An Old Stellar Population or Diffuse Nebular Continuum Emission Discovered in Green Pea Galaxies

Leonardo Clarke¹, Claudia Scarlata¹, Vihang Mehta¹, William C. Keel², Carolin Cardamone³, Matthew Hayes⁴, Nico Adams⁵, Hugh Dickinson⁵, Lucy Fortson⁶, Sandor Kruk⁶, Chris Lintott⁷, and Brooke Simmons⁸

¹ Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA
² Department of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487-0324, USA
³ Stockholm University, Department of Astronomy and Oskar Klein Centre for Cosmoparticle Physics, AlbaNova University Centre, SE-10691, Stockholm, Sweden
⁴ Tufts University, 108 Bromfield Road, Somerville, MA 02144, USA
⁵ School of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK
⁶ European Space Agency (ESA), European Space Research and Technology Centre (ESTEC), Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
⁷ Oxford Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
⁸ Department of Physics, Lancaster University, Bailrigg, Lancaster LA1 4YB, UK

Received 2020 December 12; revised 2021 April 12; accepted 2021 April 13; published 2021 May 6

Abstract

We use new Hubble Space Telescope (HST) images of nine Green Pea galaxies (GPGs) to study their resolved structure and color. The choice of filters, F555W and F850LP, together with the redshift of the galaxies (z ~ 0.25), minimizes the contribution of the nebular [O III] and Hα emission lines to the broadband images. While these galaxies are typically very blue in color, our analysis reveals that it is only the dominant stellar clusters that are blue. Each GPG does clearly show the presence of at least one bright and compact star-forming region, but these are invariably superimposed on a more extended and lower surface brightness emission. Moreover, the colors of the star-forming regions are on average bluer than those of the diffuse emission, reaching up to 0.6 magnitudes bluer. Assuming that the diffuse and compact components have constant and single-burst star formation histories, respectively, the observed colors imply that the diffuse components (possibly the host galaxy of the star formation episode) have, on average, old stellar ages (>1 Gyr), while the star clusters are younger than 500 Myr. While a redder stellar component is perhaps the most plausible explanation for these results, the limitations of our current data set lead us to examine possible alternative mechanisms, particularly recombination emission processes, which are unusually prominent in systems with such strong line emission. With the available data, however, it is not possible to distinguish between these two interpretations. A substantial presence of old stars would indicate that the mechanisms allowing large escape fractions in these local galaxies may be different from those at play during the reionization epoch.

Unified Astronomy Thesaurus concepts: Compact dwarf galaxies (281)

1. Introduction

Green Pea galaxies (GPGs) are low-redshift (z) objects first discovered by citizen scientists in the GalaxyZoo project (Lintott et al. 2008; Cardamone et al. 2009). These galaxies turned out to be much more than just an “interesting curiosity”. They have low stellar masses (≈ 10^9 M_☉), high specific star formation rates (≈ 10^-8 yr^-1), small sizes (≈ 1 kpc, completely unresolved at ground-based resolution), and their optical spectra are characterized by extreme equivalent widths (EWs) in the [O III]λ5007 and Hα emission lines (e.g., Yang et al. 2017b; Brunker et al. 2020).

These galaxies have been the focus of intense research activity since their discovery, as they are thought to be among the best-known analogs of high-redshift galaxies and those that were responsible for the reionization of the Universe at z > 6. Specifically, their compact ultraviolet (UV) morphologies, low stellar masses, low metallicities, high specific star formation rates, and high ionizing photon production rate (Cardamone et al. 2009; Amorín et al. 2010; Izotov et al. 2016a) are very similar to those of typical star-forming galaxies at z > 6 (Schaerer et al. 2016). GPGs are the only galaxy population known to have a high escape fraction of hydrogen-ionizing radiation, from direct measurement of the stellar continuum below 912 Å, with values ranging between 2% and 76% (Izotov et al. 2016a, 2016b; Faisst et al. 2016; Izotov et al. 2018a, 2018b; Bassett et al. 2019). The physical conditions that allow these high escape fractions are still poorly understood. One possibility is that these galaxies are overall deficient in neutral hydrogen (e.g., Henry et al. 2015; Eggen et al. 2021), although it is also possible that the ionizing radiation escapes along lines of sight where most of the hydrogen is ionized (Zackrisson et al. 2013; Jaskot & Oey 2014).

