

## A NEW NONPARAMETRIC APPROACH TO GALAXY MORPHOLOGICAL CLASSIFICATION

JENNIFER M. LOTZ,<sup>1</sup> JOEL PRIMACK,<sup>1</sup> AND PIERO MADAU<sup>2</sup>

Received 2003 November 14; accepted 2004 April 12

### ABSTRACT

We present two new nonparametric methods for quantifying galaxy morphology: the relative distribution of the galaxy pixel flux values (the Gini coefficient or  $G$ ) and the second-order moment of the brightest 20% of the galaxy's flux ( $M_{20}$ ). We test the robustness of  $G$  and  $M_{20}$  to decreasing signal-to-noise ratio (S/N) and spatial resolution and find that both measures are reliable to within 10% for images with average S/N per pixel greater than 2 and resolutions better than 1000 and 500 pc, respectively. We have measured  $G$  and  $M_{20}$ , as well as concentration ( $C$ ), asymmetry ( $A$ ), and clumpiness ( $S$ ) in the rest-frame near-ultraviolet/optical wavelengths for 148 bright local “normal” Hubble-type galaxies (E–Sd) galaxies, 22 dwarf irregulars, and 73  $0.05 < z < 0.25$  ultraluminous infrared galaxies (ULIRGs). We find that most local galaxies follow a tight sequence in  $G$ – $M_{20}$ – $C$ , where early types have high  $G$  and  $C$  and low  $M_{20}$  and late-type spirals have lower  $G$  and  $C$  and higher  $M_{20}$ . The majority of ULIRGs lie above the normal galaxy  $G$ – $M_{20}$  sequence because of their high  $G$  and  $M_{20}$  values. Their high Gini coefficients arise from very bright nuclei, while the high second-order moments are produced by multiple nuclei and bright tidal tails. All of these features are signatures of recent and on-going mergers and interactions. We also find that in combination with  $A$  and  $S$ ,  $G$  is more effective than  $C$  at distinguishing ULIRGs from the “normal” Hubble types. Finally, we measure the morphologies of  $491.7 < z < 3.8$  galaxies from *HST* NICMOS observations of the Hubble Deep Field North. We find that many of the  $z \sim 2$  galaxies possess  $G$  and  $A$  higher than expected from degraded images of local elliptical and spiral galaxies and have morphologies more like low-redshift ULIRGs.

*Key words:* galaxies: fundamental parameters — galaxies: high-redshift — galaxies: peculiar — galaxies: structure

*On-line material:* machine-readable tables, color figures

### 1. INTRODUCTION

The evolution of the physical structure of galaxies is one of the keys to understanding how matter in the universe assembled into the structures we see today. The most accessible tracer of a galaxy's physical structure is its morphology, i.e., the organization of its light (stars and dust), as projected into our line of sight and observed at a particular wavelength. As we examine more distant galaxies, we find that galaxy morphologies become increasingly chaotic. The disk and spheroidal structures abundant in the local universe disappear at early times in the universe (e.g., Abraham et al. 1996; Abraham & van den Bergh 2001). The emergence of the local Hubble sequence of spiral and elliptical galaxies at late times is one of the predictions of the hierarchical picture of galaxy assembly.

While the first morphological studies sought to describe the variety of galaxy shapes and forms, the goal of present-day morphological studies is to tie the spatial distribution of stars to the formation history of the galaxy. A major obstacle to this goal has been the difficulty in quantifying morphology with a few simple, reliable measurements. One tack is to describe a galaxy parametrically, by modeling the distribution of light as projected into the plane of the sky with a prescribed analytic function. For example, bulge-to-disk ( $B/D$ ) light ratios may be computed by fitting the galaxy with a two-component profile,

where the fluxes, sizes, concentrations, and orientations of the bulge and disk components are free parameters (Peng et al. 2002; Simard et al. 2002). This  $B/D$  ratio correlates with qualitative Hubble type classifications, although with significant scatter. Unfortunately, there is often a fair amount of degeneracy in the best-fitting models and  $B/D$  ratios, and structures such as compact nuclei, bars, and spiral arms introduce additional difficulty in fitting the bulge and disk components (e.g., Balcells et al. 2003). A related approach is to fit a single Sersic profile to the entire galaxy (Blanton et al. 2003a). Profiles with high Sersic indices are interpreted as bulge-dominated systems, while low Sersic indices indicate disk-dominated systems. However, not all bulges have high Sersic index values—some are exponential in nature (Carollo 1999), so not all objects with bulges will produce intermediate or high Sersic indices. Both the one-component and multiple-component fitting methods assume that the galaxy is well described by a smooth, symmetric profile—an assumption that breaks down for irregular, tidally disturbed, and merging galaxies.

Nonparametric measures of galaxy morphology do not assume a particular analytic function for the galaxy's light distribution and therefore may be applied to irregulars, as well as standard Hubble-type galaxies. Abraham et al. (1994, 1996) introduced the concentration index  $C$  (which roughly correlates with a galaxy's  $B/D$  ratio) and Schade et al. (1995) put forward rotational asymmetry  $A$  as a way to automatically distinguish early Hubble types (E/S0/Sa) from later Hubble types (Sb/Sc) and classify irregular and merging galaxies. Subsequent authors modified the original definitions to make  $C$  and  $A$  more robust to surface-brightness selection and centering errors (Wu 1999;

<sup>1</sup> Santa Cruz Institute of Particle Physics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064; jlotz@scipp.ucsc.edu, joel@scipp.ucsc.edu.

<sup>2</sup> Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064; pmadau@ucolick.org.

Bershady et al. 2000; Conselice et al. 2000). The third quantity in the “CAS” morphological classification system is a measure of a galaxy’s residual clumpiness  $S$ , which is correlated with a galaxy’s color and star formation rate (Isserstedt & Schindler 1986; Takamiya 1999; Conselice 2003). Other more computer-intensive approaches to galaxy classification, such as artificial neural networks and shapelet decomposition have also been applied to local and distant galaxies. Artificial neural networks are trained by an astronomer on a set of galaxies of known morphological type and use a combination of size, surface brightness, concentration, and color to classify galaxy types (Odewahn et al. 1996; Naim et al. 1997). “Shaplets” deconstruct each galaxy’s image into a series of Hermite polynomials (Refregier 2003; Kelly & McKay 2004). The eigenshapes produced by shapelet decomposition are often difficult to interpret by themselves, and the additional step of principle component analysis is performed to classify galaxies.

While CAS is perhaps the most straightforward of the non-parametric methods, it is not without its weaknesses. Because concentration is measured within several circular apertures about a predefined center, it implicitly assumes circular symmetry, making it a poor descriptor for irregular galaxies. Asymmetry is more sensitive to merger signatures than concentration, but not all merger remnant candidates are highly asymmetric, and not all asymmetric galaxies are mergers (e.g., dusty edge-on spirals). Finally, the clumpiness determination requires one to define a galaxy smoothing length, which must be chosen carefully to avoid systematic effects dependent on a galaxy image’s point-spread function (PSF), pixel scale, distance, and angular size. Also, the bulges of highly concentrated galaxies give strong residuals that are not due to star-forming regions and must be masked out when computing  $S$ .

In this paper we examine two new nonparametric ways of quantifying galaxy morphology that circumvent some of the problems with the CAS system. We use the Gini coefficient, a statistic used in economics to describe the distribution of wealth within a society. It was first adapted for galaxy morphology classification by Abraham et al. (2003) to quantify the relative distribution of flux within the pixels associated with a galaxy. It is correlated with concentration, but does not assume that the brightest pixels are in the geometric center of the galaxy image. We also define a new indicator,  $M_{20}$ , which describes the second-order moment of the brightest 20% of the galaxy. While similar to the concentration index,  $M_{20}$  is more sensitive to merger signatures like multiple nuclei and does not impose circular symmetry. In § 2 we modify Abraham’s definition of the Gini coefficient to make it applicable to distant galaxies, and we define  $M_{20}$ . In § 3 we test the robustness of these statistics to decreasing signal-to-noise ratio (S/N) and resolution and find that at average S/N per galaxy pixel greater than 2 and spatial resolutions less than 500 pc, they are reliable to within 10%. We also compare the robustness of  $G$  and  $M_{20}$  to CAS. In § 4 we compare the ability of  $G$  and  $M_{20}$  to classify local Hubble type and merging galaxies to the CAS system. Finally, in § 5 we examine the near-ultraviolet/optical morphologies of 49  $1.7 < z < 3.8$  Lyman-break galaxies and attempt to classify these LBGs as ellipticals, spirals, or merger candidates.

## 2. MEASURING GALAXY MORPHOLOGIES

### 2.1. The Gini Coefficient

The Gini coefficient is a statistic based on the Lorenz curve, the rank-ordered cumulative distribution function of a

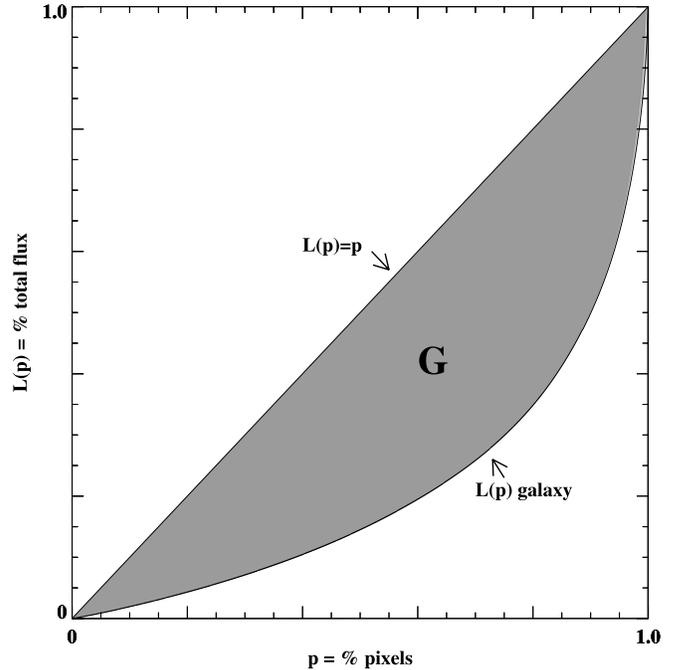


FIG. 1.—Lorenz curve: the Gini coefficient is the area between the Lorenz curve of the galaxy’s pixels and that of equitable distribution (*shaded region*). The given curve is for S0 NGC 4526,  $G = 0.59$ .

population’s wealth, or in this case a galaxy’s pixel values (Abraham et al. 2003). The Lorenz curve is defined as

$$L(p) = \frac{1}{\bar{X}} \int_0^p F^{-1}(u) du, \quad (1)$$

where  $p$  is the percentage of the poorest citizens or faintest pixels,  $F(x)$  is the cumulative distribution function, and  $\bar{X}$  is the mean over all (pixel flux) values  $X_i$  (Lorenz 1905). The Gini coefficient is the ratio of the area between the Lorenz curve and the curve of “uniform equality” where  $L(p) = p$  (*shaded region*, Fig. 1) to the area under the curve of uniform equality ( $=\frac{1}{2}$ ). For a discrete population, the Gini coefficient is defined as the mean of the absolute difference between all  $X_i$ :

$$G = \frac{1}{2\bar{X}n(n-1)} \sum_{i=1}^n \sum_{j=1}^n |X_i - X_j|, \quad (2)$$

where  $n$  is the number of people in a population or pixels in a galaxy. In a completely egalitarian society,  $G$  is zero, and if one individual has all the wealth,  $G$  is unity. A more efficient way to compute  $G$  is to first sort  $X_i$  into increasing order and calculate

$$G = \frac{1}{\bar{X}n(n-1)} \sum_i^n (2i - n - 1)X_i \quad (3)$$

(Glasser 1962).

For the majority of local galaxies, the Gini coefficient is correlated with the concentration index and increases with the fraction of light in a compact (central) component. In a study of 930 Sloan Digital Sky Survey (SDSS) Early Data Release galaxies, Abraham et al. (2003) found  $G$  to be strongly correlated with both concentration and surface brightness. However, unlike  $C$ ,  $G$  is independent of the large-scale spatial

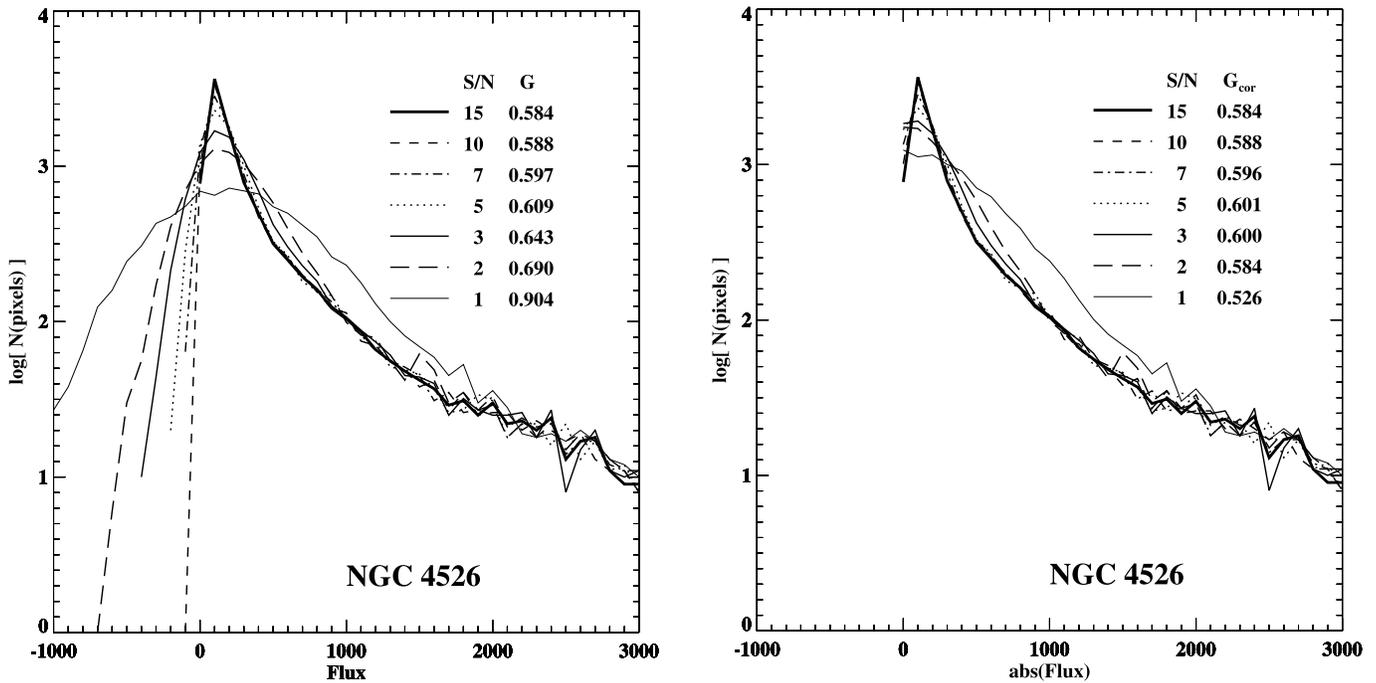


FIG. 2.—Pixel flux value distribution as a function of the average S/N per galaxy pixel for S0 galaxy NGC 4526. *Left:* As  $\langle S/N \rangle$  decreases, more faint galaxy pixels are scattered below the background sky level. *Right:* Corrected Gini coefficients calculated from the distribution of absolute pixel flux values.

distribution of the galaxy’s light. The correlation between  $C$  and  $G$  exists because highly concentrated galaxies have much of their light in a small number of pixels. High  $G$  values may also arise when very bright galaxy pixels are not found in the center of a bulge. Therefore,  $G$  differs from  $C$  in that it can distinguish between galaxies with shallow light profiles (which have both low  $C$  and  $G$ ) and galaxies where much of the flux is located in a few pixels not at the projected center (which have low  $C$  but high  $G$ ).

