TRACING GALAXY ASSEMBLY: TADPOLE GALAXIES IN THE HUBBLE ULTRA-DEEP FIELD
Amber N. Straughn, Seth H. Cohen, Russell E. Ryan Jr, Nimish P. Hathi, Rogier A. Windhorst, & Rolf A. Jansen
Department of Physics and Astronomy, Arizona State University, Tempe, AZ 85281; amber.straughn@asu.edu

ABSTRACT

In the Hubble Ultra Deep Field (HUDF) an abundance of galaxies is seen with a knot at one end plus an extended tail, resembling a tadpole. These “tadpole galaxies” appear dynamically unrelaxed — presumably in an early merging stage — where tidal interactions likely created the distorted knot-plus-tail morphology. Here we systematically select tadpole galaxies from the HUDF and study their properties as a function of their photometric redshifts. In a companion HUDF variability study presented in this issue, Cohen et al. (2005) revealed a total of 45 variable objects believed to be Active Galactic Nuclei (AGN). Here we show that this faint AGN sample has no overlap with the tadpole galaxy sample, as predicted by recent theoretical work. The tadpole morphology — combined with the lack of overlap with the variable objects — supports the idea that these galaxies are in the process of an early-stage merger event, i.e., at a stage that likely precedes the “turn-on” of any AGN component and the onset of any point-source variability. We show that the redshift distribution of tadpole galaxies follows that of the general field galaxy population, indicating that — if most of the tadpole galaxies are indeed dynamically young — the process of galaxy assembly generally kept up with the reservoir of available field galaxies as a function of cosmic epoch. These new observational results highlight the importance of merger-driven processes throughout cosmic history, and are consistent with a variety of theoretical and numerical predictions.

Subject headings: galaxies: formation — galaxies: mergers — galaxies: active galactic nuclei — cosmology

1. INTRODUCTION

The origin of disk galaxies has long been thought to occur through the process of dissipational collapse in a Cold Dark Matter (CDM) universe (White & Rees, 1978). Within this paradigm, hierarchical clustering (Navarro, Frenk, & White 1997) produces dark matter halos in which dissipational collapse of the residual gas occurs. The resulting disks retain the kinematic information of their host dark matter potential wells (Blumenthal et al. 1986). Recent numerical simulations have resolved some long-standing discrepancies in the standard dissipational collapse scenario by including previously-neglected energetic feedback from central supermassive black holes during galaxy merging events (e.g. Robertson et al. 2005). In particular, they emphasize the relationship between the central black hole mass and the stellar velocity dispersion, which confirms the link between the growth of black holes and their host galaxies (di Matteo, Springel, & Hernquist 2005; Springel, di Matteo, & Hernquist 2005ab), which predict a stage of merger-driven galaxy evolution that closely resembles these tadpole galaxies in a distinct phase that does not yet show AGN activity (as discussed further in Section 7). In particular, this morphology appears to represent an early stage in the merging of 2 nearly-equal mass galaxies. We systematically selected galaxies displaying this knot-plus-tail morphology from the HUDF; a representative sample of tadpoles is shown in Figure 1 (details of sample selection are given in Section 3). All the selected tadpole galaxies contain the asymmetric, pointlike source with a diffuse tail morphology, some with multiple knots; all of which we believe are undergoing recent interactions. They are mostly linear structures, some resembling the “chain” galaxies first reported by Cowie, Hu, & Songaila (1995). When more than two clumps come together, these objects may be more akin to the luminous diffuse objects and clump clusters (Conselice et al. 2004; Elmegreen, Elmegreen, & Sheets 2004; Elmegreen, Elmegreen, & Hirst 2004), or other types of irregular objects (van den Bergh 2002). Elmegreen et al. (2005) visually classify 97 HUDF galaxies (down to 10 pixels in size) as “tadpoles” and 126 as “double-clump.” Some of the galaxies classified by Elmegreen et al. as “double-clump” were identified as tadpoles by our code, due either to the unresolved nature of one clump (which would have been detected as a “tail” in our analysis) or to the diffuse nature of one end of the object. Since our goal in selecting these tadpoles was to sample galaxies that had recently undergone interaction,
inclusion of some of these “double-clumps” is warranted. One high redshift object in our sample has been studied in detail by Rhoads et al. (2005). A few objects with multiple knots are detected by our selection software, but the majority have a single prominent knot with an extended tail.

