \approx 6. Hence, dwarf galaxies, and not quasars, likely completed the reionization epoch at z \approx 6. The Pop II stars in dwarf galaxies therefore cannot have started shining pervasively much before z \approx 7-8, or no neutral H-I would be seen in the foreground of z \approx 6 quasars (Fan et al. 2003), and so dwarf galaxies may have ramped up their formation fairly quickly from z \approx 9 to z \approx 6.
HST ACS grism LF from the HST GRAPES and PEARS surveys to AB $\lesssim 27$ mag (Ryan et al. 2007). **Fig. 5b** Faint-end LF-slope or Schechter-$\alpha$ evolution (Ryan et al. 2007) from $z=0$ to $z=6$ for a variety of surveys, that reach several magnitudes below $L^*$ in the rest-frame FUV or rest-frame optical.

**Galaxy Assembly:** JWST will measure how galaxies of all types formed over a wide range of cosmic time, by accurately measuring their distribution over rest-frame optical type and structure as a function of redshift or cosmic epoch. HST/ACS has made significant progress at $z=6$, surveying very large areas (GOODS, GEMS, COSMOS), or using very long integrations (HUDF, Beckwith et al. 2006). Quantitative galaxy decomposition (Odewahn et al. 2002) can measure galaxy morphology and structure, and the the presence and evolution of bars, rings, spiral arms, and other structural features at higher redshifts (e.g., Jogee et al. 2004). Such techniques will allow JWST to measure the detailed history of galaxy assembly in the epoch $z=1–3$, when most of today’s giant galaxies were made. JWST will be able to do this out to $z\approx 10-15$ at least (see Windhorst et al. 2006), hence enabling to quantitatively trace galaxy assembly.

The red boundaries in Fig. 4b indicate part of the galaxy and QSO LF that a ground-based 8m class telescope with a wide-field IR-camera can explore to $z\approx 9$ and AB $\lesssim 25$ mag. A ground-based wide-field near-IR survey to AB $\lesssim 25–26$ mag can sample $L>>L^*$ galaxies at $z\lesssim 9$, which is an essential ingredient to study the co-evolution of supermassive black-holes and proto-bulges for $z\gtrsim 9$, and an essential complement to the JWST First Light studies. The next generation of wide-field near-IR cameras on ground-based 8–30 m class telescopes can do such surveys over many $\deg^2$ to AB$\approx 25–27$ mag, complementing JWST, which will survey GOODS-sized areas to AB $\lesssim 31.5$ mag (Fig. 4b).

HST has begun to measure the faint-end LF-slope evolution of the galaxy luminosity function. This Schechter $\alpha$-evolution is a fundamental issue, like the local low-mass end of the initial mass function (IMF). With accurate ACS grism+broad-band redshifts to AB $\lesssim 27$ mag, Ryan et al. (2007) measured the faint-end LF-slope at $z\approx 1$, and outlined the faint-end LF-slope evolution from $z\approx 0$ to $z=6$. As modeled by Khochfar et al.(2007), this provides new physical constraints to hierarchical formation theories. Star-formation and SN-feedback processes can further produce different faint-end slope-evolution (Khochfar et al. 2007). JWST will provide fainter spectra (AB $\approx 29$) and spectro-photometric redshifts to much higher $z$ ( $\gtrsim 15–20$), and will trace the faint-end LF-slope evolution for $z\approx 12$, hence constraining hierarchical formation theories.

Similarly, SKA, and perhaps LOFAR, will trace low-mass-end of the HI mass-function of dwarf galaxies and its $\alpha$-slope evolution for $z\approx 3–6$. This will allow us to measure environmental impact on faint-end LF-slope $\alpha$ directly. As known for nearby galaxy samples (Fig. 5a), significant scatter is likely due to different clustering environments or cosmic variance. We expect convergence to a slope $|\alpha|\approx 2$ predicted by hierarchical models at $z>6$, before the process of SN feedback starts destroying dwarf galaxies, or before the process of AGN feedback starts producing major outflows and AGN-driven winds in massive galaxies. With a combination of JWST and very deep SKA images, it may be possible to constrain onset of Pop III SNe epoch at $z \approx 8–10$, and the onset to the Type II & Type Ia SN-epochs at $z \approx 6$, and thereby constraining the major physical processes that regulate star-formation as function of cosmic time.
CONCLUSIONS

HST has led the study of galaxy assembly, showing that galaxies form hierarchically through repeated mergers with sizes growing steadily over time as $r_{hl}(z) \propto r_{hl}(0) \cdot (1+z)^{-s}$ and $s \approx 1$. The Hubble sequence thus gradually emerges at $z \approx 1$–2, when the epoch-dependent merger rate starts to wind down. The global onset of Pop II-star dominated galaxies ended the process of reionization at $z \approx 6$. High resolution rest-frame UV–optical imaging of high redshift galaxies is best done from space, because faint galaxies are small ($r_{hl} \leq 0^\prime.15$), while the ground-based sky is too bright and the PSF not stable enough to obtain good high-resolution images at faint fluxes (AB $\geq 27$ mag). JWST will extend these studies into the epoch of reionization and First Light, and trace galaxy SED’s in the rest-frame-optical for $z \approx 20$.

Such surveys will slowly approach the natural confusion limit (Fig. 1), where objects at the faintest (nJy) fluxes start to unavoidably blend with their neighbors, not because of instrumental resolution, but because of their own intrinsic sizes. SKA will be critical to help disentangle the non-negligible fraction of such objects at $\mu$Jy–nJy levels, and provide unique HI-redshifts for each component. For this, it will need to have beam-sizes as small as $0^\prime.3$ FWHM (see Fig. 1). In order to get the next radio facilities it wants, the science community will need to unify behind current & future radio facilities, such as SKA and LOFAR — building on proto-types that are currently being build (see this Vol.) — and define the essential synergy of SKA with other future facilities, such as the JWST, LSST, and the TMT’s.

ACKNOWLEDGMENTS

This work was supported by HST grants from STScI, which is operated by AURA for NASA under contract NAS 5-26555, and by NASA JWST grant NAG 5-12460. Other JWST studies are at: www.asu.edu/clas/hst/www/jwst/ .

REFERENCES