In practice, the application of the Gini coefficient to galaxy observations requires some care. One must have a consistent definition of the pixels belonging to the galaxy to measure the distribution of flux within those pixels and compare that distribution with other galaxies. The inclusion of “sky” pixels will systematically increase  $G$ , while the exclusion of low surface brightness “galaxy” pixels will systematically decrease  $G$ . Abraham et al. (2003) measure  $G$  for galaxy pixels that lie above a constant surface brightness threshold. This definition makes the direct comparison between high-redshift galaxies and the local galaxy population difficult because of the  $(1+z)^4$  surface brightness dimming of distant galaxies. Therefore, we attempt to create a segmentation map of the galaxy pixels in a way that is insensitive to surface brightness dimming. The mean surface brightness  $\mu(r_p)$  at the Petrosian radius  $r_p$  is used to set the flux threshold above which pixels are assigned to the galaxy. The Petrosian radius is the radius  $r_p$  at which the ratio of the surface brightness at  $r_p$  to the mean surface brightness within  $r_p$  is equal to a fixed value, i.e.,

$$\eta = \frac{\mu(r_p)}{\bar{\mu}(r < r_p)}, \quad (4)$$

where  $\eta$  is typically set to 0.2 (Petrosian 1976). Because the Petrosian radius is based on a curve of growth, it is largely insensitive to variations in the limiting surface brightness and S/N of the observations. This revised definition should allow

better comparison of  $G$  values for galaxies with varying surface brightnesses, distances, and observed S/N.

The galaxy image is sky-subtracted and any background galaxies, foreground stars, or cosmic rays are removed from the image. The mean ellipticity and position angle of the galaxy is measured using IRAF task ELLIPSE. The Petrosian “radius” (or semimajor-axis length) is measured for increasing elliptical apertures, rather than circular apertures. While the Petrosian radius determined by the curve of growth within circular apertures is similar to that determined from elliptical apertures for most galaxies, elliptical apertures more closely follow the galaxy’s true light profile and can produce very different  $r_p$  values for edge-on spirals. To create the segmentation map, the cleaned galaxy image is first convolved with a Gaussian with  $\sigma = r_p/5$ . This step raises the signal of the galaxy pixels above the background noise, making low surface brightness galaxy pixels more detectable. Then the surface brightness  $\mu$  at  $r_p$  is measured and pixels in the smoothed image with flux values  $\geq \mu(r_p)$  and less than  $10\sigma$  from their neighboring pixels are assigned to the galaxy. The last step assures that any remaining cosmic rays or spurious noise pixels in the image are not included in the segmentation map. This map is then applied to the cleaned but unsmoothed image, and the pixels assigned to the galaxy are used to compute the Gini coefficient.

Even when the pixels assigned to a galaxy are robustly determined, the distribution of flux within the pixels will depend on the S/N, as noise smears out the flux distribution in the faintest pixels. This is illustrated in the left of Figure 2 by adding increasing Poisson sky noise to the S0 galaxy NGC 4526 image and recalculating the segmentation map and Gini coefficient. We define the average signal-to-noise per galaxy pixel  $\langle S/N \rangle$  as

$$\langle S/N \rangle = \frac{1}{n} \sum_n^i \frac{S_i}{\sqrt{\sigma_{\text{sky}}^2 + S_i}}, \quad (5)$$

where  $S_i$  is pixel  $i$ 's flux,  $\sigma_{\text{sky}}$  is the sky noise, and  $n$  is the number of galaxy pixels in the segmentation map. As  $\langle S/N \rangle$  decreases, the distribution of measured flux values in the faintest pixels becomes broader. The measured Gini coefficient increases because low surface brightness galaxy pixels are scattered to flux values below the mean sky level, resulting in negative flux levels for the faintest pixels assigned to the galaxy by our smoothed segmentation map. We note that, while the Poisson noise redistributes all the pixel flux values, the effects are significant only for pixels with intrinsic flux values  $\leq 3 \sigma_{\text{sky}}$ . Therefore, as a first-order correction we compute the Gini coefficient of the distribution of *absolute* flux values:

$$G = \frac{1}{|\bar{X}|n(n-1)} \sum_i^n (2i - n - 1)|X_i|. \quad (6)$$

Low surface brightness galaxy pixels with flux values scattered below the sky level are reassigned positive values (*right*, Fig. 2). This correction recovers the “true” Gini coefficient to within 10% for images with  $S/N > 2$ ; at very low  $S/N$  values, even the brightest galaxy pixels are strongly affected by noise and the Gini coefficient is not recoverable. In Figures 3–4 we show the final segmentation maps used to compute the Gini coefficient as contour maps for eight galaxies of varying morphological type (Table 1).

### 2.2. The Moment of Light

The total second-order moment  $M_{\text{tot}}$  is the flux in each pixel  $f_i$  multiplied by the squared distance to the center of the galaxy, summed over all the galaxy pixels assigned by the segmentation map:

$$M_{\text{tot}} = \sum_i^n M_i = \sum_i^n f_i [(x_i - x_c)^2 + (y_i - y_c)^2], \quad (7)$$

where  $x_c, y_c$  is the galaxy's center. The center is computed by finding  $x_c, y_c$  such that  $M_{\text{tot}}$  is minimized.

The second-order moment of the brightest regions of the galaxy traces the spatial distribution of any bright nuclei, bars, spiral arms, and off-center star clusters. We define  $M_{20}$  as the normalized second-order moment of the brightest 20% of the galaxy's flux. To compute  $M_{20}$ , we rank-order the galaxy pixels by flux, sum  $M_i$  over the brightest pixels until the sum of the brightest pixels equals 20% of the total galaxy flux, and then normalize by  $M_{\text{tot}}$ :

$$M_{20} \equiv \log_{10} \left( \frac{\sum_i M_i}{M_{\text{tot}}} \right), \quad \text{while } \sum_i f_i < 0.2 f_{\text{tot}}. \quad (8)$$

Here  $f_{\text{tot}}$  is the total flux of the galaxy pixels identified by the segmentation map and  $f_i$  are the fluxes for each pixel  $i$ , order such that  $f_1$  is the brightest pixel,  $f_2$  is the second brightest pixels, and so on. The normalization by  $M_{\text{tot}}$  removes the dependence on total galaxy flux or size. We find that defining  $M$  with brighter flux thresholds (e.g., 5% of  $f_{\text{tot}}$ ) produces moment values that are unreliable at low spatial resolutions (§ 2.3), while lower flux threshold lead to a less discriminating statistic.

While our definition of  $M_{20}$  is similar to that of  $C$ , it differs in two important respects. First,  $M_{20}$  depends on  $r^2$  and is more heavily weighted by the spatial distribution of luminous regions. Second, unlike  $C$ ,  $M_{20}$  is not measured within circular or elliptical apertures, and the center of the galaxy is a free parameter. We will see in § 3 that these differences make  $M_{20}$

more sensitive than  $C$  to merger signatures such as multiple nuclei. In Figures 3–4 we display the segmentation maps and the regions containing the brightest 20% of the flux for the eight test galaxies.

### 2.3. Concentration, Asymmetry, and Smoothness

Concentration is defined in slightly different ways by different authors, but the basic function measures the ratio of light within a circular or elliptical inner aperture to the light within an outer aperture. We adopt the Bershady et al. (2000) definition as the ratio of the circular radii containing 20% and 80% of the “total flux”:

$$C = 5 \log \left( \frac{r_{80}}{r_{20}} \right), \quad (9)$$

where  $r_{80}$  and  $r_{20}$  are the circular apertures containing 80% and 20% of the total flux, respectively. For comparison with the most recent studies of galaxy concentration, we use Conselice's (2003) definition of the total flux as the flux contained within  $1.5 r_p$  of the galaxy's center (as opposed to Bershady's definition as the flux contained within  $2 r_p$ ). For the concentration measurement, the galaxy's center is that determined by the asymmetry minimization (see below). In Figures 3–4 we overplot  $r_{80}$  and  $r_{20}$  for eight galaxies of varying morphological type in the far left-hand panels.

The asymmetry parameter  $A$  quantifies the degree to which the light of a galaxy is rotationally symmetric.  $A$  is measured by subtracting the galaxy image rotated by  $180^\circ$  from the original image (Abraham et al. 1996; Wu 1999; Conselice et al. 2000):

$$A = \frac{\sum_{i,j} |I(i,j) - I_{180}(i,j)|}{\sum_{i,j} |I(i,j)|} - B_{180}, \quad (10)$$

where  $I$  is the galaxy's image and  $I_{180}$  is the image rotated by  $180^\circ$  about the galaxy's central pixel, and  $B_{180}$  is the average asymmetry of the background.  $A$  is summed over all pixels within  $1.5 r_p$  of the galaxy's center. The central pixel is determined by minimizing  $A$ . The asymmetry due to the noise must be corrected for, and it is impossible to reliably measure the asymmetry for low  $S/N$  images. In Figures 3–4 we display the residual  $I - I_{180}$  image and the  $1.5 r_p$  aperture in the second column. Objects with very smooth elliptical light profiles have a high degree of rotational symmetry. Galaxies with spiral arms are less symmetric, while extremely irregular and merging galaxies are often (but not always) highly asymmetric.

The smoothness parameter  $S$  has been recently developed by Conselice (2003), inspired by the work of Takamiya (1999), to quantify the degree of small-scale structure. The galaxy image is smoothed by a boxcar of given width and then subtracted from the original image. The residual is a measure of the clumpiness due to features such as compact star clusters. In practice, the smoothing scale length is chosen to be a fraction of the Petrosian radius.

$$S = \frac{\sum_{i,j} |I(i,j) - I_S(i,j)|}{\sum_{i,j} |I(i,j)|} - B_S, \quad (11)$$

where  $I_S$  is the galaxy's image smoothed by a boxcar of width  $0.25 r_p$ , and  $B_S$  is the average smoothness of the background. Like  $A$ ,  $S$  is summed over the pixels within  $1.5 r_p$  of the galaxy's center. However, because the central regions of most

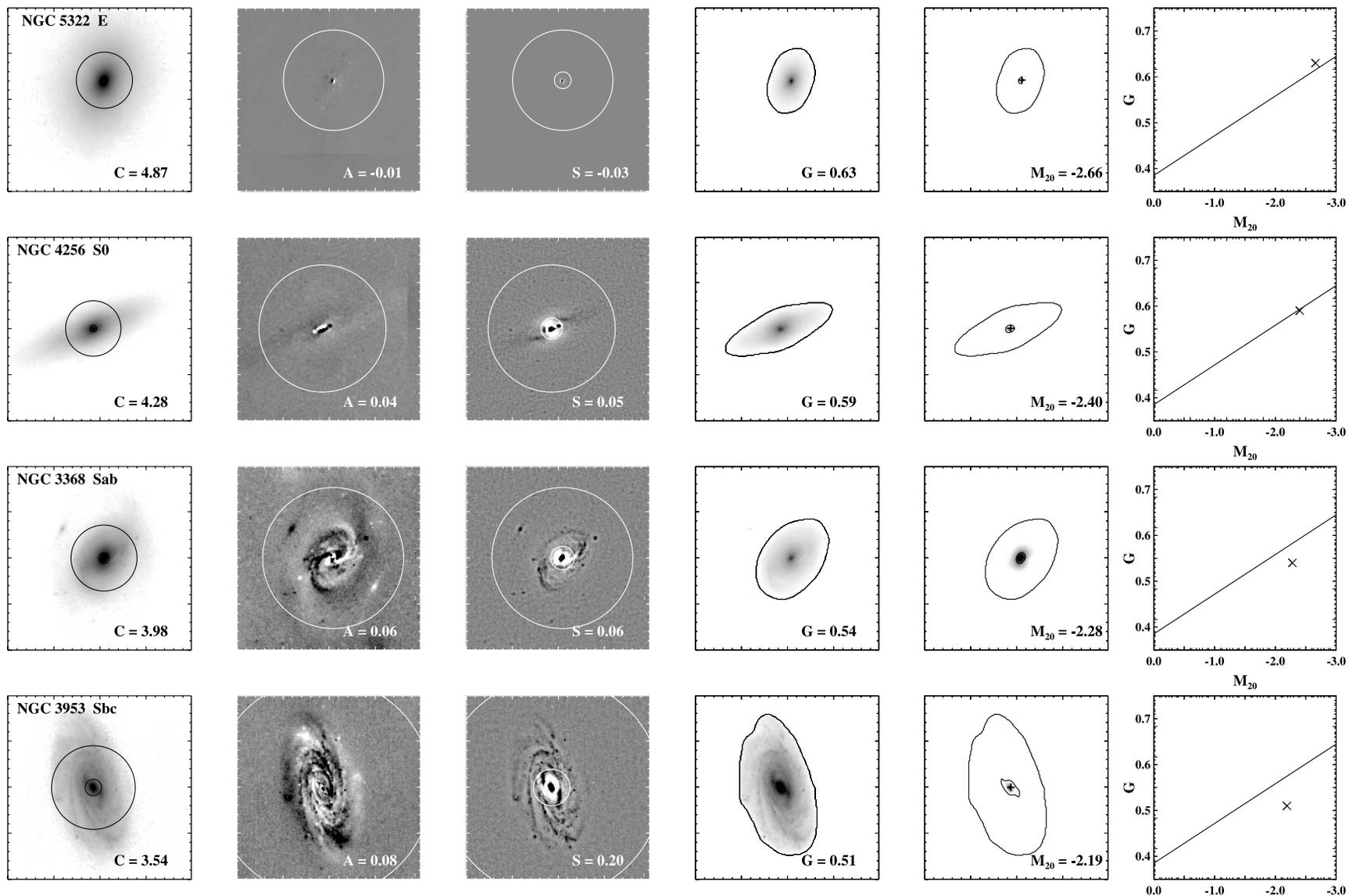


FIG. 3.—Test galaxy morphological measurements  $C$ ,  $A$ ,  $S$ ,  $G$ , and  $M_{20}$  for rest-frame  $\sim 6500 \text{ \AA}$  images (Table 1). In the first panel inner and outer circles enclose 20% and 80% of the flux within  $1.5 r_p$ . The second panel shows the residual  $I - I_{180}$  image, with the circle at  $1.5 r_p$ . The third panel shows the residual  $I - I_S$  image, with the inner and outer circles at  $0.25$  and  $1.5 r_p$ . The fourth panel images are the original galaxy images scaled such that the minimum surface brightness matches that used to create the galaxy segmentation maps. The outer edge of the segmentation map are the outer contour plotted in the fourth and fifth panels. The inner contours plotted in the fifth panel trace each galaxy's brightest 20% of it flux, while the crosses indicate each galaxy's center. The final panel plots each galaxy's  $G$  and  $M_{20}$ , where the solid line is for reference.