In this paper, we present the photometric redshift distribution of tadpole galaxies, and compare it with the redshift distribution of the general field galaxy population. This paper is organized as follows. In §2, we describe the HUDF data, and in §3 the tadpole sample selection. In §4, we discuss why the majority of tadpole galaxies are likely not chance alignments, but instead mostly dynamically young objects. In §5, we discuss their redshift distribution, in §6 their relation to galaxy assembly, and in §7 their possible relation to AGN growth.

2. HUBBLE ULTRA DEEP FIELD DATA

The Hubble Ultra Deep Field is a 400 orbit survey in four filters carried out using the Advanced Camera for Surveys (ACS) aboard the Hubble Space Telescope (HST) of a single field centered on RA(J2000)=03h32m39.0s, Dec(J2000)=−27°47′29″. The 144-orbit F775W (i′) image is deepest, followed by F850LP (i′); 47 orbits), F606W (V; 56 orbits), and F435W (B; 56 orbits). The HUDF reaches ∼1.0 mag deeper in B and V and ∼1.5 mag deeper in i′ than the equivalent filters in the Hubble Deep Field (Williams et al. 1996). From the ∼10,000 objects detected in the HUDF (Koekemoer 2004), we will select the sample of tadpole galaxies and analyze their properties using the i′-band image, because it provides the highest sensitivity of the four filters. Yan & Windhorst (2004b) discuss how this results in a bias against objects at z ≥ 5.5. This bias is small, and only concerns the high redshift tail of the redshift distribution. Note, however, that tadpole galaxies at z ≥ 5.5 do exist (e.g., Rhoads et al. 2005).

3. TADPOLE SAMPLE SELECTION

The first step in this analysis is to systematically select the galaxies that have the characteristic tadpole shape. We selected sources in the F775W (i′) band to \( I_{AB} = 28.0 \) mag using SExtractor (Bertin & Arnouts 1996). The objects of interest all have a bright “knot” at one end with an extended “tail” at the other. SExtractor selects objects from an image based on different input parameters, and adjusting them results in the desired selection of sources. The crucial parameter at this stage is DEBLEND_MINCONT, which governs the manner in which nearby peaks in flux are considered part of a single object and thus are counted as one source. With the deblending parameter set to a high value, SExtractor will separate nearby flux maxima into separate sources. In contrast, when the deblending is set to a low value, the program will count the nearby maxima largely as one source. Two different source catalogs are thus generated: the highly deblended catalog will contain many point-like sources, including the knots of potential tadpole galaxies.

The catalog with low-deblending will contain extended sources, including the tadpole galaxies’ tails. The catalogs contain many more sources than the desired ones, and the correctly shaped objects must be selected from these two initial catalogs. The desired tadpole galaxies have a nearly unresolved knot or concentration, and an extended tail, so these types of sources must be selected from the initial catalog, and related spatially such that they represent real objects. All of the following procedures were performed using IDL.

Both input SExtractor catalogs, described above, contain the following information about the selected sources: x and y pixel locations, length of the semi-major and semi-minor axes (a and b) of SExtractor ellipses, and the angle (θ) of the semi-major axis from north through east. The following input parameters were calibrated using a training set of tadpole galaxies that were manually selected by visual inspection of portions of the HUDF. First, the knots of the tadpole galaxies were selected by setting an axis-ratio limit. A “knot” was defined to be a source from the highly deblended catalog with an axis ratio greater (i.e., rounder) than some critical value (in our case, b/a > 0.70). In the same way, the “tails” needed to be elongated objects, so a similar procedure was performed on the objects from the catalog with low-deblending, but with the criterion that their b/a < 0.43. The two new lists of correctly shaped objects had to be related physically on the image, thus, a new set of objects was defined where a knot was within a certain distance of the geometrical center of a tail. This distance was taken to be < 4a (in semi-major axis units of the tail). We also required that the knot be at least > 0.1a from the tail’s geometrical center, since we are searching for asymmetric objects, and want to eliminate upfront as many of the true edge-on mid-type spiral disks as possible. The objects also must have the knot near one end of the tail, and this was accomplished by selecting only those tails and knots that had a relative angle θ — measured with respect to the semi-major axis of the tail — that was ≤ 20°. This step prevented including knots and tails that were close together on the image, but not physically part of the same galaxy.