Little is known about the star formation history (SFH) of GPGs, and whether or not they host old stellar populations (e.g., Amorín et al. 2012). The debate about the SFH of dwarf starburst galaxies is an old one, with I Zw 18 as the prototypical example (Izotov et al. 1997; Aloisi et al. 2007; Papaderos & Östlin 2012). This galaxy is characterized by a very compact and young starburst, although the detection of red giant branch stars suggests that an older (≈ 1 Gyr) stellar component also exists. Should the GP galaxies be similar to I Zw 18, then the mechanisms that drive their high escape fraction may be different than those in the young galaxies during the reionization epoch.

A limiting factor to detailed studies of their stellar populations is that GPGs’s global light is dominated by the...
blinding emission of young stars in the UV and by the intense emission lines in broadband optical filters. Additionally, most of the results come from unresolved data from ground-based telescopes (where most observations are seeing-limited). Space-based observations have so far been mostly limited to the UV spectral range, using the imaging capabilities of the Cosmic Origin Spectrograph (COS). These studies reveal a UV morphology characterized by bright star-forming regions in the center of the GPGs, and possibly the presence of exponential disks with scale lengths between 0.6 and 1.4 kpc (e.g., Izotov et al. 2016a). However, these studies are hampered by the limited unvignetted portion of the COS aperture (≤0.5 radius), and by the fact that they mostly trace the spatial distribution of young stars.

In this study, we analyze new images of nine GPGs taken with the Advanced Camera for Surveys (ACS) on board the Hubble Space Telescope (HST). The galaxies were chosen to be at \( z \sim 0.25 \) — at this redshift the F555W and F850LP filters exclude the strongest emission lines in typical Green Pea (GP) spectra (e.g., [O III] and H\( \alpha \), see Figure 1). The high-resolution optical imaging allows for a detailed look at the stellar morphology of these galaxies, without contamination from nebular emission that can be substantial in these objects (e.g., Guseva et al. 2017). In Section 2, we present the data and the analysis. In Section 3, we present our results. Section 4 discusses the color maps and possible physical interpretations of these data. Throughout this Letter, magnitudes are expressed in the AB magnitude system (Oke 1990). At \( z \sim 0.25 \) the physical scale is 4 kpc/′, assuming a Planck 2015 cosmology (Ade et al. 2016).

2. Observations and Analysis

The targets of this study were selected from the Sloan Digital Sky Survey (SDSS) spectroscopic catalog (Data Release 12) using the same color criteria described in Cardamone et al. (2009) and limited to star-forming galaxies using the Kauffmann et al. (2003) classification, with the additional redshift constraint of \( z = 0.250 \pm 0.006 \) to ensure that [O II], [O III], and H\( \alpha \) are outside the chosen filters. From the resulting sample, we removed galaxies with bright nearby stars that could compromise the accuracy of the photometry.

The galaxies were observed with the HST-ACS camera as part of the “Gems of the Galaxy Zoo” (Zoo Gems) program (PID: 15445, PI: Keel). For each galaxy, two 337 s images were taken, one each in the F555W and F850LP filters. At redshift \( z = 0.250 \), the central wavelength of the F555W and F850LP filters correspond to rest frame wavelengths of 4500 Å and 7200 Å, respectively. In what follows, we will refer to the F555W and F850LP filters as the B- and I-band, respectively. The images were bias-subtracted, dark-corrected, and flat-fielded using calacs, version 10.2.1. This new version includes a pixel-based correction for charge transfer efficiency losses that mitigates the amplification of readnoise (Anderson & Ryon 2018).

Cosmic ray rejection was performed using a Laplacian edge detection algorithm developed by Van Dokkum (2001). After the cosmic ray removal, we performed sky subtraction on all of the images. The sky value was calculated for each image using the Astrodizzle9 tool.

