Radio Galaxy Zoo: new giant radio galaxies in the RGZ DR1 catalogue

H. Tang [1]⋆ A. M. M. Scaife,[1,2] O. I. Wong,[3,4,5] A. D. Kapińska,[6] L. Rudnick [7] S. S. Shabala [8] N. Seymour[9] and R. P. Norris [10,11]

[1]*Jodrell Bank Centre for Astrophysics, University of Manchester, Manchester M13 9PL, UK*
[2]*The Alan Turing Institute, Euston Road, London NW1 2DB, UK*
[3]*ICRAS-M468, University of Western Australia, Crawley, WA 6009, Australia*
[4]*CSIRO Astronomy and Space Science, PO Box 1130, Bentley, WA 6102, Australia*
[5]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[6]*National Radio Observatory Astronomy (NRAO), PO Box 0, Socorro, NM 87801-0180, USA*
[7]*School of Physics and Astronomy, University of Minnesota, 116 Church St. SE, Minneapolis, MN 55455, USA*
[8]*School of Natural Sciences, Private Bag 37, University of Tasmania, Hobart, TAS 7001, Australia*
[9]*International Centre for Radio Astronomy Research, Curtin University, Perth, 6009, Australia*
[10]*CSIRO Astronomy and Space Science, Australia Telescope National Facility, PO Box 76, Epping, NSW 1710, Australia*
[11]*Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia*

ABSTRACT

In this paper, we present the identification of five previously unknown giant radio galaxies (GRGs) using Data Release 1 of the Radio Galaxy Zoo citizen science project and a selection method appropriate to the training and validation of deep learning algorithms for new radio surveys. We associate one of these new GRGs with the brightest cluster galaxy (BCG) in the galaxy cluster GMBCG J251.67741+36.45295 and use literature data to identify a further 13 previously known GRGs as BCG candidates, increasing the number of known BCG GRGs by > 60 per cent. By examining local galaxy number densities for the number of all known BCG GRGs, we suggest that the existence of this growing number implies that GRGs are able to reside in the centres of rich ($\sim 10^{14}\,M_\odot$) galaxy clusters and challenges the hypothesis that GRGs grow to such sizes only in locally underdense environments.

Key words: methods: data analysis – catalogues – radio continuum: galaxies.

1 INTRODUCTION

Giant radio galaxies (GRGs) are the largest radio galaxies in the Universe. Originally defined to be those radio galaxies with projected linear sizes greater than 1 Mpc, in a cosmology with $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Willis, Strom & and Wilson 1974), the GRG size limit is now equivalent to 700 kpc in a Λcold dark matter (ΛCDM) cosmology with the Planck 2016 parameters (Planck Collaboration XIII 2016; Dabhade et al. 2020a). It is thought that GRG sizes might be caused by high kinetic jet power (Wiita et al. 1989) and it has been shown that the size of a radio source is positively correlated with source radio luminosity and jet power (Shabala & Godfrey 2013). Alternatively, it has also been proposed that the gigantic size of GRGs might be caused by the comparative longevity of their jets (Subrahmanyan et al. 1996), or due to the radio source growing in a low-density environment (Malarecki et al. 2015).

The role of local environment in GRG formation was first considered by Ishwara-Chandra & Saikia (1999) who compared 53 GRGs in the literature with 3CR radio sources (Laing, Riley & Longair 1983) of smaller sizes. Ishwara-Chandra & Saikia (1999) found that GRGs share a marginally higher separation ratio of hotspot distances from the nucleus, which Ishwara-Chandra & Saikia (1999) suggest

might be caused by the interaction of energy carrying beams and cluster-sized density gradients far from the source host galaxy. This has in turn led to GRGs being used as probes of the low ambient density warm–hot intergalactic medium (Safouris et al. 2009; Peng, Chen & and Strom 2015).

Another environmental consideration is the local galaxy density around GRGs. GRGs have typically been found in underdense environments, and it has been proposed that such reduced galaxy densities facilitate these radio galaxies to grow larger (Malarecki et al. 2015). However, a number of other studies have found that there is no correlation between radio source linear size and local galaxy density (Komberg & Pashchenko 2009; Ortega-Minakata, Torres-Papaqui & Andernach 2013; Kuźmicz et al. 2018). Moreover, the recent discovery of more than 20 GRGs that not only reside in galaxy cluster environments as found in Seymour et al. (2020), but are also the brightest galaxy in these clusters (brightest cluster galaxies, BCGs) has also challenged this hypothesis (Dabhade et al. 2017, 2020a).