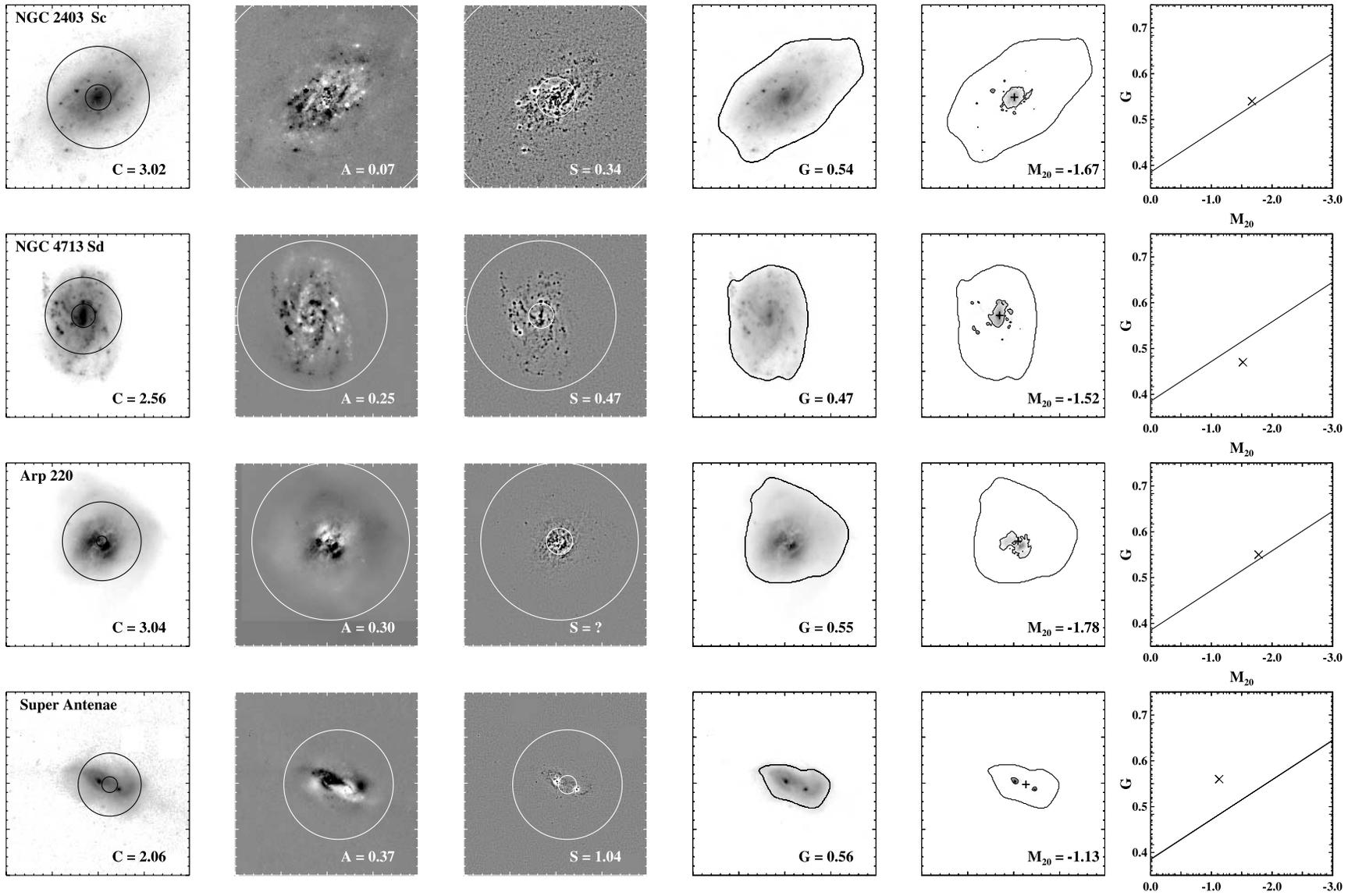


FIG. 4.—Same as Fig. 3.

TABLE 1  
 TEST GALAXIES

Galaxy	Type <sup>a</sup>	Dist. (Mpc)	S/N	Res. (pc pixel <sup>-1</sup> )	$C_R$	$A_R$	$S_R$	$G_R$	$M_{20,R}$	Notes
NGC 5332 <sup>b</sup>	E4(E/S0)	28.7	7.1	56	4.87	-0.01	-0.03	0.63	-2.66	...
NGC 4526 <sup>c</sup>	S0_3_(6)	17.0	15.1	111	4.28	0.04	0.05	0.59	-2.40	Virgo Cluster
NGC 3368 <sup>c</sup>	Sab(s)II	11.2	14.6	73	3.98	0.06	0.06	0.54	-2.28	Leo group
NGC 3953 <sup>c</sup>	SBbc(r)I-II	18.6	8.4	122	3.54	0.08	0.20	0.51	-2.19	Ursa Major group
NGC 2403 <sup>c</sup>	Sc(s)III	3.2	5.8	19	3.02	0.07	0.34	0.54	-1.67	M81 group
NGC 4713 <sup>b</sup>	SAB(rs)d	17.0	6.4	33	2.56	0.25	0.47	0.47	-1.52	Virgo Cluster
Arp 220 <sup>d</sup>	ULIRG	77.0	3.7	37	2.92	0.30	0.43	0.55	-1.64	IRAS 15327+2340
Super Antenna <sup>d</sup>	ULIRG	245.4	3.1	119	2.06	0.37	1.04	0.56	-1.13	IRAS 19254-7245

<sup>a</sup> See Sandage & Bedke (1994).

<sup>b</sup> From Abazajian et al. (2003).

<sup>c</sup> From Frei et al. (1996).

<sup>d</sup> From Borne et al. (2000).

galaxies are highly concentrated, the pixels within a circular aperture equal to the smoothing length  $0.25 r_p$  are excluded from the sum. In Figures 3–4 we display the residual  $I - I_S$  images, and the  $0.25$  and  $1.5 r_p$  apertures in the third column.  $S$  is correlated with recent star formation (Takamiya 1999; Conselice 2003). However, because of its strong dependence on resolution, it is not applicable to poorly resolved and distant galaxies.

### 3. RESOLUTION AND NOISE EFFECTS

To make a fair comparison of the measured morphologies of different galaxies, we must understand how noise and resolution affect  $G$  and  $M_{20}$ . This is particularly important when comparing local galaxies to high-redshift galaxies, as the observations of distant galaxies are generally of lower S/N and resolution than those of local galaxies. We have defined  $G$  and  $M_{20}$  in the previous sections in an attempt to minimize systematic offsets with noise and resolution. Nevertheless, any measurement is ultimately limited by the S/N of the observations. Also, the PSF and finite pixel size of the images may introduce increasing uncertainties to the morphologies as the resolution decreases and small-scale structures are washed out.

We have chosen eight galaxies of varying morphological type (Figs. 3–4; Table 1) to independently test the effects of decreasing S/N per pixel and physical resolution (parsecs per pixel) on the measurements of  $G$ ,  $M_{20}$ ,  $C$ ,  $A$ , and  $S$ . For the S/N tests, random Poisson noise maps of increasing variance were added to the original sky-subtracted image. For each noise-added image, we measured  $r_p$ , created a new segmentation map, measured  $\langle S/N \rangle$  for galaxy pixels assigned by the segmentation map, and measured  $G$ ,  $M_{20}$ ,  $C$ ,  $A$ , and  $S$ . Noisy galaxy images were created and measured 20 times at each S/N level, and the mean changes in the morphological values with  $\langle S/N \rangle$  are plotted in Figure 5. To simulate the effect of decreasing resolution, we rebinned the galaxy images to increasingly large pixel sizes. Rebinning the original galaxy images increases the S/N per pixel, so additional Poisson sky noise ( $\sigma_{\text{sky}}$ ) was added to the rebinned image such that average  $\langle S/N \rangle$  was kept constant with decreasing resolution. Again, we measured  $r_p$ , created a segmentation map, and computed the average change  $G$ ,  $M_{20}$ ,  $C$ ,  $A$ , and  $S$  with resolution for 20 simulations at each resolution step (Fig. 6).

We find that  $G$ ,  $M_{20}$ , and  $C$  are reliable to within  $\sim 10\%$  ( $\Delta \leq 0.05, 0.2$ , and  $0.3$ , respectively) for galaxy images with

$\langle S/N \rangle \geq 2$ .  $A$  systematically decreases with  $\langle S/N \rangle$ , but generally shows offsets less than  $0.1$  at  $\langle S/N \rangle \geq 5$ .  $S$  also systematically decreases with  $\langle S/N \rangle$  and has decrements less than  $0.2$  at  $\langle S/N \rangle \geq 5$ . Decreasing resolution, however, has much stronger effects on the morphology measurements.  $C$  and  $M_{20}$  show systematic offsets greater than  $\sim 15\%$  ( $\Delta \geq 0.5$  and  $0.3$ , respectively) at resolution scales worse than  $500$  pc, as the cores of the observed galaxies become unresolved.  $G$ ,  $A$ , and  $S$ , on the other hand, are relatively stable to decreasing spatial resolution down to  $1000$  pc. As a galaxy’s image becomes less resolved, the observed curve of growth changes, resulting in larger  $r_p$  values and therefore producing slightly higher  $G$  values as the segmentation map grows accordingly. At the lowest resolutions the observed biases in  $C$ ,  $A$ , and  $S$  appear to be a function of Hubble type: the E–Sbc galaxies are biased to higher  $A$  and  $S$  and lower  $C$ , while both the Sd and mergers are biased toward lower  $A$  and the merger remnants are biased to higher  $C$ . On the other hand, on the Sc and Sd galaxies show  $G$  offsets greater than  $20\%$  ( $\Delta \sim 0.1$ ) at resolutions between  $1000$  and  $2000$  pc.

## 4. LOCAL GALAXY MORPHOLOGIES

### 4.1. Frei and SDSS Local Galaxy Samples

We have measured  $G$ ,  $M_{20}$ ,  $C$ ,  $A$ , and  $S$  at both  $\sim 4500$  and  $\sim 6500$  Å for 104 local galaxies taken from the Frei et al. (1996) catalog. The Frei catalog galaxies are a representative sample of bright, well-resolved, Hubble-type galaxies (E–S0–Sa–Sb–Sc–Sd) and have been used as morphological standards by a number of authors (Takamiya 1999; Wu 1999; Bershady et al. 2000; Conselice et al. 2000; Simard et al. 2002). The galaxies were observed by Frei et al. (1996) with either the 1.5 m telescope at Palomar Observatory or the 1.1 m telescope at Lowell Observatory. The Palomar images were taken in the Thuan–Gunn  $g$  and  $r$  filters ( $\lambda_{\text{eff}} = 5000, 6500$  Å) at plate scale =  $1''.19$  pixel<sup>-1</sup> and typical PSF FWHM  $\sim 2''$ – $3''$ . The Lowell images were taken in the  $B_J$  and  $R$  passbands ( $\lambda_{\text{eff}} = 4500, 6500$  Å) at a plate scale =  $1.35''$  pixel<sup>-1</sup> and typical PSF FWHM  $\sim 3''$ – $5''$ . In Table 2 we give  $G$ ,  $M_{20}$ ,  $C$ ,  $A$ , and  $S$  as measured in  $R/r$  and  $B_J/g$  for each of the galaxies.

We have also obtained the images of nine Frei galaxies and 44 other galaxies selected by their  $u$ -band brightness ( $u < 14$ ) from the SDSS Data Release 1 database (Abazajian et al. 2003). The morphologies of the SDSS sample were measured in the  $u$ ,  $g$ , and  $r$  bands ( $\lambda_{\text{eff}} = 3600, 4400$ , and  $6500$ , respectively; Table 3). The SDSS plate scale is  $0''.4$  pixel<sup>-1</sup>

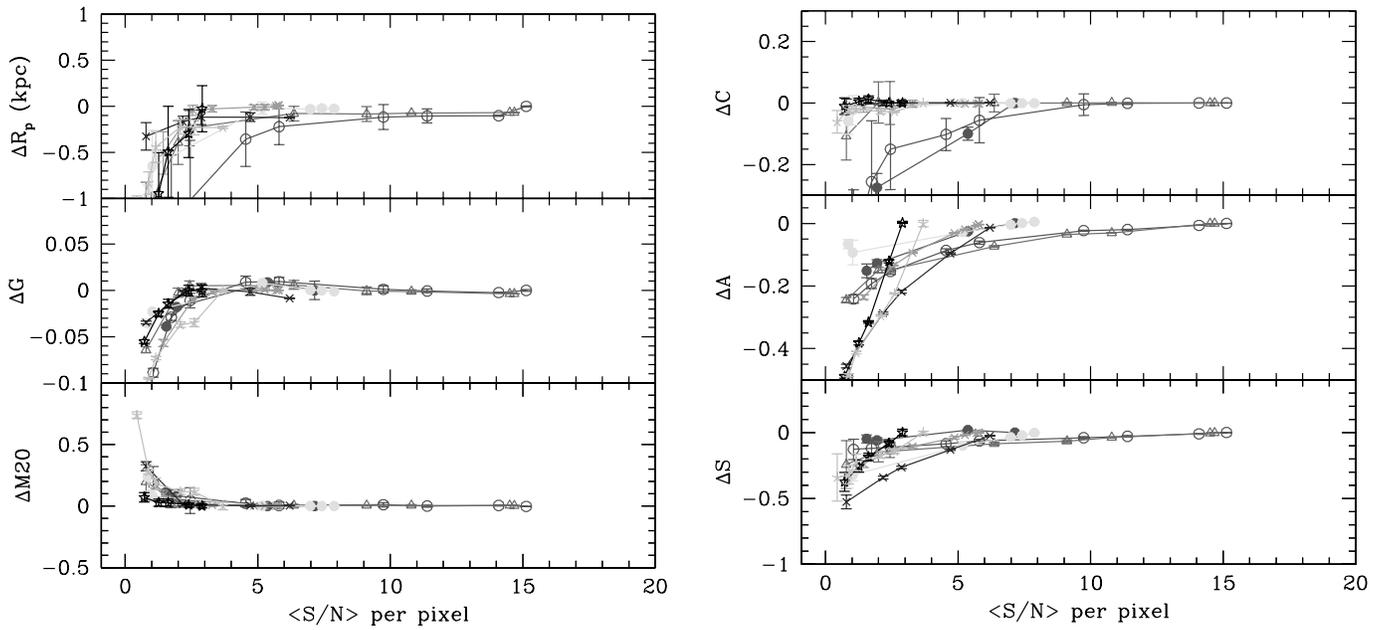


FIG. 5.— $\Delta R_p$ ,  $\Delta G$ ,  $\Delta M_{20}$ ,  $\Delta C$ ,  $\Delta A$ , and  $\Delta S$  vs.  $\langle S/N \rangle$  per pixel: E/NGC 5322 (filled circles), S0/NGC 4526 (open circles), Sab/NGC 3368 (triangles), Sbc/NGC 3953 (squares), Sc/NGC 2403 (light crosses), Sd/NGC 4713 (dark crosses), Arp 220 (light stars), Super Antenna (black stars).

and the  $r$ -band PSF FWHM values are typically  $\sim 1''.3$ – $1''.8$  (Stoughton et al. 2002). We find that the mean absolute difference between the SDSS and Frei observations are:

$$\begin{aligned}
 & r/R \text{ band:} \\
 & \delta G = 0.02 \quad \delta M_{20} = 0.12 \quad \delta C = 0.11 \\
 & \delta A = 0.04 \quad \delta S = 0.09, \\
 & g/B \text{ band:} \\
 & \delta G = 0.02 \quad \delta M_{20} = 0.11 \quad \delta C = 0.14 \\
 & \delta A = 0.05 \quad \delta S = 0.14. \tag{12}
 \end{aligned}$$

In addition, we have analyzed  $B$ -band images of 22 nearby dwarf irregular galaxies from the Van Zee (2001) sample (Table 4). We have selected galaxies from the original Van Zee sample with minimal foreground star contamination and  $\langle S/N \rangle \geq 2$ . These images were obtained at the Kitt Peak 0.9 m telescope and have PSF FWHM  $\sim 1''.4$ – $2''.3$  and a plate scale =  $0''.688$ .