The images were then photometrically calibrated to AB magnitudes (Oke 1990) and corrected for foreground Galactic extinction at the position of each target using the Schlafly & Finkbeiner (2011) recalibration of the Schlegel et al. (1998) extinction map. We used the NASA/IPAC Extragalactic Database extinction calculator10 to compute the Galactic extinction in the HST filter bandpasses.

Colors were computed after the homogenization of the point-spread function (PSF) across the two bands. We chose to convolve the F555W images to match the broader PSF of the F850LP band. First, we generated model PSFs at the position of each target in the detector for each ACS filter using the HST Tiny Tim software (Krist et al. 2011). We then generated the F555W-to-F850LP PSF homogenization kernel using Pypher11 (Boucaud et al. 2016). Color maps were then created using the convolved images.

3. Results

In Figure 2 we present the individual images of the nine GPGs as well as the B – I color maps. All GPGs are clearly resolved and show a range of morphologies. In most galaxies we identify multiple stellar clusters, likely star-forming regions responsible for the characteristic emission line spectra, in addition to a diffuse and extended continuum. In a few objects, e.g., J1004+2017 and J1020+2937, we see a single prominent star cluster that lies in the galaxy center, giving rise to a relatively symmetric morphology, while in J1336+6255 the cluster is off-centered with respect to the diffuse continuum.

We automatically detected the stellar clusters within the GPGs using a contrast-based image analysis algorithm adopted from Guo et al. (2015) with some modifications (Meha et al. 2020). Briefly, the B and I stamps were smoothed with a 8px (0.4′) boxcar kernel, and a contrast image for each filter was generated by subtracting the smoothed image from the original. The contrast image was then filtered to mask out all pixels below 2\( \sigma \) (estimated in blank sky regions) to create a filtered image. Finally, the clusters were detected on the B+I multi-band filtered images using SExtractor (Bertin & Arnouts 1996). Only those stellar clusters within the SExtractor segmentation map generated from the smoothed I image of the galaxy were retained for further analysis. The positions of the resulting clusters are marked on the images in Figure 2. Galaxies J0353,
J1015, and J1214 host multiple star clusters, while J3336, J1633, and J1004 are examples of objects with only one well-defined cluster located either off-center with respect to the underlying continuum (e.g., J3336) or close to the galaxy center (e.g., J1004).

The $B-I$ color maps shown in Figure 2 reveal that the stellar clusters have colors significantly bluer than the rest of the host galaxy. In order to quantify this difference, we compute the galaxies’ and stellar clusters’ colors as follows. First, we use the contrast images to create a segmentation map of each galaxy, with stellar clusters and host pixels identified. We then compute the host-galaxy color, $(B-I)_G$, by summing the flux of all pixels not assigned to star clusters. For the stellar clusters, we provide two estimates of their colors, with and without the subtraction of the underlying host-galaxy contribution. To subtract the galaxy contribution, we compute the average host-galaxy surface brightness within the host segmentation map and assume that this value corresponds to the galaxy color at the position of the cluster. Judging by the color maps in Figure 2, this assumption appears to be justified as no strong gradients, or variations in colors, are visible within the host galaxies. The local background flux is subtracted from the cluster aperture, and the fluxes of the stellar clusters and the corresponding colors, $(B-I)_{SC}$, were then computed within circular apertures of 0′′2 radius.

We also measured the clusters’ colors without removing the underlying galaxy contribution, $(B-I)_{ObsSC}$, using the same 0′′2 aperture. The colors of the host-galaxy and the stellar clusters are reported in Table 1.

In Figure 3, we present a comparison between the $(B-I)$ colors of the stellar clusters and their corresponding host galaxies. Filled symbols indicate $(B-I)_{SC}$, while open symbols show the colors observed at the position of the stellar clusters $(B-I)_{ObsSC}$. For galaxies with more than one clump, the stellar-cluster colors are connected by a vertical line. Right away, we quantitatively confirm the impression derived from the color maps that the stellar clusters appear to be bluer than the underlying host galaxy. Some of the clusters are up to 0.6 magnitudes bluer than the underlying host. This conclusion does not change whether we use the $(B-I)_{SC}$ colors or the $(B-I)_{ObsSC}$ ones, as the galaxy subtracted stellar-cluster colors are only 0.1 magnitudes bluer than the observed colors at the cluster positions.