From a galaxy evolution perspective, GRGs represent the tail of the radio galaxy size distribution. A comprehensive study of the shape of this distribution requires consistent sampling of both GRGs and smaller radio galaxies. However, traditional methods of cross-matching large-scale radio surveys, like the Faint Images of the Radio Sky at Twenty-Centimeters (FIRST; Becker, White & Helfand 1995) survey, with optical/infrared surveys such as those obtained using

⋆ E-mail: hongming.tang@manchester.ac.uk

the Wide-field Infrared Survey Explorer (*WISE*; Wright et al. 2010), e.g. the AllWISE image atlas and catalogue (Cutri & et al. 2013), are complicated by scale-dependent observational selection effects, as well as the uncertainties in cross-matching which arise when dealing with diffuse or complex radio emission.

The physical size of a source can be calculated if its host galaxy redshift (z) and its largest angular size (LAS) are available. These require validated cross-identification of radio components and their host galaxy. Traditionally, a limited number of experts would first identify radio source components and then cross-match their optical/infrared hosts (e.g. Subrahmanyan et al. 1996; Machalski, Jamrozy & Zola 2001; Schoenmakers et al. 2001; Lara et al. 2001a; Saripalli et al. 2005; Machalski, Koziel-Wierzbowska & and Jamrozy 2007; Solovyov & and Verkhodanov 2011). Recently, thanks to the availability of large optical and radio surveys, Dabhade et al. (2020a) discovered 225 new GRGs using the Value Added Catalogue (VAC; Williams et al. 2019) of the LOw Frequency ARray (LOFAR; van Haarlem et al. 2013). Most compact sources in the VAC catalogue are selected by cross-matching the LOFAR Two-metre Sky Survey Data Release 1 catalogue (LoTSS DR1; Shimwell et al. 2017, 2019) with a catalogue of matches between the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS; Kaiser et al. 2002, 2010; Chambers et al. 2016) catalogue and the AllWISE catalogue, using a likelihood ratio method (Williams et al. 2019; Dabhade et al. 2020a). However, diffuse and complex sources in the catalogue are cross-matched by visual inspection using the citizen science LOFAR Galaxy Zoo project (LGZ; Williams et al. 2019). Among the 231 716 sources of LoTSS DR1 that have optical/IR identifications, only 0.1 per cent are found to be GRGs (Williams et al. 2019; Dabhade et al. 2020a).

Citizen science offers an alternative to more traditional methods of building large cross-matched radio galaxy catalogues. Radio Galaxy Zoo (RGZ; Banfield et al. 2015) is an online citizen science project that aims to cross-match extended radio sources from the FIRST survey (Becker et al. 1995) and the Australia Telescope Large Area Survey (ATLAS; Franzen et al. 2015) with their host galaxies in the infrared waveband, using data from the AllWISE survey and the SIRTF Wide-Area Infrared Extragalactic Survey (SWIRE; Lonsdale et al. 2003). RGZ offers its volunteers a 3×3 arcmin^2 cutout from the FIRST survey with radio contours starting at $3\sigma_{rms}$ on top of a *WISE* 3.4 μm image. Project participants are asked (a) to identify radio components of a source from an image, (b) to select the infrared host galaxy of the corresponding radio source, and (c) to check if there are additional sources without existing identifications present in the same image (Banfield et al. 2015). The project is intended to provide the foundation of a large cross-matched radio galaxy catalogue.

Both citizen science and traditional astronomy methods are now being used as the foundation for recent studies that employ automatic radio morphology classification using deep learning algorithms. Using a number of radio galaxy catalogues that include radio morphology classification (e.g. FRICAT, Capetti, Massaro & Baldi 2017a; FRIICAT, Capetti, Massaro & Baldi 2017b; the Combined NVSS-FIRST Galaxies (CoNFIG) sample, Gendre & Wall 2008; Gendre, Best & Wall 2010; Mingo et al. 2019), several automatic radio morphology deep learning classifiers have been developed (e.g. Aniyan & Thorat 2017; Alhassan, Taylor & Vaccari 2018; Lukic et al. 2018, 2019; Ma et al. 2019). These automatic classifiers are built to extract morphological features from input images for classification. For general radio galaxy classification (FR I/FR II; Fanaroff & Riley 1974), these applications have achieved model accuracies comparable to visual inspection. However, these deep learning algorithms require their individual image inputs either to