These selection criteria provided a list of the tadpole-shaped galaxies. The final number of tadpoles selected depended on the selection program’s input parameters for the limiting axis-ratios, the distance at which the knots and tails were considered related, and the angle difference between the knot and the semi-major axis of the tail. The values of these parameters are given in Table 1. With these values, the tadpole galaxy selection program detected 154 sources total. This sample was then refined as follows.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>b/a limit: knots</td>
<td>&gt;0.70</td>
</tr>
<tr>
<td>b/a limit: tails</td>
<td>&lt;0.43</td>
</tr>
<tr>
<td>Distance to center (in a-axis units)</td>
<td>&lt;4</td>
</tr>
<tr>
<td>Angle difference θ (tail-knot)</td>
<td>≤ 20°</td>
</tr>
<tr>
<td>Total number of tadpoles automatically selected</td>
<td>154</td>
</tr>
<tr>
<td>DEBLEND_MINCONT (knots)</td>
<td>0.000005</td>
</tr>
<tr>
<td>DEBLEND_MINCONT (tails)</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 1 lists SExtractor parameters used in the tadpole sample selection. b & a are semi-major and semi-minor axis lengths of the ellipses used in selecting tadpole knots and tails. DEBLEND_MINCONT (knots/tails) is the SExtractor parameter that governs in what manner nearby peaks in flux are considered part of a single object and are thus counted as one source.
A large majority of the 154 tadpole galaxy candidates selected had the characteristic elongated knot-plus-tail morphology, although there were some anomalies. In total, 14 (9%) obvious mis-detections were visually rejected because they were very faint, on the edge of the image, or in the outskirts of large face-on spiral galaxies, where both knotty and diffuse regions are common and spatially close together. A visual examination of the field also produced 25 more tadpole galaxies not found by the selector program due to the inability of SExtractor to correctly separate particular point-like sources within these galaxies. These extra selected tadpole objects visually obeyed the morphological criteria that were used to define the main sample. Our total final sample thus contains 165 tadpole galaxies, a subset of which is shown in Figure 1.

In our final sample, less than 10% of the selected tadpoles appear as normal edge-on disk galaxies; the vast majority have the highly asymmetric morphology. In terms of visual vs. automatic selection, we find our sample to be about 91% (140/154) reliable and about 86% (154/179) complete. The final set of 165 tadpoles galaxies will now be studied as a separate class of dynamically unrelaxed objects and compared to the general field galaxy population in the HUDF.

4. WHY TADPOLE GALAXIES ARE NOT CHANCE ALIGNMENTS

In this section we demonstrate that these tadpole galaxies are likely not chance alignments of tails and unrelated knots. We first select all elongated diffuse structures (“tails”) in the HUDF, and then measure the angle $\theta$ of the nearest off-centered knot within a radius $r \leq 4a$ ($\leq 2''$). Chance alignments of unrelated tails and knots would show a random distribution of angles; however, Figure 2 shows that there clearly is an excess of knots near $|\theta| \simeq 0^\circ$. The excess peak contains 154 knots, while the average number of knots with $|\theta| \geq 10^\circ$ is $\sim 15$ per $5^\circ$-bin. Figure 2 thus shows a significant overabundance of knots near the end of elongated diffuse structures as compared to randomly distributed knots. Hence, this physically meaningful result suggests that the majority of tadpole galaxies are not just chance alignments of unrelated knots. Instead, we believe they are mostly linear structures which are undergoing interactions. When compared to models of galaxy mergers (di Matteo, Springel, & Hernquist 2005; Springel, di Matteo, & Hernquist 2005; Robertson et al. 2005, Hopkins et al. 2005), these objects strongly resemble dynamically young objects in the early stages of merging.