4. Discussion

The color differences between the stellar clusters and the host galaxy presented in the previous section can be due to a combination of effects, such as stellar age, metallicity, SFH as well as dust extinction. With only two pass-bands, however, disentangling these effects of stellar population properties and/or nebular continuum is not possible. Keeping this caveat in mind, however, we can discuss the implications of the relative differences in the observed colors, under some empirically motivated assumptions.

Existing constraints on the stellar population of GPGs come from the analysis of spatially unresolved data, either ground-based optical observations or far-UV spectra obtained with the
We list each galaxy component. On the right-hand side, we enumerate the clusters corresponding to each host galaxy along with their characteristics. The HST-COS spectrograph. These data show that the GPGs are characterized by very young stellar ages of only a few million years, low stellar and nebular metallicities, and low dust content (e.g., Guseva et al. 2020). These are, however, luminosity-weighted results, and the derived physical properties are biased by the emission of the brightest youngest stars (e.g., Chisholm et al. 2019). It is reasonable to assume that the SFH of these clusters is well approximated by a single burst caught at a very young age, given their compact morphology and high EW emission lines. For the underlying galaxy, we assume that stars have formed over a more prolonged period of time. We therefore interpret the color differences using two SFHs, which we chose to likely bracket the range of expected colors: constant star formation (CSF) rate and simple stellar population (SSP).

In Figure 4 we compare the observed \((B-I)\) colors with the colors of synthetic stellar population synthesis models. The top panel of Figure 4 shows the distribution of colors for the stellar clusters (blue histogram) and galaxies (red histogram). The blue histogram includes a larger number of objects than the red histogram because some galaxies host multiple clumps. For each SFH, the theoretical colors are computed for a range of ages (between a few Myr and the age of the Universe) and two metallicities (20% and 40% solar). Subsolar metallicities in GPGs are implied both by the analysis of the nebular lines in the optical spectra (e.g., Izotov et al. 2016a; Yang et al. 2017a) and by the modeling of the far-UV (FUV) spectra observed with COS. The model templates were generated using Flexible Stellar Population Synthesis (FSPS) with a Chabrier (2003) initial mass function and the default MILES stellar library with MIST isochrones (Conroy et al. 2009; Conroy & Gunn 2010). We show only dust-free models. The effect of an absorption of \(A_V=1\) would be to redder the \((B-I)\) colors by \(\approx 0.5\) magnitudes. We show calculations with and without the nebular contribution (from both emission lines and continuum) as solid and dashed lines, respectively. Although the emission lines fall