have a common input image size or to be resized to a common size (Lukic et al. 2019). Since GRG identification requires LAS estimation, training a deep learning based GRG classifier under these constraints would require an image training data set with image sizes large enough to enable an algorithm to estimate source LAS for very extended objects. Considering the memory limits of state-of-the-art GPUs, the image sizes required for such an algorithm in the GRG case are likely to make such a general approach highly computationally expensive. Consequently, in the case where image size is restricted due to memory limitations, as well as the potential for confusion due to multiple objects in the field, careful consideration must be given to the effects of selection bias in the use of such machine learning approaches.

In this paper, we identify five new GRGs selected from the RGZ Data Release 1 (RGZ DR1; Wong et al., in preparation). RGZ DR1 is a manually cross-matched radio galaxy catalogue, using the efforts of more than 12 000 citizen scientist volunteers (Wong et al., in preparation). Unlike previous GRG identification studies, this work uses a process compatible with the constraints imposed by current deep learning algorithms. In Section 2, we describe the initial source selection process; in Section 3, we describe the validation process for identifying GRGs; in Section 4, we draw comparisons with other GRG identification studies, including their comparative selection effects and resulting impact on deep learning algorithms. We also discuss the characteristics and environments of these newly identified GRGs in a wider context; and in Section 5, we draw our conclusions.

In this work, we assume a ΛCDM cosmology with $\Omega_m = 0.31$ and a Hubble constant of $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Planck Collaboration XIII 2016). The AllWISE magnitudes we adopted in the work follow the Vega magnitude system (Wright et al. 2010) Finally, we adopted the spectral index definition as $S \propto \nu^\alpha$ throughout the work, where α is the spectral index.

2 SOURCE SELECTION

RGZ DR1 is a catalogue from the first 2.75 yr of the RGZ project (Wong et al., in preparation). Within the catalogue, 99.2 per cent of classifications used radio data from the FIRST survey, and the remainder used data from the ATLAS survey. Each source classification has a user-weighted consensus fraction (consensus level) >0.65 (Wong et al., in preparation). The LAS of each source in the RGZ DR1 is estimated by measuring the hypotenuse of a rectangle that encompasses the entire radio source at the lowest radio contour (Banfield et al. 2015). This method is generally reliable if the radio lobes of a source are correctly identified and the source is not severely bent.

The RGZ DR1 catalogue contains information on individual radio galaxies and their associated radio components, and also a table of cross-matched host galaxies. In this study, we used the catalogue of cross-matched host galaxies as our primary input sample. The original catalogue contains ~140 000 entries. From this catalogue we removed sources without FIRST data available or without host galaxy redshift data in the Sloan Digital Sky Survey (SDSS; Alam et al. 2015). This reduced the sample to 11 549 entries.

Although the RGZ DR1 catalogue requires entries to have a minimum consensus level of 0.65, we found duplicate entries that had (i) the positions of two radio sources separated by less than the pixel size of the FIRST survey (1.8 arcsec), or (ii) multiple host galaxies identified within the same extended radio source. We identified 186 of the first instance and 147 of the second instance. We visually inspected all these source pairs. We then eliminated
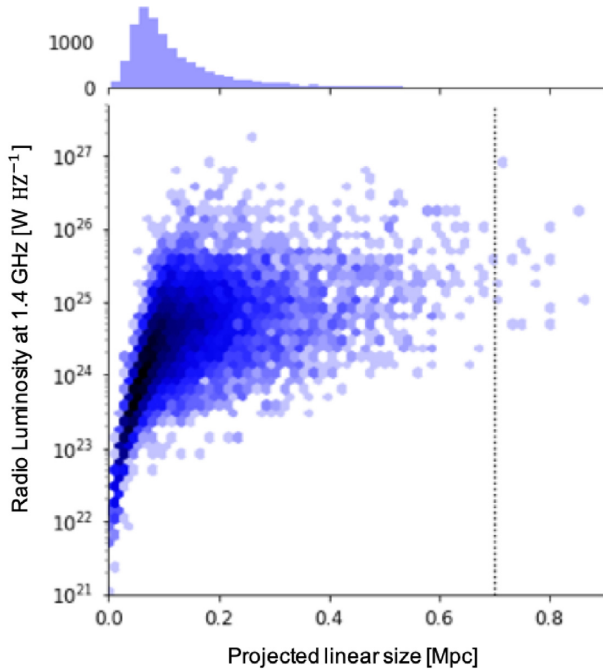
Figure 1. Upper: The projected linear size histogram of the adopted 11 237 RGZ DR1 candidates. Lower: The size–radio luminosity diagram of the same candidates. The colour of each hexagon in the diagram represents the source number density with corresponding size and radio luminosity at 1.4 GHz. The dashed line refers to 700 kpc of linear size.