In Figures 7–8 we examine the dependence of  $C$ ,  $A$ ,  $S$ ,  $G$ , and  $M_{20}$  on the observed near-ultraviolet/optical wavelength. For the majority of galaxies, the differences between the observed morphologies at  $\sim 4500 \text{ \AA}$  ( $B/g$ ) and  $6500 \text{ \AA}$  ( $R/r$ ) are comparable to the observational offsets between the SDSS and

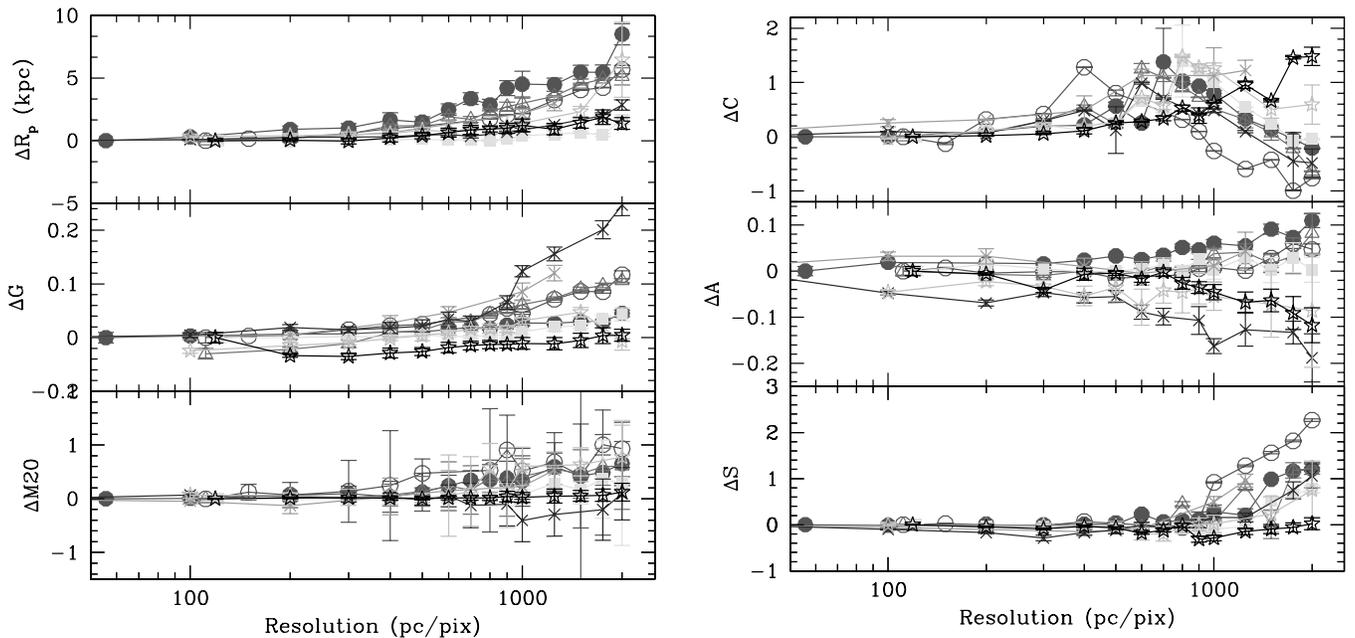


FIG. 6.— $\Delta R_p$ ,  $\Delta G$ ,  $\Delta M_{20}$ ,  $\Delta C$ ,  $\Delta A$ , and  $\Delta S$  vs. resolution (parsecs per pixel). Symbols are same as Fig. 5.

TABLE 2  
FREI GALAXY CATALOG<sup>a</sup>

Galaxy	Type <sup>b</sup>	$M_B^c$	$m - M$	S/N <sub>B</sub>	$C_B$	$A_B$	$S_B$	$G_B$	$M_{20,B}$	S/N <sub>R</sub>	$C_R$	$A_R$	$S_R$	$G_R$	$M_{20,R}$
NGC 2768.....	S0_1/2	-20.9	31.75 <sup>f</sup>	7.7	4.33	-0.01	0.08	0.59	-2.39	7.9	4.32	-0.02	0.06	0.59	-2.45
NGC 3377.....	E6	-19.0	30.25 <sup>f</sup>	5.6	4.77	-0.01	-0.01	0.63	-2.58	5.0	4.99	-0.02	-0.01	0.64	-2.67
NGC 3379.....	E1	-19.9	30.12 <sup>f</sup>	9.0	4.61	-0.01	-0.02	0.59	-2.52	7.8	4.83	-0.01	-0.02	0.61	-2.54
NGC 4125 <sup>d</sup> .....	E6/S0_1/2	-21.2	31.89 <sup>f</sup>	7.1	4.30	0.02	0.04	0.60	-2.28	3.4	4.70	-0.05	0.00	0.63	-2.38
NGC 4365.....	E3	-21.0	31.55 <sup>f</sup>	3.4	4.40	-0.08	-0.06	0.61	-2.43	7.6	4.53	-0.01	0.01	0.59	-2.50

NOTE.—Table 2 is presented in its entirety in the electronic edition of the *Astronomical Journal*. A portion is shown here for guidance regarding its form and content.

<sup>a</sup> From Frei et al. (1996); see also <http://www.astro.princeton.edu/~frei/catalog.htm>.

<sup>b</sup> From Sandage & Bedke (1994).

<sup>c</sup>  $B$  from RC3.

<sup>d</sup> Also observed by SDSS Data Release 1; see Table 3.

<sup>e</sup> Distance from redshift obtained from the NASA Extragalactic Database (NED), assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

<sup>f</sup> From Jensen et al. (2003).

Frei observations of the same galaxies in the same bandpass. The observed changes in  $C$ ,  $G$ , and  $M_{20}$  from  $\sim 3600 \text{ \AA}$  ( $u$ ) to  $\sim 6500 \text{ \AA}$  ( $r$ ) are also consistent with observational scatter. The SDSS  $u$ -band observations often have too low S/N to obtain reliable asymmetries. This may also produce the increased scatter in  $S$ . Nevertheless, late-type galaxies generally have higher clumpiness values and slightly higher  $M_{20}$  values at 3600 than 6500  $\text{\AA}$ . A handful of galaxies (many of which are edge-on spirals) show much larger morphological changes at bluer wavelengths. The S0 galaxy UGC 1597 has an obvious tidal tail, and it has higher  $g$ -band  $A$ ,  $S$ , and  $G$  values and a lower  $g$ -band  $C$ . Several mid-type spirals have significantly higher  $M_{20}$  values in  $B/g$  than in  $R/r$ . These include NGC 3675, an Sb with prominent dust features, and NGC 5850, an Sb with a star-forming ring.

Previous studies have noted small offsets in concentration and asymmetry from  $U$  and  $B$  to  $R$ , with much stronger shifts at wavelengths  $\leq 2500 \text{ \AA}$  (Brinchmann et al. 1998; Conselice et al. 2000; Kuchinski et al. 2001). We see similar trends of slightly higher  $B/g$  asymmetries for late-type spirals ( $\sim 0.05$ ) and lower  $B/g$  concentrations for most galaxies ( $\sim 0.1$ ). However, given that these trends are smaller than the difference between different observations of the same galaxy at the same wavelength, we conclude that morphological  $K$ -corrections to  $C$  and  $A$  are not very substantial for most normal galaxies observed redward of rest-frame  $\sim 3500\text{--}4000 \text{ \AA}$ . The late-type spirals show small but systematic trends of stronger clumpiness and higher second-order moments at bluer wavelengths.

In Figure 9 we examine the  $G$ - $M_{20}$  morphologies of local galaxies observed in both the  $R/r$  and  $B/g$  bands. The distribution of local galaxies is very similar at both wavelengths, with E/S0's showing high  $G$  and low  $M_{20}$  values, Sa–Sbc at intermediate  $G$  and  $M_{20}$  values, and most late-type spirals and dI's with low  $G$  and higher  $M_{20}$  values. Most edge-on galaxies (*barred symbols*) show  $G$  and  $M_{20}$  values consistent with the mean values for their Hubble type. One notable exception is the S0 NGC 4710, which has a prominent dust lane and  $G = 0.50$ ,  $\sim 0.1$  lower than for other E/S0's. The majority of local galaxies lie below the rough dividing line plotted in Figure 9. Four out of the 22 dI's lie above this line. Two of these are classified as starbursting dwarfs (UGC 11755 and UGC A439), and a third has the bluest  $U - B$  color gradient in the sample (UGC 5288; Van Zee 2001). The other outliers are UGC 10991, which appears to have a tidal tail and star-forming knots, and UGC 10310, which has two very bright knots in its outer arms that may be foreground stars. As we discuss in the

next section, most ULIRGs lie above this dividing line. While a few truly starbursting dI's are  $\sim 0.04$  in  $G$  above the normal galaxy sequence at blue wavelengths, it appears that dI's will not seriously contaminate the merger/interacting galaxies classified by  $G$ - $M_{20}$ .

#### 4.2. Merger Indicators

One of the primary goals of morphological studies is to quantitatively identify interacting and merging galaxies. Toward this end, Abraham et al. (1996) and Conselice et al. (2000, 2003) have used combinations of concentration, asymmetry, and smoothness to roughly classify “normal” galaxies as early and late types, as well as to distinguish mergers from these normal types. Abraham (2003) also found that for a large sample of normal galaxies, the Gini coefficient is strongly correlated with concentration, color, and surface brightness, and therefore may be as efficient as concentration at quantifying galaxy morphologies. Here we compare the effectiveness of our definition of the Gini coefficient (eq. [6]) to  $C$  at classifying local galaxy types and identifying merger candidates. We also expect that  $M_{20}$  will be strongly correlated to  $C$ , because of their similar definitions, and therefore examine the  $G$ - $M_{20}$  correlation and compare it with the  $C$ - $G$  relation found by Abraham et al. (2003).

In Figures 10–14 we compare the  $R/r$ -band morphological distributions of our local galaxy sample with archival *HST* WFPC2 F814W observations of 73 ultraluminous infrared galaxies (ULIRGs) with  $11.5 \leq \log(L_{\text{FIR}}/L_{\odot}) \leq 12.5$  and  $\langle \text{S/N} \rangle \geq 2$  (Borne et al. 2000, *HST* Cycle 6 program 6346, Table 5). ULIRGs often show morphological signatures of ongoing or recent merger events in the form of high asymmetries, multiple nuclei, and tidal tails (Wu et al. 1998; Borne et al. 2000; Conselice et al. 2000; Cui et al. 2001). We have divided the ULIRG sample into objects with “single,” “double,” or “multiple” nuclei as classified by Cui et al. (2001) by counting the number of surface brightness peaks with  $\text{FWHM} > 0''.14$  and  $M_I < -17.0$  separated by less than 20 kpc projected. We also identify ULIRGs in projected pairs as *IRAS* sources with projected separations greater than 20 kpc and less than 120 kpc. The ULIRG sample has a mean redshift of  $\sim 0.2$ ; therefore, the F814W bandpass ( $\lambda_{\text{eff}} = 8200 \text{ \AA}$ ) samples the rest-frame light at  $\sim 6800 \text{ \AA}$ . Given the  $0''.14$  PSF of the WF camera, ULIRGs at  $z < 0.25$  are spatially resolved to better than  $\sim 500 \text{ pc}$  and may be directly compared with the local galaxy  $r/R$ -band observations.

Most ULIRGs lie above the  $G$ - $M_{20}$  correlation for normal galaxies (Fig. 10, *bottom*), while many ULIRGs overlap with

TABLE 3  
SDSS *u*-SELECTED CATALOG<sup>a</sup>

Galaxy	Type <sup>b</sup>	$M_B^c$	$m - M$	$S/N_u$	$C_u$	$A_u$	$S_u$	$G_u$	$M_{20,u}$	$S/N_g$	$C_g$	$A_g$	$S_g$	$G_g$	$M_{20,g}$	$S/N_r$	$C_r$	$A_r$	$S_r$	$G_r$	$M_{20,r}$
NGC 3640.....	E3	-19.0	30.37 <sup>f</sup>	1.6	4.26	-0.14	-0.13	0.60	-2.34	6.1	4.55	0.00	0.00	0.63	-2.45	8.3	4.45	0.01	0.00	0.62	-2.45
NGC 4073.....	E5	-22.2	34.65 <sup>c</sup>	0.7	3.95	-0.22	-0.04	0.47	-2.02	2.1	4.51	-0.08	-0.02	0.62	-2.58	2.8	4.64	-0.06	-0.03	0.62	-2.58
NGC 4125 <sup>d</sup> .....	E6/S0_1/2	-21.2	31.89 <sup>f</sup>	0.9	4.37	-0.17	-0.11	0.56	-2.36	5.2	4.54	-0.04	0.07	0.63	-2.51	7.1	4.64	-0.04	0.07	0.63	-2.54
NGC 4261.....	E3	-21.1	32.50 <sup>f</sup>	0.9	4.56	-0.07	0.48	0.55	-2.55	4.4	4.64	-0.03	-0.08	0.62	-2.59	6.8	4.71	-0.01	-0.05	0.62	-2.58
NGC 4636 <sup>d</sup> .....	E0/S0_1_(6)	-20.4	30.83 <sup>f</sup>	0.6	3.88	-0.23	-0.04	0.46	-1.63	2.9	4.28	-0.06	-0.10	0.60	-2.54	3.9	4.33	-0.05	-0.08	0.60	-2.56

NOTE.—Table 3 is presented in its entirety in the electronic edition of the *Astronomical Journal*. A portion is shown here for guidance regarding its form and content.

<sup>a</sup> See Abazajian et al. (2003); see also <http://www.sdss.org/dr1>.

<sup>b</sup> From Sandage & Bedke (1994).

<sup>c</sup>  $B$  from RC3.

<sup>d</sup> Also observed by Frei et al. (1996); see Table 2.

<sup>e</sup> Distance from redshift obtained from NED, assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

<sup>f</sup> From Jensen et al. (2003).