5. THE REDSHIFT DISTRIBUTION OF TADPOLE GALAXIES

To investigate the occurrence of tadpole galaxies throughout the history of the universe, we calculate photometric redshifts of all HUDF galaxies to $i'_{AB} = 28.0$ mag. All photometric redshifts were calculated from the HUDF $BViz$ photometry using HyperZ (Bolzonella et al. 2000). In order to investigate associated redshift errors, we compared our photometric redshifts to published spectroscopic redshifts for CDFS 70 objects. We find an rms scatter of 0.15 for the fractional photometric redshift error $\delta = (\text{photoz-specz})/(1+\text{specz})$ if all 70 objects are included, and 0.10 when we reject a few of the most obvious outliers. This result is fully consistent with prior claims of photometric redshift accuracy in the literature (Lanzetta et al. 1997, Mobasher et al. 2004). The accuracy of our photometric redshift estimates depends on the accuracy of the measured magnitudes in each of the available filters. It also is largely independent of the shape of an object (although magnitude errors for more extended, lower surface brightness objects tend to be somewhat larger than those for more concentrated, higher surface brightness objects of the same total magnitude).

The redshift distribution of all galaxies in the HUDF (solid line in Figure 3) is as expected, with the primary peak at $0.5 \lesssim z \lesssim 1.0$ and a generally declining tail at $z \simeq 4-5$. These trends were also seen in the general HDF redshift distribution of faint field galaxies (Driver et al. 1998). Also apparent is a lack of objects at $z \simeq 1-2$ due to unavailable UV spectral features crossing the $BViz(+JH)$ filters. This occurs because the HUDF does not have deep enough F300W or $U$-band (ultraviolet) data, unlike the situation in the HDF (Williams et al. 1996). This redshift bias, however, is the same for both the tadpole and the general field galaxy populations. In Figure 3 the tadpole galaxy distribution is multiplied by a factor of 16 for best comparison with that of the field galaxies. Within the available statistics, the redshift distribution shape of the tadpole galaxies follows that of the general field galaxies quite closely. This suggests that if tadpole galaxies are indeed dynamically young objects related to early-stage mergers, they may occur in the same proportion to the field galaxy population at all redshifts. Tadpole galaxies may therefore be good tracers of the galaxy assembly process. The ratio of the two redshift distributions $N(z)$ was calculated as well, and the resulting percentage of tadpole galaxies is plotted in Figure 4 as a function of redshift together with the statistical errors. Overall, the percentage of tadpole galaxies is roughly constant at $\sim 6\%$ with redshift to within the statistical errors for the redshift range probed in our study ($0.1 \lesssim z \lesssim 4.5$).