one source from each pair of the first instance if their LAS and host galaxy redshift were identical. In the second instance, we retained the source in each pair which had a position closer to the *WISE* host galaxy position. This process removed a further 312 objects, which reduced the number of objects in the sample to 11 237 entries.

Fig. 1 shows the size–luminosity diagram of this sample. The projected linear size and 1.4 GHz radio luminosity for each object were calculated using the catalogued RGZ DR1 source LAS, integrated flux density measured from FIRST images, and the host galaxy redshift, assuming a typical spectral index of $\alpha = -0.7$. The majority of the samples share modest radio luminosity and projected linear size. Sources with larger size tend to have higher radio luminosity. Within the sample there are 17 objects that have a projected physical size greater than 700 kpc. The FIRST images for each of these entries were visually inspected, and one additional repeated object was identified and removed. We then crossed-matched the remaining 16 objects with the GRG catalogues of Kuźmicz et al. (2018), Dabhade et al. (2020a), and Kozieł-Wierzbowska, Goyal & Żywucka (2020), and found that Kuźmicz et al. (2018) had previously recorded three of the objects: J0929+4146, J1511+0751, and J1521+5105. We also cross-matched the recent Proctor (2016) inspired GRG candidate catalogue from Dabhade et al. (2020b), and found no overlap. The remaining 13 candidate objects were not found to match any previously known GRG.

3 GIANT RADIO GALAXY IDENTIFICATIONS

For the 13 candidate GRGs, we refined their identifications and measurements using manual inspection of the data. This inspection was used to clean the data set in three steps:

(i) We examined the relationship between the radio structure and the assigned infrared host galaxy of each entry using a *WISE* 3.4 μm image centred on the estimated central radio emission position of the radio galaxy. This process removed three objects where no clear relationship between the radio lobes and the host was seen.

(ii) We re-calculated the LAS of each radio galaxy using the HEALPix Ximview software (HEALPIX; Górski et al. 2005) and the FIRST images, and compared these values with the LAS recorded in the RGZ DR1 catalogue. This process identified three fields containing two radio galaxies that had been misidentified as a single source. In two further fields, we found that the source LAS was overestimated due to confusion with neighbouring sources and that this had also caused the host galaxy to be misidentified. These five objects were removed from the candidate list.

(iii) For objects with misidentified host galaxies [point (i) above], we made a renewed host galaxy search using the NASA Extragalactic Database (NED) and the SDSS Sky Server (SDSS DR15; Aguado et al. 2019). The mid-point of the radio emission was chosen to be the search centre in each case. Given that each image had a side of 3 arcmin, we searched within a radius of 1.5 arcmin. In those cases where a host redshift was found in SDSS DR15, we re-measured the projected linear size of each radio source. This check showed that none of the misidentified sources had a projected linear size larger than 700 kpc.

This three-step data cleaning process resulted in a final sample of five GRGs. Fig. 2 shows the images of these sources; Tables 1–3 summarize the redshift, LAS, linear size, infrared, and radio properties of each object. Redshifts in the tables are extracted from SDSS DR15. Source LASs have been manually re-measured, but are generally consistent (typically 0.5 per cent larger) with those from the original DR1 catalogue.

For the newly identified GRGs, we used visual inspection of the FIRST data to classify each source by morphology and found that four of the five objects to be FR II type. The fifth source, J1646+3627, has an ambiguous morphology.

All five sources have comparatively high radio luminosities, with $\log P_{1.4}\,[\mathrm{W\,Hz^{-1}}] > 25.1$, the mean total radio luminosity of FR II objects as determined by Kozieł-Wierzbowska & Stasińska (2011). Since the host galaxies in each case have W1 − W2 < 0.8 and W2 − W3 < 3.5, where W1, W2, and W3 are the *WISE* observed source magnitudes at 3.4, 4.6, and 12 μm (Cutri & et al. 2013), they are likely to be either elliptical or intermediate disc galaxies (Jarrett et al. 2017).