TABLE 4  
 VAN ZEE DWARF IRREGULAR GALAXY SAMPLE<sup>a</sup>

Galaxy	Type <sup>b</sup>	$M_B$	$S/N_B$	$C_B$	$A_B$	$S_B$	$G_B$	$M_{20,B}$
UGC 00290.....	Sdm	-14.48	2.1	3.75	0.03	0.30	0.54	-1.78
UGC 00634.....	SABm	-17.67	2.5	2.85	0.14	0.61	0.49	-1.15
UGC 00685.....	SAm	-14.74	5.0	2.81	0.18	0.39	0.52	-1.47
UGC 00891.....	SABm	-15.53	3.1	2.90	0.06	0.38	0.49	-1.56
UGC 01104.....	Im	-16.08	9.2	3.18	0.20	0.36	0.50	-1.69
UGC 01175.....	BCD/E	-14.13	3.8	2.84	0.24	0.27	0.47	-1.38
UGC 03647.....	IBm	-17.06	2.0	2.87	0.12	0.56	0.54	-1.47
UGC 04117.....	IBm	-14.86	4.7	2.89	0.28	0.67	0.48	-0.94
UGC 05205.....	SBm pec	-16.29	4.5	2.72	0.23	0.28	0.48	-1.77
UGC 05288.....	Sdm	-14.44	4.2	3.32	0.13	0.43	0.59	-1.54
UGC 09219.....	Im	-16.33	5.1	2.78	0.36	0.57	0.50	-1.53
UGC 09240.....	IAm	-14.60	5.1	2.54	0.24	0.55	0.50	-1.30
UGC 10054.....	SBdm	-18.13	3.6	3.06	0.08	0.45	0.47	-1.62
UGC 10310.....	SB(s)m	-16.81	3.7	2.39	0.41	1.29	0.50	-0.70
UGC 10351.....	Sdm	-16.06	5.9	3.29	0.15	0.34	0.56	-1.58
UGC 10445.....	SBc	-17.53	4.7	2.53	0.30	0.73	0.50	-1.04
UGC 10991.....	Im	-16.35	2.3	2.68	0.41	0.78	0.54	-1.08
UGC 11755.....	BCD/E	-17.14	9.1	4.17	0.15	0.39	0.61	-1.79
UGC 12713.....	S0/a	-14.76	6.7	2.96	0.18	0.31	0.53	-1.47
UGC A009.....	IB(s)m	-14.17	3.0	2.69	0.01	0.34	0.49	-1.30
UGC A015.....	IB(s)m	-12.93	2.4	2.49	0.10	0.47	0.47	-1.09
UGC A439.....	BCD	-16.73	17.5	3.93	0.12	0.29	0.57	-1.30

<sup>a</sup> See Van Zee (2001).

<sup>b</sup> Morphological types from RC3.

the  $C$ - $G$  and  $C$ - $M_{20}$  correlations for normal galaxies (Fig. 11). Normal local galaxies also segregate more cleanly from the ULIRGs sample in  $G$ - $A$  and  $G$ - $S$  than  $C$ - $A$  and  $C$ - $S$  (Figs. 12–13). In particular, the Gini coefficient of edge-on spirals galaxies is more consistent with the values obtained for face-on spirals. In addition, ULIRGs with double or multiple nuclei

generally have higher Gini coefficients relative to their concentrations than most normal galaxies.  $G$ - $M_{20}$  is slightly less effective at identifying single-nuclei ULIRGs than  $G$ - $A$  and  $G$ - $S$ ; however,  $M_{20}$  is a more robust indicator at low  $S/N$  than  $A$  and  $S$  and at low resolution than  $S$  (Figs. 5–6), and therefore may be applied to fainter galaxy populations. We find that  $M_{20}$ , in combination with  $C$ ,  $A$ , and  $S$ , is not effective at separating the ULIRGs from the normal galaxy population (Figs. 12–14).

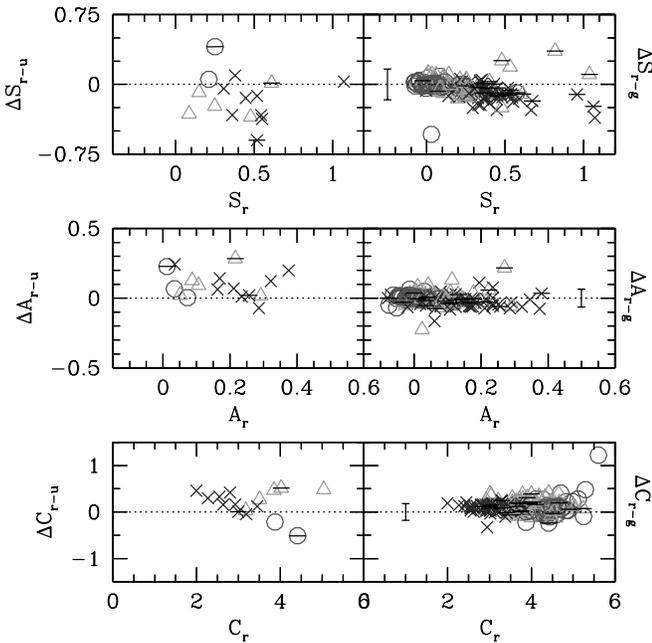


FIG. 7.—Change in  $CAS$  morphology from  $\sim 6500 \text{ \AA}$  ( $R/r$ ) to  $\sim 3600 \text{ \AA}$  ( $u$ ) for SDSS  $u$ -selected sample with  $S/N_u > 2.0$  (left) and to  $\sim 4500 \text{ \AA}$  ( $B/g$ ) for Frei and SDSS galaxies with  $S/N > 2.0$  (right). The error bars are  $(\delta_r^2 + \delta_g^2)^{1/2}$ , where  $\delta$  is the average difference between SDSS and Frei et al. observations of the same galaxies: E/S0 (circles), Sa–Sbc (triangles), and Sc–Sdm (crosses).

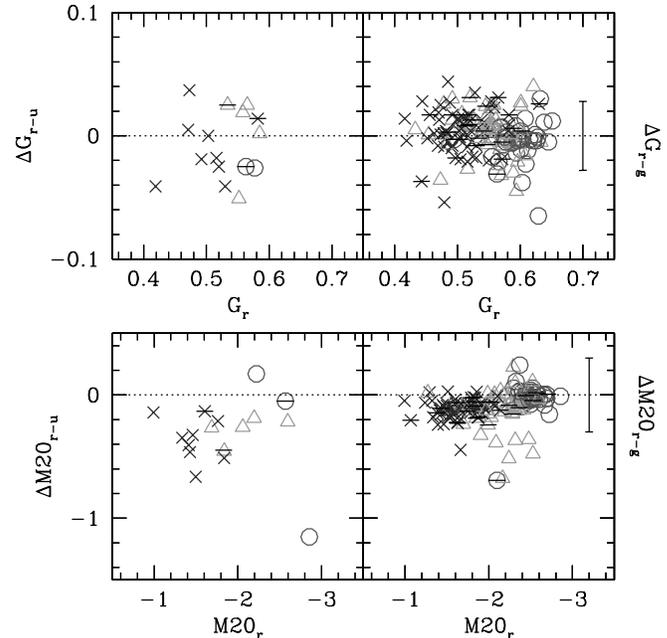


FIG. 8.—Change in  $G, M_{20}$  morphology from  $\sim 6500 \text{ \AA}$  ( $R/r$ ) to  $\sim 3600 \text{ \AA}$  ( $u$ ) for SDSS  $u$ -selected sample with  $S/N_u > 2.0$  (left) and to  $\sim 4500 \text{ \AA}$  ( $B/g$ ) for Frei and SDSS galaxies with  $S/N > 2.0$  (right). Error bars and point symbols are same as Fig. 7.

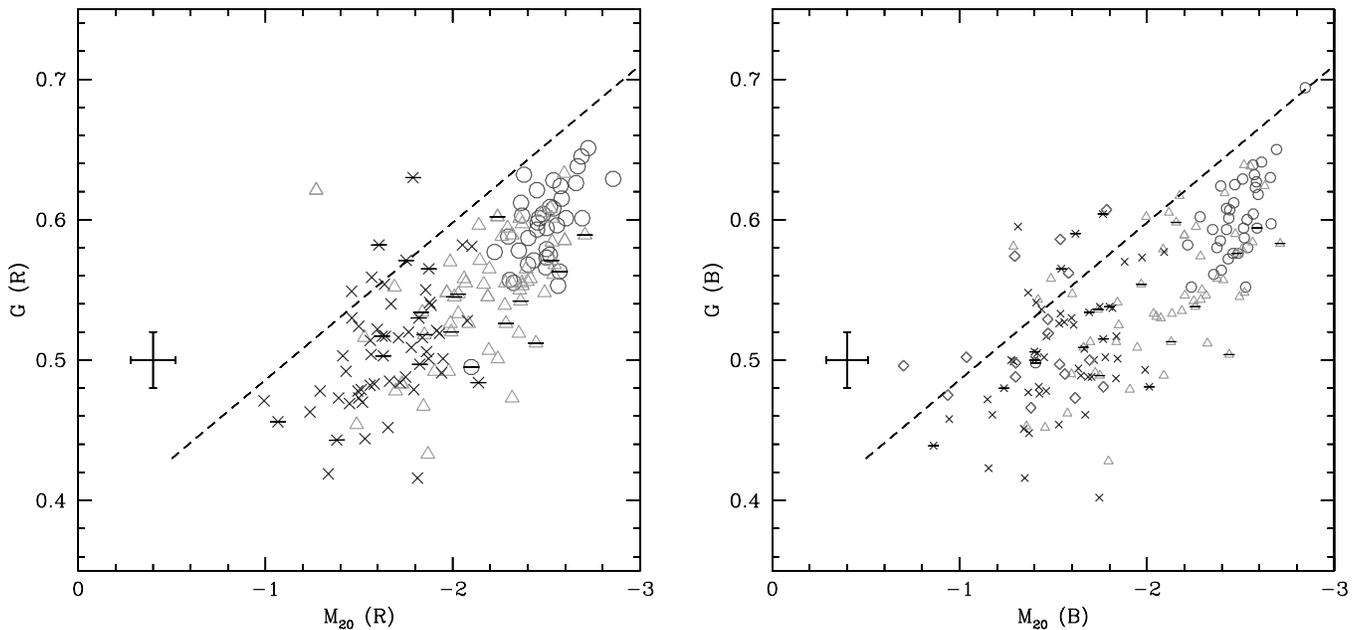


FIG. 9.— $M_{20}$  vs.  $G$  for rest-frame  $\sim 6500 \text{ \AA}$  (left) and  $4400 \text{ \AA}$  (right) observations of local galaxies (circles: E/S0; triangles: Sa–Sbc; crosses: Sc–Sd; diamonds: dl; bars: edge-on spirals). The error bars are mean difference in  $G$  and  $M_{20}$  between SDSS  $r$ -band and Frei  $R/r$  observations of the same objects. Almost all the “normal” galaxies lie below the dashed line in the  $R$ -band plot. The outlying Sb galaxy NGC 5850 has a strong star-forming ring and is in a close pair with NGC 5846. Three of the outlying dl’s in the  $B$ -band plot are starbursting.

In Table 6, we give the results of a series of two-dimensional Kolmogorov-Smirnov (K-S) tests (Fasano & Franceschini 1987) applied to the ULIRGs and  $R/r$ -band normal galaxy observations for each combination of  $C$ ,  $A$ ,  $S$ ,  $G$ , and  $M_{20}$ . For all the tests, the probability that the ULIRGs and normal galaxies are drawn from the same parent sample is less than  $10^{-6}$ .

While the ULIRG population as a whole occupies a different region of  $C$ - $A$ - $S$ - $G$ - $M_{20}$  space than our SDSS and Frei galaxy samples, we also find significant differences between ULIRGs in well-separated pairs, ULIRGs with single nuclei, and ULIRGs with double or multiple nuclei (Table 6). ULIRGs in pairs show the smallest offsets from the normal galaxy sample. Double- and multi-nuclei ULIRGs show the greatest changes in morphology, with typically large  $M_{20}$  and  $A$  values. Single-nucleus ULIRGs appear similar to paired ULIRGs but can also have higher  $G$  and  $C$ . Two-dimensional K-S tests show that the multi- and double-nuclei ULIRGs are distinct from the single-nucleus ULIRGs and paired ULIRGs with greater than 97% and 90% confidence, respectively. The multi- and double-nuclei ULIRGs have a greater than 5% probability of being drawn from the same sample, while single-nucleus ULIRGs and ULIRGs in pairs have a greater than 12% probability of being drawn from the same sample.

## 5. GALAXY MORPHOLOGIES AT REDSHIFT $> 2$

One of the major successes of the hierarchical paradigm of galaxy formation has been the discovery of large fractions of morphologically irregular galaxies at  $z > 1$  (e.g., Driver et al. 1995; Abraham et al. 1996; Odewahn et al. 1996; Abraham & van den Bergh 2001). Many of these galaxies are excellent merger candidates and suggest merger fractions between 25%–40% at  $0.5 < z < 2$ . However, morphological studies of the most distant galaxies—the Lyman-break galaxies (LBGs)—have produced confusing and conflicting conclusions. Initial *HST* WFPC2 observations of the rest-frame far-ultraviolet morphologies of 20  $z > 3$  galaxies found that they possessed

one or more compact “cores” with sizes similar to present-day spiral bulges (Giavalisco et al. 1996). More recent ACS observations of large numbers of  $2 < z < 6$  LBGs have confirmed ultraviolet half-light radii between 1.5 and 3.5 kpc and concentrations similar to local bulges and ellipticals (Ferguson et al. 2004). However, these LBGs have an ellipticity distribution more like disk galaxies than ellipsoids, leading to the conclusion that LBGs are drawn from a mixture of morphological

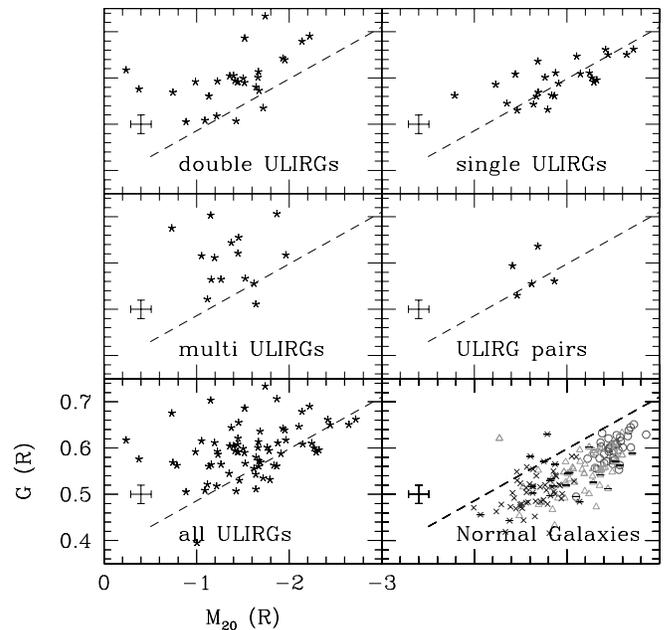


FIG. 10.— $M_{20}$  vs.  $G$  for rest-frame  $\sim 6500 \text{ \AA}$  observations of local galaxies (circles: E/S0; triangles: Sa–Sbc; crosses: Sc–Sd; stars: ULIRGs; bars: edge-on spirals). The error bars are mean difference in  $G$  and  $M_{20}$  between SDSS  $r$ -band and Frei  $R/r$  observations of the same objects. Almost all the “normal” galaxies lie below the dashed line.