6. TADPOLE GALAXIES AS TRACERS OF GALAXY ASSEMBLY

The fact that about $6\%$ of all field galaxies are seen in the tadpole stage is a measurement with potentially important consequences. In light of simulations by Springel et al. (2005) that predict a tadpole-like stage $\simeq 0.7$ Gyr after a major merger begins, we suggest that this particular tadpole morphology represents an early-merger stage of two galaxies with comparable mass. If this $6\%$ indicates the fraction of time that an average galaxy in the HUDF spends in an early-merger stage during its lifetime, and if most of these low-luminosity objects started forming the bulk of their stars at the end of the reionization epoch at $z \simeq 6-7$ (e.g. Yan & Windhorst 2004a, 2004b), then each galaxy would spend about $6\%$ of $12.9$ Gyr (i.e. $0.8$ Gyr) since $z \simeq 7$ in a distinctly recognizable merger or tadpole stage. At the median redshift at which the tadpoles are seen ($z_{med} \simeq 1.6$; see Figure 3), each object is then seen at an age of about $4$ Gyr if born at $z \simeq 7$. Each tadpole is $\simeq 1''$ (or $\simeq 8$ kpc) across (Figure 1), and given the fluxes measured, each clump in a tadpole has roughly $M \simeq 10^8-10^9 M_{\odot}$ in stars (see, e.g., Papovich, Dickinson, & Ferguson 2001 who estimated stellar masses of lyman break galaxies). For these rough estimates of
their physical parameters, the freefall timescale for each tadpole is roughly $\tau \lesssim (\text{few} \times 10^7) - 10^8$ years, or $\approx 6\%$ of the galaxy lifetime at that redshift. Hence, if every galaxy is seen in a tadpole stage for $\approx 0.8$ Gyr of its lifetime, then it may have undergone $\lesssim 0.8$ Gyr of its lifetime. During the early stage of each merger, it would be temporarily seen as a tadpole. More complex mergers may lead to irregular/peculiar and trainwreck type objects and the luminous diffuse objects or clump clusters, which are among the type of objects that dominate the galaxy counts at faint magnitudes (Driver et al. 1998). In this paper, we limit the sample selection to two clumps passing by each other, which we believe leads to the more uniquely classifiable tadpole morphology. Given that tadpoles only trace a certain type and stage of merging galaxies, the above statistics are likely a lower limit on the number of all mergers. In summary, each galaxy seen today may have had of order one- to two-dozen mergers since most of its Population II stars were born at $z \approx 7$, and given the small masses and short merger timescales involved, would then be seen as tadpole galaxies for about $6\%$ of their lifetime.

Figure 4 suggests that tadpole galaxies — if indeed dynamically young objects — appear to occur in the same proportion to the field galaxy population at all redshifts probed in this study. Tadpole galaxies may therefore be good tracers of the process of galaxy as-

---

**Fig. 1.** — F775W ($i'$) band mosaic of a subset of the tadpole galaxy sample in the HUDF. Stamps retain the orientation of the HUDF; north is toward the top of the page and east is to the left. Index numbers are displayed in the upper-left corner of the stamps; photometric redshifts are given in upper-right corners. Parentheses indicate photometric redshifts with errors $>1$ based on HyperZ calculations. Stamps are 3 arcsec on a side. A table of coordinates for the entire tadpole sample is given in Section 3; the entire sample of 165 tadpoles appears in color in the online supplement. The vast majority of our tadpole sample contains the distinctive knot-plus-tail morphology, while sample contamination by normal (non-interacting) edge-on disk galaxies is less than 10%.

**Fig. 2.** — Distribution of (the absolute value of) angles ($\theta$) of all off-centered knots found within a radius $r \leq 4a$ ($\leq 2''$) from the center of each elongated diffuse structure in the HUDF, showing a clear excess of knots near $|\theta| \approx 0^\circ$.

**Fig. 3.** — Photometric redshift distribution of galaxies in the HUDF. The solid black histogram shows the redshift distribution of all HUDF field galaxies to $i_{AB} = 28.0$, while the dashed histogram shows the redshift distribution of the tadpole galaxies. The latter was multiplied by $16 \times$ for best comparison of its shape with the redshift distribution of the field galaxies.
assembly. This implies that the process of galaxy assembly — as traced by tadpole galaxies — keeps up with the reservoir of available galaxies as a function of redshift for $0.1 \leq z \leq 4.5$. Our result is in excellent agreement with the predictions of Robertson et al. (2005) that describe a merger-driven scenario to build up disk galaxies, and is consistent with Rhoads et al. (2005) who conclude that their $z=5.4$ galaxy (tadpole # 165 in Table 2) is strongly indicative of a galaxy in assembly.