The five GRG sources are as follows:

J0941+3126 This source is also known as B2 0938+31A, and is centred at J2000.0 RA 9^{h}41^{m}01^{s}24 DEC +31°26′32″.3 (Colla et al. 1970, 1972, 1973; Fanti et al. 1974). The source is hosted by SDSS J094103.62+312618.7. The source has a flux density of 20 mJy at 15.2 GHz (Waldram et al. 2010), and 7.2 ± 3.3 mJy at 30 GHz (Gawroński et al. 2010). Its host has W2 − W3 > 2, redder than is typical for elliptical galaxies and more consistent with the 'intermediate disc galaxy' designation of Jarrett et al. (2017). The host galaxy in this case currently has only photometrically determined redshifts (Alam et al. 2015; Bilicki et al. 2016; Zou et al. 2019), ranged from 0.282 to 0.398. We adopted the lowest one measured by SDSS DR12. We consequently note that this GRG candidate should be treated with caution.

J1331+2557 This source is also known as 7C 1328+2412, and is centred at J2000.0 RA 13^{h}31^{m}18^{s}12 DEC +23°57′07″.4 (Waldram et al. 1996). Its north-east lobe is also known as TXS 1328+242. The host galaxy of this source is identified as
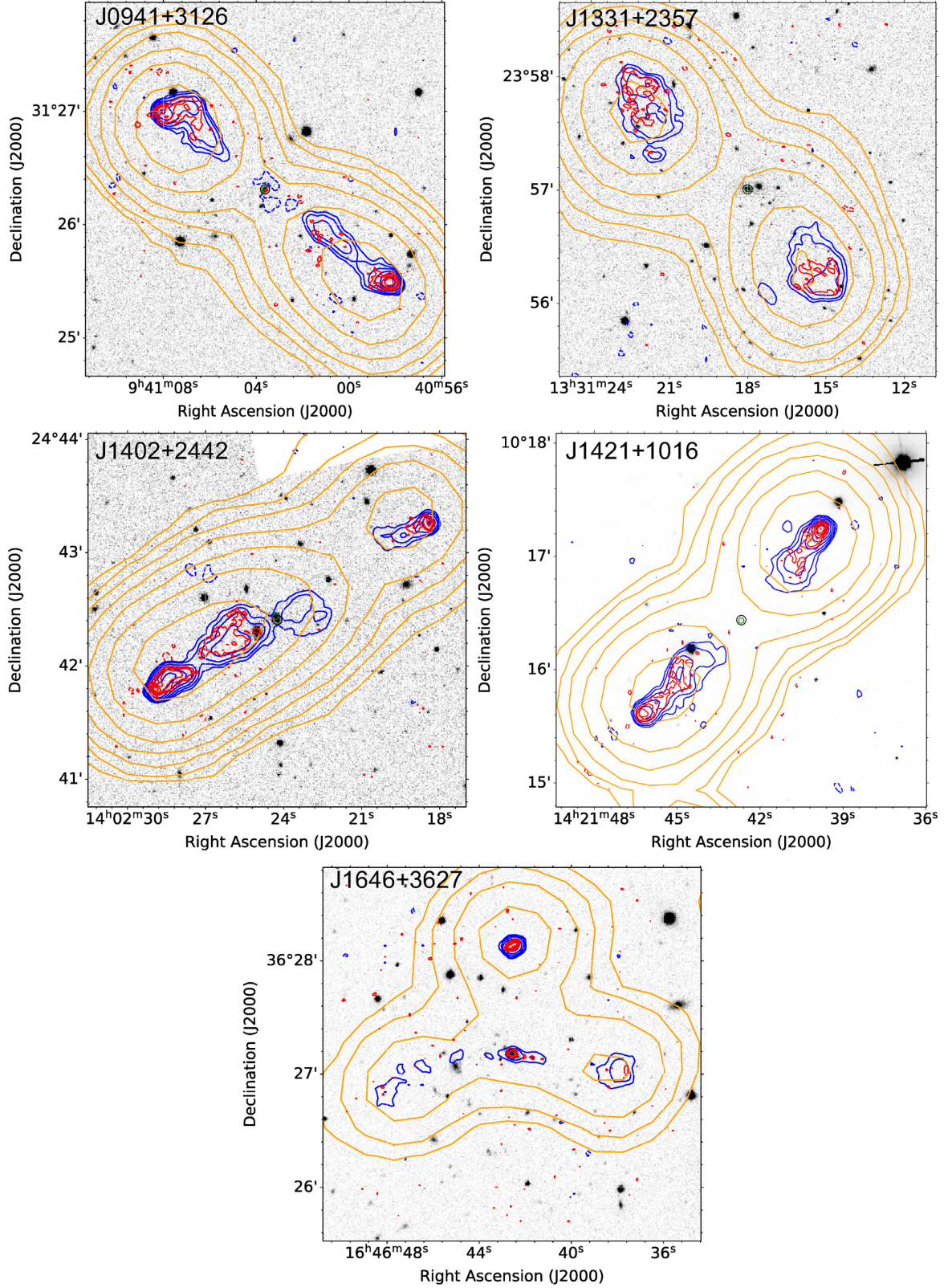
Figure 2. The new GRGs identified in this work. The figure shows radio-near-infrared overlays of these sources, using SDSS *i*-band images rather than *WISE*, given their better angular resolution. The orange, blue, and red radio contours for each source from the NVSS at 1.4 GHz, FIRST at 1.4 GHz and the Karl G. Jansky Very Large Array Survey (VLASS; Lacy et al. 2020) at 3 GHz, respectively, are shown on each image from $3\,\sigma_{rms}$ increasing in steps of a factor of 2. The dashed lines are $-3\,\sigma_{rms}$ of the same survey. The *WISE* candidate host galaxy identified by RGZ DR1 is shown as a green ring, while possible SDSS host galaxies we found are shown in a black ring. The host galaxy position of J1646+3627 concides with the peak brightness of its VLASS/FIRST images. Moreover, the diffuse and compact radio emission above J1646+3627 is irrelevant to the source.