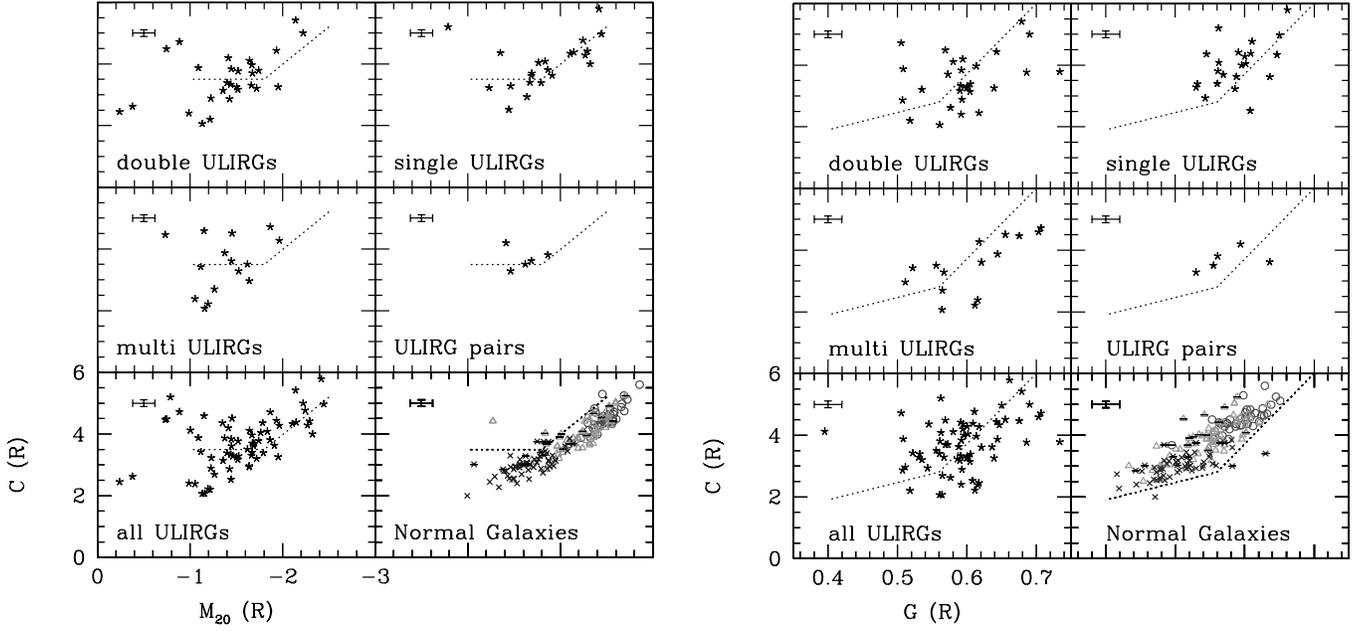


FIG. 11.— $M_{20}$  vs.  $C$  and  $G$  vs.  $C$  for rest-frame  $\sim 6500 \text{ \AA}$  observations of local galaxies. Symbols are same as Fig. 10. The majority of normal galaxies lie to one side of the dashed lines.

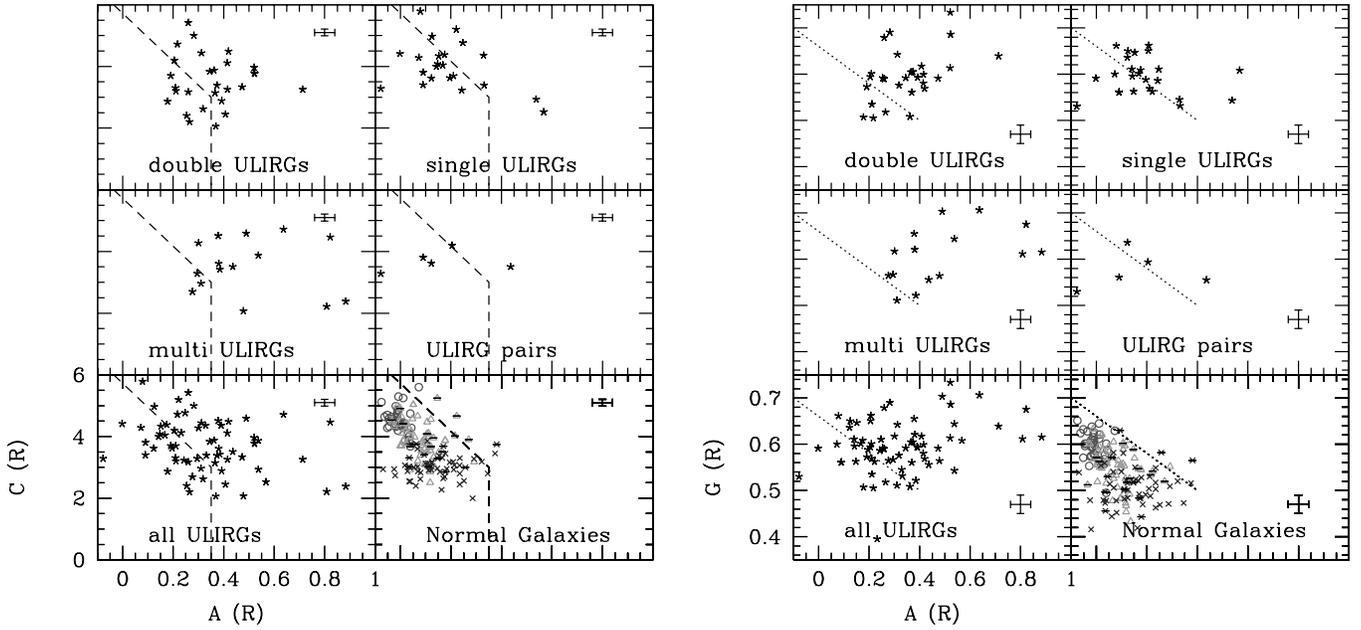


FIG. 12.— $A$  vs.  $C$  and  $G$  vs.  $A$  for rest-frame  $\sim 6500 \text{ \AA}$  observations of local galaxies. Symbols are same as Fig. 10. The majority of normal galaxies lie to one side of the dashed lines.

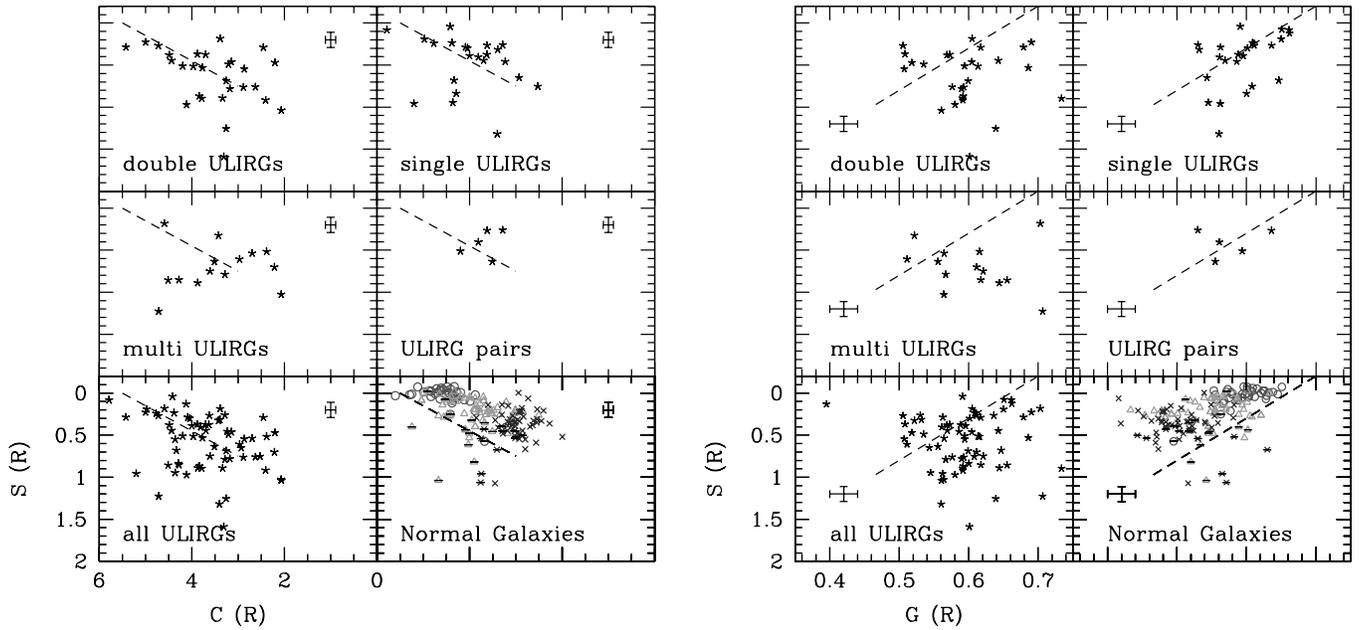


FIG. 13.— $S$  vs.  $C$  and  $G$  for rest-frame  $\sim 6500 \text{ \AA}$  observations of local galaxies. Symbols are same as Fig. 10.

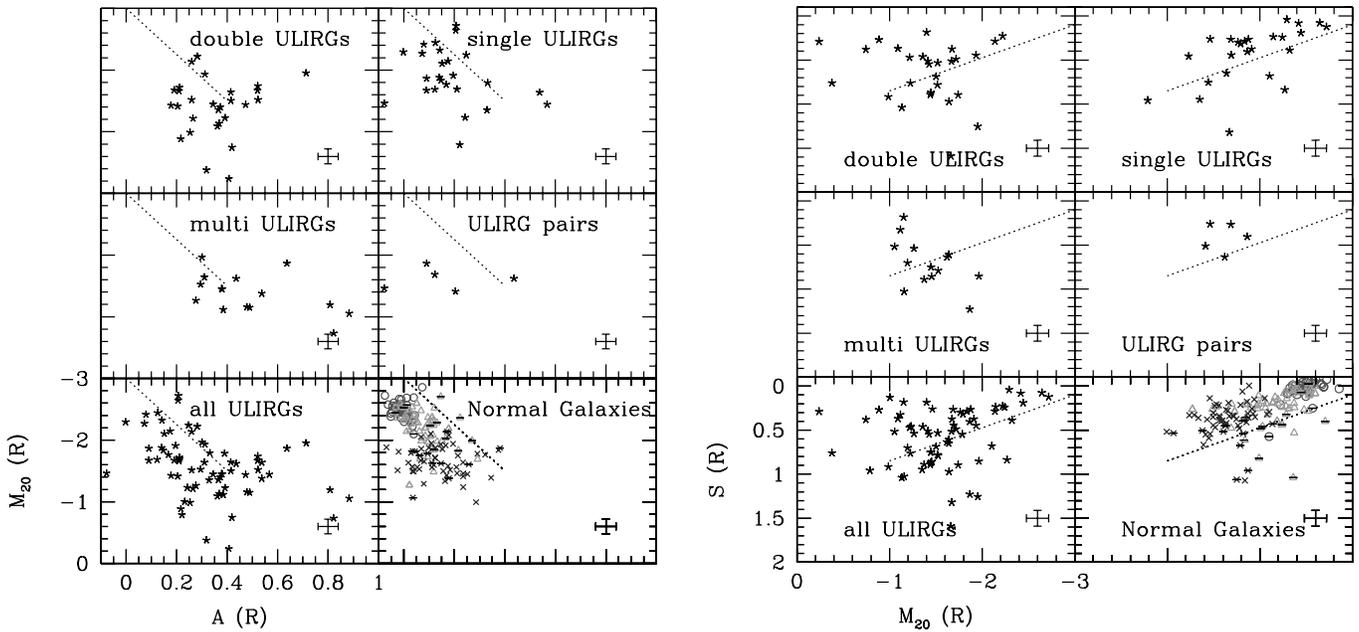


FIG. 14.— $A$  vs.  $M_{20}$  and  $M_{20}$  vs.  $S$  for rest-frame  $\sim 6500 \text{ \AA}$  observations of local galaxies. Symbols are same as Fig. 10.

TABLE 5  
BORNE ET AL. 2000 ULIRG *HST* SURVEY<sup>a</sup>

Galaxy	Redshift <sup>b</sup>	R.A.	Decl.	S/N	<i>C</i>	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$	Nuclei <sup>c</sup>
IRAS 00091–0738	0.118	00 11 43	–07 22 07	2.0	3.63	0.20	0.38	0.59	–1.91	Single*
IRAS 00104–0139	0.163	00 13 04	–01 23 05	2.1	2.20	0.26	0.47	0.52	–1.22	Double
IRAS 00161–0850	0.109	00 18 43	–08 33 36	3.7	2.69	0.28	0.54	0.56	–1.26	Multiple
IRAS 00207+1029	0.231	00 23 22	+10 46 22	3.3	3.79	0.52	0.90	0.73	–1.74	Double
IRAS 00335–2732	0.069	00 36 01	–27 15 35	5.0	4.33	0.15	0.68	0.65	–2.11	Single

NOTE.—Table 5 is presented in its entirety in the electronic edition of the *Astronomical Journal*. A portion is shown here for guidance regarding its form and content. Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds.

<sup>a</sup> See Borne et al. (2000); *HST* WFPC2 F814W observations.

<sup>b</sup> Redshifts taken from NED.

<sup>c</sup> From Cui et al. (2001) classification, except where starred.

TABLE 6  
TWO-DIMENSIONAL K-S TEST PROBABILITIES

Parameter	K-S Probabilities			
“Normal” Galaxies versus ULIRGs				
	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$
<i>C</i> .....	6.6e-15	9.3e-8	7.5e-7	8.3e-8
<i>A</i> .....	...	2.1e-16	7.7e-18	6.2e-14
<i>S</i> .....	...	...	1.6e-13	5.1e-10
<i>G</i> .....	...	...	...	5.8e-13
Double versus Multi-Nuclei ULIRGs				
	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$
<i>C</i> .....	0.29	0.25	0.15	0.66
<i>A</i> .....	...	0.047	0.17	0.086
<i>S</i> .....	...	...	0.086	0.36
<i>G</i> .....	...	...	...	0.18
Paired versus Single-Nuclei ULIRGs				
	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$
<i>C</i> .....	0.25	0.21	0.29	0.14
<i>A</i> .....	...	0.51	0.49	0.13
<i>S</i> .....	...	...	0.47	0.30
<i>G</i> .....	...	...	...	0.12
Double + Multi- versus Single-Nuclei ULIRGs				
	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$
<i>C</i> .....	8.2e-7	0.0012	0.029	3.3e-4
<i>A</i> .....	...	7.1e-6	7.2e-6	9.3e-7
<i>S</i> .....	...	...	0.0024	3.2e-5
<i>G</i> .....	...	...	...	7.6e-4
Double + Multi-Nuclei versus Paired ULIRGs				
	<i>A</i>	<i>S</i>	<i>G</i>	$M_{20}$
<i>C</i> .....	0.0041	0.017	0.078	0.057
<i>A</i> .....	...	0.0042	0.020	0.0070
<i>S</i> .....	...	...	0.096	0.016
<i>G</i> .....	...	...	...	0.056

NOTE.—Two-dimensional K-S test probabilities that two galaxy populations have the same distribution in a two-parameter space.