### Table 1
Characteristics of the Nine Target GPGs along with Their Associated Stellar Clusters

| Target Name       | Coordinates (J2000) | \(z\) | \((B-I)_G\) (mag) | \(\mu\) (mag/"a") | \(\log \left( \frac{M_d}{M_{\text{tot}}} \right)\) | \(\frac{M_d}{M_{\text{tot}}}\) | Clump ID | \((B-I)_{\text{Obs.SC}}\) (mag) | \((B-I)_{\text{NC}}\) (mag) |
|-------------------|---------------------|------|------------------|------------------|----------------------------------|------------------|----------|-------------------------------|------------------|
| J0353–0010        | 03:53:32.4636-00:10:28.88 | 0.246 | 0.06 ± 0.02 | 21.13 ± 0.01 | 9.13 ±0.09 \(\pm 0.25\) | 0.82 | C1 | −0.15 ± 0.03 | −0.2 ± 0.03 |
| J0004+2017        | 10:04:00.6406+20:17:19.25 | 0.255 | 0.25 ± 0.01 21.43 ± 0.01 | 9.41 ±0.06 \(\pm 0.03\) | 0.81 | C1 | 0.01 ± 0.01 | −0.01 ± 0.01 |
| J1015+2227        | 10:15:41.1521+22:27:27.52 | 0.243 | 0.16 ± 0.02 | 21.74 ± 0.01 | 8.77 ±0.08 \(\pm 0.09\) | 0.71 | C1 | −0.1 ± 0.02 | −0.15 ± 0.03 |
| J1020+2937        | 10:20:57.4622+29:37:26.47 | 0.256 | 0.21 ± 0.01 | 21.55 ± 0.01 | 9.25 ±0.06 \(\pm 0.07\) | 0.84 | C1 | −0.05 ± 0.02 | −0.08 ± 0.02 |
| J1055+0841        | 10:55:30.4166+08:41:32.9 | 0.252 | 0.17 ± 0.02 | 21.67 ± 0.01 | 8.91 ±0.08 \(\pm 0.06\) | 0.76 | C1 | −0.04 ± 0.02 | −0.07 ± 0.02 |
| J1214+4520        | 12:14:23.1802+45:20:40.91 | 0.255 | 0.26 ± 0.02 | 21.70 ± 0.01 | 9.04 ±0.06 \(\pm 0.07\) | 0.82 | C1 | 0.14 ± 0.02 | 0.1 ± 0.03 |
| J1336+6255        | 13:36:07.9138+62:55:30.77 | 0.252 | 0.45 ± 0.02 | 21.70 ± 0.01 | 9.48 ±0.06 \(\pm 0.05\) | 0.97 | C1 | −0.02 ± 0.03 | −0.09 ± 0.03 |
| J1504+5954        | 15:04:57.9874+59:54:07.27 | 0.250 | 0.25 ± 0.01 | 21.35 ± 0.01 | 9.32 ±0.08 \(\pm 0.08\) | 0.74 | C1 | 0.15 ± 0.01 | 0.14 ± 0.01 |
| J1633+3753        | 16:33:37.9414+37:53:14.3 | 0.252 | 0.2 ± 0.02 | 21.76 ± 0.01 | 8.62 ±0.09 \(\pm 0.07\) | 0.59 | C1 | 0.17 ± 0.02 | 0.15 ± 0.02 |

**Note.** We list each galaxy’s name followed by its coordinates, redshift, \((B-I)\) color, \(\mu\), mass of the diffuse component \((M_d)\), and fraction of the total mass in the diffuse component. On the right-hand side, we enumerate the clusters corresponding to each host galaxy along with their \((B-I)\) colors calculated in the two ways described in the text.

**Figure 3.** \((B-I)\) color comparison between each galaxy and the clumps corresponding to each galaxy. Note that points connected by vertical lines correspond to galaxies that host multiple clumps. Solid points show the colors corrected for the underlying galaxy contribution, while open symbols show the measurement of the observed colors at the position of the clump (see the text for details).
mostly outside the filters, the nebular continuum would still contribute substantially to the observed colors, particularly for younger ages. As Figure 4 shows, for the CSF rate model (red curves), the nebular emission is important for ages up to one billion years and dominates for ages below a few hundred million years.

Comparing the host-galaxy and stellar cluster colors with CSF rate and SSP models, respectively, we can draw the following conclusions. Neglecting the reddening due to dust extinction, the colors of the host galaxies are well reproduced by stellar populations that have formed continuously over at least the past billion year, up to 10 billion years, or by a burst that occurred more than 100 Myr in the past. This conclusion does not substantially depend on the metallicity of the stars or the contribution of the nebular continuum that, at these ages, is subdominant compared to the stellar continuum. The bluer galaxies, however, can also be reproduced with young CSF rate models (younger than 10 Myr) if the nebular continuum is important. The colors of the compact star-forming regions, on the other hand, suggest younger stellar ages than the underlying galaxies. Specifically, with the single-burst assumption the observed colors can be explained with ages between one and 10 Myr, although ages of a few hundred million years cannot be ruled out. The younger estimates are more in agreement with the previous studies of GPGs mentioned above.

The stellar mass hiding behind the blinding stellar clusters is possibly not negligible. For the constant SFH model, we find that the stellar mass in the diffuse red component ($M_d$) comprises between 60% and almost 100% of the total stellar mass in the GPGs (see Table 1). We are conscious, however, of the large uncertainties in the stellar population properties implied by the limited information available.