Table 1. A summary of the newly discovered GRGs found in the present work. RGZ ID for each source represents the truncated host galaxy coordinates recorded in the RGZ DR1 catalogue. RA/DEC of source host galaxies are that of the infrared host galaxies shown in Fig 2. The LAS of the sources is measured using **HEALPix Ximview**. For the first four sources, we have assigned errors of 5 arcsec (FWHM) to the LAS of each source, since their leading edges are fairly sharp. In the case of J1646+3627, we have listed the size as a lower limit as the source could be found to extend further given observations with improved sensitivity to larger scale structure. Redshift annotations − p: photometric; s: spectroscopic.

GRG	RGZ ID	RA (J2000.0) (h:m:s)	DEC (J2000.0) (°:′:″)	z	LAS (arcsec)	Size (kpc)
J0941+3126	J094103.6+312618	09:41:03.62	+31:26:18.7	0.282 ± 0.0454^p	163	717 ± 88
J1331+2357	J133117.9+235700	13:31:18.01	+23:57:00.4	0.33610 ± 0.00006^s	162	803 ± 7
J1402+2442	J140224.3+244226	14:02:24.25	+24:42:24.3	0.337 ± 0.032^p	173	810 ± 12
J1421+1016	J142142.6+101626	14:21:42.68	+10:16:26.3	0.37392 ± 0.00003^s	144	765 ± 6
J1646+3627	J164642.5+362710	16:46:42.58	+36:27:10.6	0.43425 ± 0.00010^s	130	$>754 \pm 1$

Table 2. A summary of source infrared properties. *WISE* magnitudes in the table are extracted from the AllWISE catalogue (Cutri & et al. 2013) via VizieR (Ochsenbein, Bauer & Marcout 2000).

GRG	W1	W2	W3
J0941+3126	15.165 ± 0.038	14.650 ± 0.062	11.595 ± 0.204
J1331+2357	14.704 ± 0.030	14.441 ± 0.048	>12.205
J1402+2442	14.763 ± 0.031	14.319 ± 0.045	12.488 ± 0.415
J1421+1016	15.104 ± 0.033	14.703 ± 0.054	12.841 ± 0.512
J1646+3627	13.944 ± 0.141	13.799 ± 0.031	>12.275

SDSS J133118.01+235700.4. We found the source has been observed at 1.4 GHz by the VLA archival public project AG0635 (Fig. 3). The observation has angular resolution of 19.7 arcsec × 13.7 arcsec, along with source flux density of 172 ± 8 mJy. The observation shows the source has visible radio core emission cross-matched with its host galaxy, and its core flux density is 6 ± 3 mJy. Similarly to J0941+3126, the host galaxy of the source has W2 − W3 > 2.