TABLE 7  
HDF-N LYMAN-BREAK GALAXIES<sup>a</sup>

ID <sup>a</sup>	Redshift <sup>b</sup>	R.A. (J2000.0)	Decl. (J2000.0)	$H^a$	$M_B$	$r_p$ (arcsec)	S/N	$C$	$A$	$G$	$M_{20}$
882.....	1.8	12 36 45.66	62 12 41.9	22.27	-21.6	0.62	8.2	4.4	0.08	0.64	-1.86
1118.....	1.8	12 36 51.28	62 12 33.8	24.47	-19.7	0.31	4.0	3.3	0.12	0.63	-1.45
1217.....	1.8	12 36 45.15	62 12 05.5	24.77	-19.5	0.41	3.1	3.5	0.02	0.55	-1.19
121.....	1.9	12 36 44.85	62 14 06.1	23.14	-20.9	0.36	4.2	4.0	0.07	0.68	-1.77
653.....	1.9	12 36 56.64	62 13 39.9	23.87	-20.3	0.63	4.0	...	...	...	...

NOTE.—Table 7 is presented in its entirety in the electronic edition of the *Astronomical Journal*. A portion is shown here for guidance regarding its form and content.

<sup>a</sup> From M. Dickinson (2003, private communication).

<sup>b</sup> Photometric redshifts; see Budavari et al. (2000).

types. Rest-frame optical observations in the near-infrared with NICMOS have shown that the observed LBG morphologies are not a strong function of wavelength (Papovich et al. 2001; Dickinson 1999) and that LBGs have internal far-UV–optical color dispersions much smaller than  $z \sim 1$  galaxies (Papovich 2002). LBGs are significantly bluer than local galaxies, and it is likely that their ultraviolet and optical morphologies are dominated by young stars. Their small sizes, high concentrations, and high star formation rates suggest that many are precursors to local spiral bulges. However, surface brightness dimming may prevent the detection of faint tidal tails and some appear to possess multiple nuclei. In a recent study of the optical morphologies of the Hubble Deep Field North galaxies, Conselice et al. (2003) found that seven out of 18  $z < 3$ ,  $M_B < -21$  galaxies possess corrected asymmetries greater than 0.35, implying that up to 50% are recent mergers. However, as we found in § 4, asymmetry is not as sensitive by itself at detecting merger remnants as it is in combination with  $C$  or  $G$ . Here we reexamine the optical morphologies of the HDFN high-redshift galaxy sample using  $C$ ,  $G$ , and  $M_{20}$ , and we attempt to classify these galaxies as ellipticals, disks, or recent mergers.

The Hubble Deep Field North has 27 spectroscopically confirmed high-redshift galaxies and 70 additional candidates with  $1.7 < z < 4$  and  $H < 25.0$  (Papovich et al. 2001 and references therein). At these redshifts the near-ultraviolet and optical regions of the galaxies spectral energy distributions have been shifted to redward of  $1 \mu\text{m}$  and therefore require infrared observations to directly compare their morphologies with the rest-frame near-UV/optical morphologies of local galaxies. The HDFN has been observed with the NICMOS camera 3 in the F110W ( $J$ ) and F160W ( $H$ ) bandpasses ( $\lambda_{\text{eff}} = 1.1, 1.6 \mu\text{m}$ ) down to a  $10 \sigma$  limiting magnitude of 26.5 (Dickinson 1999, *HST* Cycle 7 program 7817). Most of the HDFN LBGs are fainter than  $H = 23.0$ ; therefore, to increase their S/N per pixel, we have measured the morphologies of the LBG sample in a summed F110W and F160W image. The effective central wavelength of the summed LBG observations is  $\sim 1.3 \mu\text{m}$ . Galaxies at  $z \sim 2$  and 3 are observed at rest-frame wavelengths 4300 and 3250 Å, respectively. Out of our initial sample of 97  $H < 25$ ,  $1.7 < z \leq 3.8$  galaxies, 33 galaxies with  $1.7 < z < 2.3$ , and 16 galaxies with  $2.3 \leq z \leq 3.8$  have  $\langle \text{S/N} \rangle > 2.0$  (Table 7). We also give estimates of the rest-frame  $M_B$  in AB magnitudes, computed by interpolating between the  $J$ ,  $H$ , and  $K_s$  fluxes and assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_m = 0.3$  cosmology.

The NICMOS images offer the highest available resolution at near-UV/optical wavelengths for these galaxies. Nevertheless,

the physical resolution of the  $z > 2$  galaxies is significantly worse than that for the local galaxy images. The dithered NIC3 observations have a pixel scale =  $0''.08 \text{ pixel}^{-1}$  and a PSF FWHM =  $0''.22$ . At  $z \geq 2$  this corresponds to a physical pixel scale of  $\sim 670 \text{ pc}$  and PSF FWHM  $\sim 1.8 \text{ kpc}$ . Our simulations in § 3 showed that these resolutions produce strong biases in the measured morphologies, which are often a function of morphological type. The well-defined correlations of local galaxy morphologies are likely to change significantly with these biases. Therefore, we compare the LBG morphologies with local galaxy images that have been measured from degraded  $u$ - and  $B/g$ -band images. The galaxies are selected to lie in the same ( $M_B - M^*$ ) range (Fig. 15), assuming  $M^* = -20.1$  locally (Blanton et al. 2003b) and  $M^* = -22.9$  at  $z \geq 2$  (Shapley et al. 2001). This selection tests a “passive” evolutionary scenario, in which the local galaxies were brighter in the past but did not evolve morphologically. We select local galaxies observed in  $B/g$  with  $3.5 \geq (M_B - M^*) \geq 1.0$  to compare with a similarly selected  $z \sim 2$  sample, and local galaxies observed in  $u$  with  $2.5 \geq (M_B - M^*) \geq -0.5$  to compare with the  $z \sim 3$  sample.

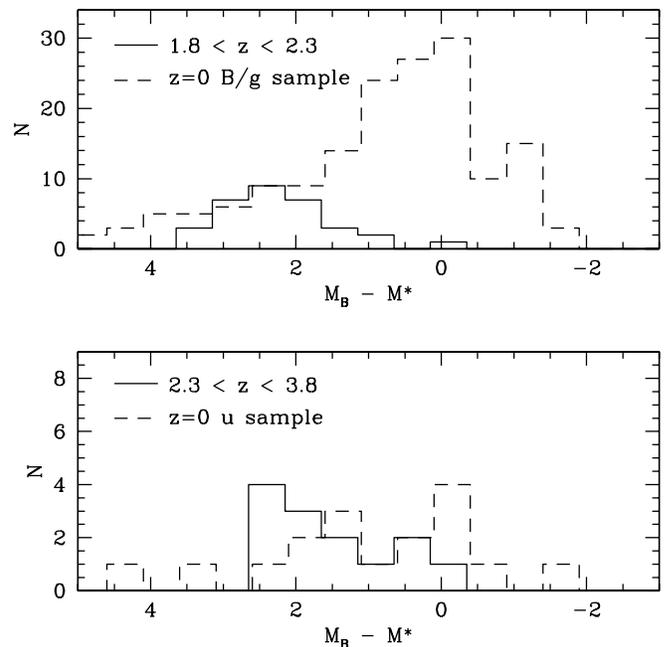


FIG. 15.— $M_B - M^*$  histograms for the  $z \sim 2$  and  $B/g$ -band local galaxy samples (top) and the  $z \sim 3$  and  $u$ -band local galaxy samples (bottom).  $M^*$  is assumed to be  $-20.1$  for the local galaxies and  $-22.9$  for the  $z \geq 2$  galaxies.

TABLE 8  
 $z = 2$  AND 3 MORPHOLOGICAL SIMULATIONS

Type	$N$	$C_z - C_{z=0}$	$A_z - A_{z=0}$	$S_z - S_{z=0}$	$G_z - G_{z=0}$	$M_{20,z} - M_{20,z=0}$
$z = 2$ , Rest-Frame $B/g$ Observations						
E/S0.....	36	+0.14 $\pm$ 0.69	+0.07 $\pm$ 0.05	+0.29 $\pm$ 0.19	-0.01 $\pm$ 0.02	+0.54 $\pm$ 0.26
Sa-Sbc.....	58	+0.36 $\pm$ 0.51	+0.08 $\pm$ 0.15	+0.05 $\pm$ 0.22	-0.01 $\pm$ 0.03	+0.23 $\pm$ 0.41
Sc-Sd.....	58	+0.46 $\pm$ 0.37	+0.03 $\pm$ 0.18	-0.10 $\pm$ 0.28	-0.01 $\pm$ 0.04	-0.06 $\pm$ 0.20
dI.....	22	+0.07 $\pm$ 0.66	-0.07 $\pm$ 0.12	+0.46 $\pm$ 0.53	+0.10 $\pm$ 0.07	-0.16 $\pm$ 0.28
$z = 3$ , Rest-Frame $u$ Observations						
E/S0.....	3	+0.04 $\pm$ 0.74	+0.09 $\pm$ 0.23	+0.56 $\pm$ 0.52	-0.02 $\pm$ 0.10	+0.95 $\pm$ 0.40
Sa-Sbc.....	5	+0.50 $\pm$ 0.72	+0.11 $\pm$ 0.13	-0.10 $\pm$ 0.31	-0.02 $\pm$ 0.03	-0.08 $\pm$ 0.49
Sc-Sd.....	8	+0.47 $\pm$ 0.58	-0.02 $\pm$ 0.13	-0.34 $\pm$ 0.40	+0.02 $\pm$ 0.07	-0.29 $\pm$ 0.30

The local galaxies images were first deconvolved in the standard way in IDL: we divide the Fourier transform of the image by the Fourier transform of the PSF and compute the inverse Fourier transform of the result. Next they were rebinned to the pixel scale of galaxies observed at  $z = 2$  (670 pc pixel<sup>-1</sup>) or  $z = 3$  (616 pc pixel<sup>-1</sup>) and convolved with the NIC3 PSF (FWHM = 0".22 = 2.75 pixels). The galaxy fluxes were scaled to the count rate for an  $M^*$  galaxy at  $z = 2$  or 3 observed by NICMOS in F110W + F160W. Finally, a blank region of NICMOS HDFN combined F110W + F160W image was added to the redshifted galaxy images to simulate the effects of sky noise. (Note that we do not conserve the luminosities of the local galaxy sample. Many local galaxies would not be visible at  $z = 2-3$  in the rest-frame  $u$  or  $B$ , and Lyman-break galaxies are typically 2 mag brighter than local galaxies. Our simulations in § 3 suggest that at  $\langle S/N \rangle > 2$  spatial resolution will dominate any morphological biases.)

We find that the poor spatial resolution of  $z = 2-3$  galaxies is expected to significantly bias their observed morphologies. In Table 8, we give the simulated mean biases in  $C$ ,  $A$ ,  $S$ ,  $G$ , and  $M_{20}$  for early-, mid-, and late-type galaxies at  $z = 2$  and 3 observed at  $\sim 1.3 \mu\text{m}$ . A scatter of  $\sim 0.13$  is introduced to the  $A$  measurements, making it ineffective at distinguishing between early and late-type galaxies.  $S$  also has large uncertainties at these resolutions and large biases for E/S0's as a result of their unresolved centers.  $C$  and  $M_{20}$  are also significantly biased as a function of morphological type but have a greater dynamical range and therefore are still useful.  $G$  remains a reliable unbiased diagnostic out to at least  $z \sim 3$  for the NICMOS HDFN plate scale and PSF (Table 8).

Given these biases, some of the observed LBG morphologies appear to be similar to the morphologies of local early-type galaxies (Figs. 16–17). However, some of the  $z \sim 2-3$  galaxies have higher Gini coefficients and/or asymmetries than expected from the degraded local galaxy images (Fig. 16), and one  $z \sim 3$  object has a double nucleus, resulting in a much higher asymmetry and  $M_{20}$  than any of the degraded local galaxy images. We have applied a series of two-dimensional K-S tests to the LBG sample and degraded local galaxy simulations, similar to the ones used in § 4 (Table 9). We find that the  $z \sim 2$  LBG sample has a less than 0.4% probability of matching the degraded  $B/g$ -band local galaxy morphologies for all combinations of  $C$ - $A$ - $G$ - $M_{20}$ , except for  $C$ - $M_{20}$ , where systematic biases are the strongest. The  $z \sim 3$  LBGs are more likely to be drawn from a populations of galaxies with “normal” morphologies ( $>2\%$  probability); however, fewer galaxies are observed in the  $z \sim 3$  and the  $u$ -band local galaxy

samples, and one galaxy is highly asymmetric. Therefore, it is highly unlikely that the  $z \sim 2$  galaxies have morphologies identical to local elliptical/S0 or spiral galaxies; rather their high  $G$  and moderate  $A$  values suggest that they are more like the ULIRG population (Figs. 10–12).

## 6. SUMMARY

We have redefined the Gini coefficient in equation (6) as a statistic for measuring the distribution of flux values within a galaxy's image and introduced  $M_{20}$  (eq. [8]), the second-order moment of the brightest 20% of the galaxy's flux. These two indices are complementary, nonparametric morphology measures. We have tested robustness of  $G$  and  $M_{20}$  to decreasing S/N and resolution and found them to change by less than 10% for average S/N per pixel  $\geq 2$  and resolutions better than 500 pc. At worse resolutions  $C$ ,  $A$ , and  $M_{20}$  have systematic biases that are a function of Hubble type, while  $S$  becomes unreliable.  $G$ , on the other hand, appears to be remarkably stable at low resolutions and therefore is a powerful tool for classifying the morphologies of high-redshift galaxies.

We have measured  $C$ ,  $A$ ,  $S$ ,  $G$ , and  $M_{20}$  from the near-UV/optical images of 170 local E–S0–Sa–Sb–Sc–Sd–dI galaxies, 73  $z \sim 0.2$  ULIRGs, and 49  $1.7 < z < 3.8$  Lyman-break galaxies. We find that:

1. Normal Hubble-type galaxies follow a tight  $G$ - $M_{20}$ - $C$  sequence. Early-type and bulge-dominated systems have high Gini coefficients and concentrations and low second-order moments as a result of their bright and compact bulges. Shallower surface brightness profiles, spiral arms, and off-center star clusters give late-type disks lower Gini coefficients and concentrations and higher second-order moments.

2. In combination with  $A$  and  $S$ ,  $G$  is more effective than  $C$  at distinguishing ULIRGs from normal Hubble types. We also find that most ULIRGs lie above the  $G$ - $M_{20}$  sequence and can be identified by their higher  $G$  and  $M_{20}$  values. The high Gini coefficients arise from very bright compact nuclei, while multiple nuclei and bright tidal features produce large second-order moments.