Finally, we consider the alternative possibility that there is only one stellar component, but the young stars and the ionized gas have different spatial distribution (as observed, e.g., in I Zw 18 by Papaderos & Östlin 2012). In this scenario, the observed color maps would be the result of a varying relative contribution of the nebular spectrum to the total color in different places of the ionized nebula.

With our choice of filters, the nebular emission (line +continuum) has a red $B-I$ color, of $\approx 0.76$, and would only get redder in the presence of dust. If this interpretation is correct, the blue emission at the position of the stellar clusters would be dominated by stellar continuum, with minimal contribution from the ionized gas. This would be possible if, for example, we were looking at the stellar clusters through clear channels in the gas distribution. The emission in the diffuse component of the GPGs would then result from a declining surface brightness in hot stars, and a correspondingly relatively-more-important contribution of the nebular emission. UV-resolved observations of GPGs suggest that the surface brightness of young stars declines exponentially with radius (Izotov et al. 2016b).

We proceed to calculate the expected contribution of nebular continuum over the segmented area of the image. We computed the predicted nebular continuum flux density from the observed Hα flux, assuming an Hα $EW_{neb} = 4740$ Å, appropriate for a pure nebular spectrum. With this assumption, $f_{\text{neb}} = \frac{4740}{1.1 EW_{H\alpha}}$ (the 10% correction is applied to convert the continuum from the Hα wavelength to the central wavelength of the filter, 8750 Å). We use the galaxy half-light radius measured in the I-band images ($R_{1/2}$) and compute the expected surface brightness due to the nebular continuum as $I_{\text{neb}} = 0.5 f_{\text{neb}}$. This calculation shows that for a flat surface brightness profile in the nebular gas, the nebular continuum reaches a comparable surface brightness to the diffuse I-band emission ($i_{\mu}$ in Table 1) that is sufficient in all but one galaxy (where it accounts for about 85% of the emission), and gives $\approx$four times as many photons as needed in the highest extreme. More realistic nebular surface brightness profiles will decline with radius, leading to lower contributions.

5. Conclusion

We analyzed HST images of nine compact, star-forming GPGs at $z \sim 0.25$. The choice of broadband filters avoids the prominent emission lines from these galaxies and allows us to measure resolved colors for the first time. In all galaxies, we find one or more dense bright stellar clusters (typically located in the central regions of the galaxies) superimposed on a more diffuse component. The bright clumps are associated with blue colors, systematically bluer than the diffuse component.

Interpretating the color differences depends on assumptions of the star formation history in each component, a parameter that is poorly constrained in this class of galaxies. Assuming that the diffuse and compact components have constant and single-burst SFHs, respectively, the observed colors imply that the diffuse components (possibly the host galaxy of the star formation episode) have, on average, old stellar ages ($\gtrsim 1$ Gyr), while the star clusters are younger than 500 Myr. We consider the possibility that the diffuse red emission is due to a varying relative contribution of nebular emission. Although the latter

Figure 4. Top panel: histogram of galaxy and clump colors. Bottom panel: stellar age versus $B-I$ color for constant star formation rate (red) and simple stellar population (blue) star formation histories of varying metallicities. Solid and dashed lines represent calculations with and without the contribution of nebular emission (see the text for details).
explanation is unlikely, it is not possible to distinguish between these two interpretations with the available data.

GPGs are the focus of a large investment of resources because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape in large fraction. As such, these objects are often referred to as the best local analogs of the sources that reionized the Universe in large fraction. As such, these objects are often referred to as the best local analogs of the sources that reionized the Universe. These objects are often referred to as the best local analogs of the sources that reionized the Universe. Because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape. These objects are often referred to as the best local analogs of the sources that reionized the Universe. As such, these objects are often referred to as the best local analogs of the sources that reionized the Universe. These objects are often referred to as the best local analogs of the sources that reionized the Universe. Because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape. These objects are often referred to as the best local analogs of the sources that reionized the Universe. As such, these objects are often referred to as the best local analogs of the sources that reionized the Universe. Because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape because they are the only galaxy population for which hydrogen-ionizing radiation is consistently observed to escape.