7. RELATION BETWEEN TADPOLE GALAXIES AND ACTIVE GALACTIC NUCLEI

In a companion paper in this issue, Cohen et al. (2005) present a study of the variable objects in the HUDF which have a point source component with a measurably variable flux on timescales of 0.4–3.5 months, which roughly corresponds to 0.5–5.5 weeks in the rest-frame at the median redshift of the sample ($z_{\text{med}} \approx 1.5 - 2$). In particular, they found 45 plausibly variable objects among 4644 galaxies to $V_{\text{AB}} = 28.0$ mag. They argue that these objects are most likely variable because they host weak AGN. Sometimes these AGN are in the galaxy center, but often they occur off-center in a dynamically relaxed system. This prompts the question: What fraction of tadpole galaxies contains a variable weak AGN in its knots? This is a critical issue, because it is widely believed that the process of merging in galaxies can also disturb the inner accretion disk around the supermassive black hole (SMBH) and switch on the AGN. Among our 165 tadpole galaxies, none coincide with the sample of 45 variable objects or with the x-ray sources in the Chandra Deep Field South (Alexander, D.M. et al. 2005). Recent state-of-the-art hydrodynamical models (di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604) suggest that during (major) mergers, the black hole accretion rate peaks considerably after the merger started, and after the star formation rate (SFR) has peaked. Specifically, their models suggest that, for massive galaxies, a tadpole stage is seen typically about 0.7 Gyr after the merger started, but $\sim 0.9$ Gyr before the SMBH accretes most of its mass, which is when the galaxy displays strong AGN activity. Since the lifetimes of QSO’s and radio-galaxies are known to be $\lesssim (\text{few} \times 10^7) - 10^8$ years (Martini, P. 2004, Grazian et al. 2004), these hydrodynamical models thus imply that the AGN stage is expected to occur considerably after (i.e., $\geq 1$ Gyr) the early-merger event during which the galaxy is seen in the tadpole stage. The observed lack of overlap between the tadpole galaxies and the AGN sample in the HUDF provide direct observational support for this prediction. Recent studies by Grogin et al. (2005) find asymmetry (A) values that are similar between AGN and non-AGN samples; this result is consistent with our study in light of the theoretical models mentioned above, which indicate that AGN activity is seen only well after the merger has taken place and the galaxy has settled into a more dynamically relaxed state. In addition, Hopkins et al. (2005) have quantified the timescales that quasars will be visible during merging events, noting that for a large fraction of the accretion time, the quasar is heavily obscured. In particular, their simulations show that during an early merging phase — our observed tadpole phase — the intrinsic quasar luminosity peaks, but is completely obscured. Only after feedback from the central quasar clears out the gas will the object become visible as an AGN. This should be observable by the Spitzer Space Telescope in the mid-infrared (IR) as a correspondingly larger fraction of IR-selected obscured faint QSO’s.

In conclusion, tadpole galaxies are a class of easily-identifiable, dynamically young objects that exist throughout the history of the Universe and are good tracers of galaxy assembly. They provide strong observational support for the validity of recent numerical simulations, and highlight the importance of mergers to the process of galaxy assembly and AGN growth.

8. ACKNOWLEDGEMENTS

We thank the STScI staff, and in particular Steve Beckwith, Ray Lucas, and Massimo Stiavelli, for their persistent efforts to implement the Hubble Ultra Deep Field in the best possible way. This research was partially funded by NASA grants GO-9793.08-A and AR-10298.01-A, awarded by STScI, which is operated by AURA for NASA under contract NAS 5-26555; as well as by the NASA Space Grant program at ASU, and the Harriet G. Jenkins Predoctoral Fellowship Program.
van den Bergh, S., 2002, PASP 114, 797
### TABLE 2
Global Properties of Tadpole Galaxies

<table>
<thead>
<tr>
<th>ID</th>
<th>RA (h : m : s)</th>
<th>Dec (° : ′ : ″)</th>
<th>$i'_{AB}$ (mag)</th>
<th>$z^\dagger$</th>
<th>ID</th>
<th>RA (h : m : s)</th>
<th>Dec (° : ′ : ″)</th>
<th>$i'_{AB}$ (mag)</th>
<th>$z^\dagger$</th>
<th>ID</th>
<th>RA (h : m : s)</th>
<th>Dec (° : ′ : ″)</th>
<th>$i'_{AB}$ (mag)</th>
<th>$z^\dagger$</th>
</tr>
</thead>
</table>

**NOTE:** parentheses indicate estimated aperture magnitudes for visually selected objects

1 From redshift from Rhoads et al. (2005)
2 Photometric redshifts computed from HyperZ (Bolzonella et al. 2000)