3. ULIRGs with double and multiple nuclei have a statistically different distribution in morphology space than single-nuclei ULIRGs. ULIRGs with double/multiple nuclei typically have higher second-order moments and asymmetries and slightly lower concentrations than single nuclei ULIRGs. Singly nucleated ULIRGs are more likely to possess low asymmetries and low second-order moments and often have higher

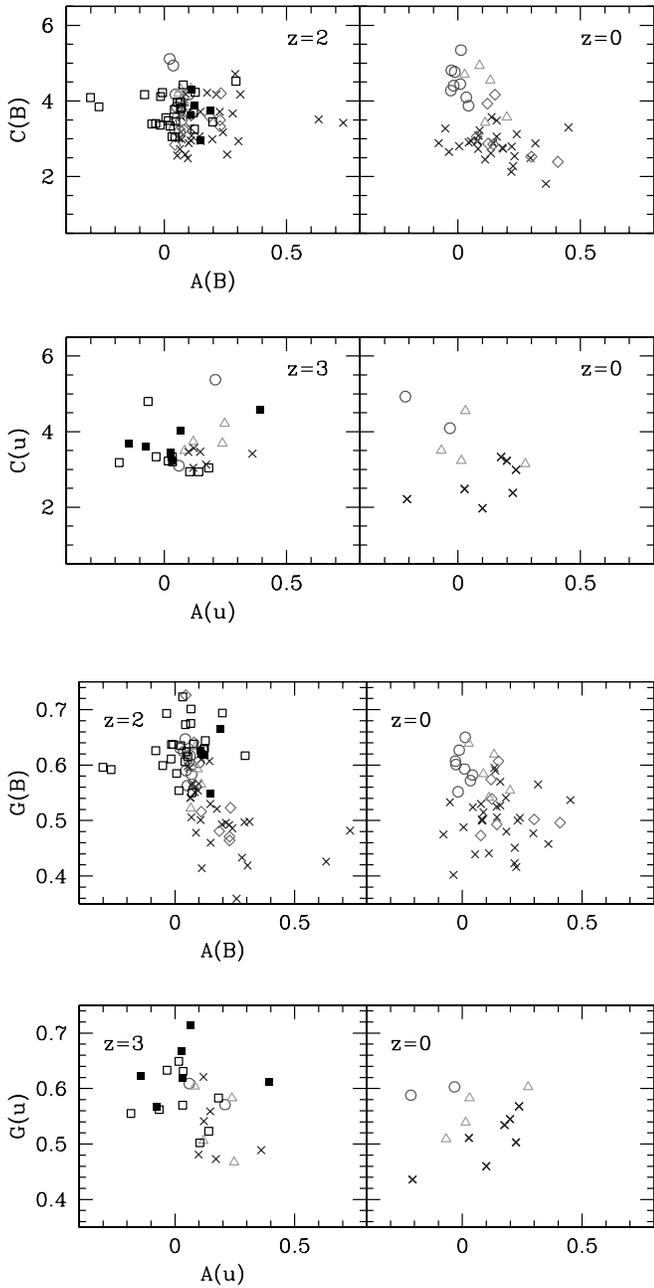


FIG. 16.— $A$  vs.  $C$  and  $A$  vs.  $G$  for the HDF-N Lyman-break galaxies (*open squares*: LBGs with spectroscopic redshifts; *filled squares*: LBGs with photometric redshifts). The right-hand panels show observed morphologies of normal local galaxies (*circles*: E/S0, *triangles*: Sa–Sbc, *crosses*: Sc–Sd, *diamonds*: dl). The left-hand panels show the observed LBG morphologies and the morphologies of local galaxies expected for  $z = 2$ – $3$  galaxies at the NICMOS HDF-N image resolution.

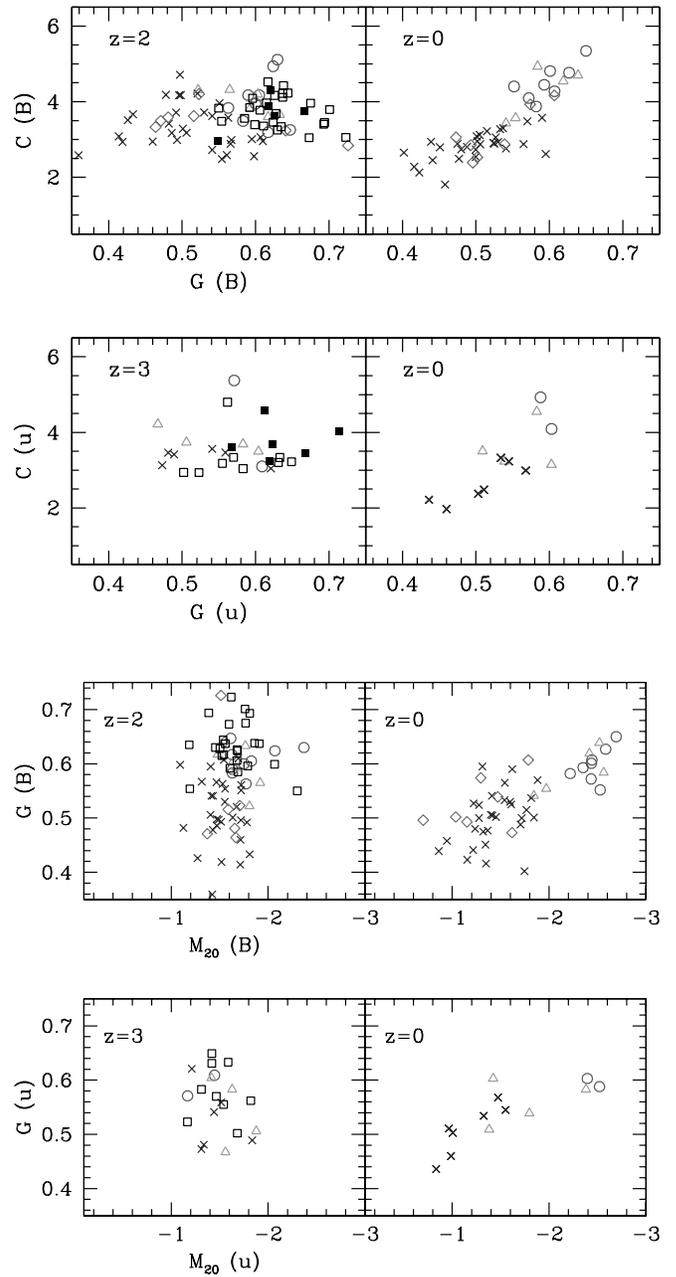


FIG. 17.— $G$  vs.  $C$  and  $M_{20}$  vs.  $G$  for the HDF-N Lyman-break galaxies. Symbols are same as Fig. 16.

TABLE 9  
LYMAN-BREAK GALAXY K-S TEST PROBABILITIES

Parameter	$z = 2$ LBGs versus Local Galaxies			$z = 3$ LBGs versus Local Galaxies		
	$A$	$G$	$M_{20}$	$A$	$G$	$M_{20}$
$C$ .....	5.1e-4	4.6e-6	0.075	0.065	0.032	0.27
$A$ .....	...	2.1e-8	0.0039	...	0.017	0.13
$G$ .....	...	...	6.6e-6	...	...	0.10

NOTE.—Two-dimensional K-S test probabilities that two galaxy populations have the same distribution in a two-parameter space.

concentrations and Gini coefficients than ULIRGs in well-separated galaxy pairs.

4. Many of HDFN galaxies at  $z \sim 2$  have higher rest-frame  $B$ -band Gini coefficients and asymmetries than expected for local elliptical and spiral galaxies degraded to the same resolution. Instead, these objects are most similar in morphology to local ULIRGs.

Our revised Gini coefficient has proved to be a highly robust and unbiased nonparametric morphological indicator for  $z > 2$  galaxies observed at *HST* NICMOS resolution and therefore has opened a window into the morphologies and assembly of the earliest galaxies. At lower redshifts, and in combination with  $M_{20}$ ,  $A$ , and  $S$ , the Gini coefficient allows us to more precisely classify galaxy morphologies and identify merger candidates. In our next paper we analyze a suite of hydrodynamical galaxy merger simulations to predict the evolution of merging galaxies in  $G$ - $M$ - $C$ - $A$ - $S$  morphology space. These simulations will explore a range of merger mass ratios, orbital parameters, and star formation feedback efficiencies and will trace the spatial distribution of dark matter, gas, and old and new stars as a function of time (Cox et al. 2004).

We would like to thank T. J. Cox and P. Jonsson for their valuable input and careful reading of this manuscript. We also

gratefully acknowledge C. Conselice for his comments and access to his morphology analysis code, M. Dickinson for use of the NICMOS Hubble Deep Field North observations, and the anonymous referee for their useful suggestions. Support for J. L. was provided by NASA through grant 9515 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. J. P. acknowledges support from NSF through grant AST 02-05944 and NASA through NAG5-12326. P. M. acknowledges support by NASA through grants NAG5-11513 and GO-09425.19A from STScI. This work is based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained from the data archive at STScI, and observations from the Sloan Digital Sky Survey. Funding for the SDSS has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institut für Astronomie, the Max-Planck-Institut für Astrophysik, New Mexico State University, University of Pittsburgh, Princeton University, the US Naval Observatory, and the University of Washington.

#### REFERENCES

- Abazajian, K., et al. 2003, *AJ*, 126, 2081  
 Abraham, R., van den Bergh, S., & Nair, P. 2003, *ApJ*, 588, 218  
 Abraham, R. G., Tanvir, N. R., Santiago, B. X., Ellis, R. S., Glazebrook, K., & van den Bergh, S. 1996, *MNRAS*, 279, L47  
 Abraham, R. G., Valdes, F., Yee, H.K.C., & van den Bergh, S. 1994, *ApJ*, 432, 75  
 Abraham, R. G., & van den Bergh, S. 2001, *Science*, 293, 1273  
 Balcells, M., Graham, A., Dominguez-Palmero, L., & Peletier, R. 2003, *ApJ*, 582, L79  
 Bershady, M., Jangren, A., & Conselice, C. 2000, *AJ*, 119, 2645  
 Blanton, M., et al. 2003a, *ApJ*, 594, 186  
 ———. 2003b, *ApJ*, 592, 819  
 Borne, K. D., Bushouse, H., Lucas, R. A., & Colina, L. 2000, *ApJ*, 529, L77  
 Brinchmann, J., et al. 1998, *ApJ*, 499, 112  
 Budavari, T., Szalay, A., Connolly, A., Csabai, I., & Dickinson, M. 2000, *AJ*, 120, 1588  
 Carollo, M. 1999, *ApJ*, 523, 566  
 Conselice, C. 2003, *ApJS*, 147, 1  
 Conselice, C., Bershady, M., Dickinson, M., & Papovich, C. 2003, *ApJ*, 126, L1183  
 Conselice, C., Bershady, M., & Jangren, A. 2000, *ApJ*, 529, 886  
 Cox, T. J., et al. 2004, in preparation  
 Cui, J., Xia, X.-Y., Deng, Z.-G., Mao, G., & Zou, Z.-L. 2001, *AJ*, 122, 63  
 Dickinson, M. 1999, in *After the Dark Ages: When Galaxies were Young* (The Universe at  $2 < z < 5$ ), ed. S. Holt & E. Smith (College Park, MD: AIP Press), 122  
 Driver, S. P., Windhorst, R. A., & Griffiths, R. 1995, *ApJ*, 453, 48  
 Drozdovsky, I. O., & Karachentsev, I. D. 2000, *A&AS*, 142, 425  
 Eastman, R. G., Schmidt, B. P., & Kirshner, R. 1996, *ApJ*, 466, 911  
 Fasano, G., & Franceschini, A. 1987, *MNRAS*, 225, 155  
 Ferguson, H. C., et al. 2004, *ApJ*, 600, L107  
 Freedman, W. L., et al. 2001, *ApJ*, 553, 47  
 Frei, Z., Guhathakurta, P., Gunn, J., & Tyson, J. A. 1996, *AJ*, 111, 174  
 Garcia, A. M., Fournier, A., di Nella, H., & Paturel, G. 1996, *A&A*, 310, 412  
 Gavazzi, G., Boselli, A., Scodreggio, M., Pierini, D., & Belsole, E. 1999, *MNRAS*, 304, 595  
 Gialalisco, M., Steidel, C. C., & Macchetto, F. D. 1996, *ApJ*, 470, 189  
 Glasser, G. J. 1962, *J. Am. Stat. Assoc.*, 57, 648  
 Isserstedt, J., & Schindler, R. 1986, *A&A*, 167, 11  
 Jensen, J. B., et al. 2003, *ApJ*, 583, 712  
 Kelly, B. C., & McKay, T. A. 2004, *AJ*, 127, 625  
 Karachentsev, I., & Drozdovsky, I. O. 1998, *A&AS*, 131, 1  
 Karachentsev, I., Musella, I., & Grimaldi, A. 1996, *A&A*, 310, 722  
 Karachentsev, I., et al. 2002, *A&A*, 383, 125  
 Kuchinski, L. E., Madore, B. F., Freedman, W. L., & Trewhela, M. 2001, *ApJ*, 122, L729  
 Lorenz, M. O. 1905, *J. Am. Stat. Assoc.*, 9, 209  
 Naim, A., Ratnatunga, K., & Griffiths, R. 1997, *ApJ*, 476, 510  
 Odewahn, S. C., Windhorst, R. A., Driver, S. P., & Keel, W. 1996, *ApJ*, 472, L13  
 Papovich, C. 2002, Ph.D. thesis, Johns Hopkins Univ.  
 Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, 559, 620  
 Parodi, B. R., Saha, A., Sandage, A., & Tammann, G. A. 2000, *ApJ*, 540, 634  
 Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, 124, 266  
 Petrosian, V. 1976, *ApJ*, 209, L1  
 Pierce, M. J. 1994, *ApJ*, 430, 53

- Pierce, M. J., McClure, R. D., & Racine, R. 1992, *ApJ*, 393, 523  
Pierce, M. J., & Tully, R. B. 1988, *ApJ*, 330, 579  
Refregier, A. 2003, *MNRAS*, 338, 35  
Russell, D. G. 2002, *ApJ*, 565, 681  
Sandage, A., & Bedke, J. 1994, *The Carnegie Atlas of Galaxies, Volume 1* (Washington: Carnegie Inst.)  
Schade, D., Lilly, S. J., Crampton, D., Hammer, F., LeFevre, O., & Tresse, L. 1995, *ApJ*, 451, L1  
Schoniger, F., & Sofue, Y. 1994, *A&A*, 283, 21  
Shanks, T. 1997, *MNRAS*, 290, L77  
Shapley, A., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, *ApJ*, 562, 95  
Simard, L., et al. 2002, *ApJS*, 142, 1  
Stoughton, C., et al. 2002, *AJ*, 123, 485  
Takamiya, M. 1999, *ApJS*, 122, 109  
Theureau, G., Hanski, M., Ekholm, T., Bottinelli, L., Gouguenheim, L., Paturel, G., & Teerikorpi, P. 1997, *A&A*, 322, 730  
Van Zee, L. 2001, *AJ*, 121, 2003  
Wu, H., Zou, Z. L., Xia, X. Y., & Deng, Z. G. 1998, *A&AS*, 132, 181  
Wu, K. 1999, Ph.D. thesis, Univ. California, Santa Cruz  
Yasuda, N., Fukugita, M., & Okamura, S. 1997, *ApJS*, 108, 417