J1402+2442 This source is also known as B2 1400+24, and is centred at J2000.0 RA $14^h02^m25^s.87$ DEC $+24°41'53''.0$ (Colla et al. 1970, 1972, 1973; Fanti et al. 1974). The host of this source is a close pair of galaxies, SDSS J140224.25+244224.3 and SDSS J140224.31+244226.8. The latter has a photometric redshift $z = 0.299 \pm 0.067$ (Alam et al. 2015). We note that although we identify the above galaxy pair as the host for this source, SDSS J140225.03+244218.1 is also in close proximity, see Fig. 2. This source has a photometric redshift of $z = 0.208 \pm 0.018$ (Alam et al. 2015).

J1421+1016 This source is also known as MRC 1419+104, and is centred at J2000.0 RA $14^h21^m42^s.03$ DEC $+10°16'17''.3$ (Large et al. 1981; Large, Cram & Burgess 1991). This source was mentioned by Amirkhanyan, Afanasiev & Moiseev (2015), but not previously identified as a GRG due to differences in the estimation of both the LAS and redshift. This source has host galaxy SDSS J142142.68+101626.2, which is not visible in Fig. 2 where we show the SDSS-i image, but can be seen clearly in *WISE* 3.4 μm data.

J1646+3627 The host galaxy of this source is 2MASX J16464260+3627107. It is the BCG in the galaxy cluster GMBCG J251.67741+36.45295 (Hao et al. 2010) and has a slightly bent morphology, see Fig. 2. This morphology is consistent with the findings of Garon et al. (2019) who used 4304 extended radio sources from RGZ to determine that BCGs have higher probabilities than other cluster members to have slightly bent morphologies.

4 ANALYSIS AND DISCUSSION

The overall occurrence of GRGs in the RGZ DR1 catalogue is 0.08 per cent, which is slightly lower than that of LoTSS DR1. There are two potential reasons for this difference. First, the RGZ citizen scientists are provided with only small-sized images to classify (3 × 3 arcmin^2), which limits the LAS of radio galaxies that can be fully contained in the image cut-outs. Among the 11 237 galaxies considered in this work, the maximum source LAS is 195 arcsec. For GRGs to have angular sizes smaller than this requires them to lie at redshifts $z \geq 0.213$. Under a similar restriction, Dabhade et al. (2020a) would have missed 26.3 per cent of their discovered GRGs. Correspondingly, the Kozieł-Wierzbowska et al. (2020) and Kuźmicz et al. (2018) samples would have missed as much as 62.5 per cent and 66.4 per cent of their catalogued GRGs within the sky area covered by RGZ DR1, respectively. We also take the redshift limitation given by SDSS DR12 into account, while SDSS DR12 is able to detect quasar as far as $z = 6.440$ (Alam et al. 2015), which is further than any detected GRG. Secondly, GRGs have historically been poorly detected in radio surveys like FIRST in part due to their synchrotron spectral index. The radio lobes of GRGs share relatively steep spectral indices, i.e. they are brighter at lower frequencies and thus in principle can more easily be found at MHz frequencies compared to GHz (Dabhade et al. 2020a).

In addition, finding GRGs in radio surveys like FIRST is limited by instrumental considerations. Interferometers with comparatively long baselines (as a function of wavelength) may not be sensitive to the large-scale emission associated with extended or diffuse radio sources (Saxena et al. 2018), nor may it always be encompassed by the comparatively small field of view for single-pixel centimetre-wave receivers. Such issues have in part been alleviated by radio telescopes such as the Expanded Very Large Array (EVLA; Sahr, Hunt & Cornwell 2002), and LOFAR at MHz-frequencies, and by telescopes with large instantaneous fields of view due to Phased Array Feed (PAF) technology, such as the Australian SKA Pathfinder (ASKAP; Johnston et al. 2008); however, whilst these instruments may be powerful probes of GRGs in the future (Peng et al. 2015) instrumental selection effects will always persist.

Consideration of selection effects is of particular importance in the context of developing automated deep learning based GRG classifiers. Such algorithm development is complicated by a lack of large, uniform, and reliable cross-matched radio source catalogues that contain source information characterized in a consistent manner appropriate for the formation of computationally tractable training data. Furthermore, a key aspect of the development of potential machine learning based GRG classification algorithms, as well as radio galaxy classification more generally, is a clear understanding of the biases that are introduced by this training data selection. In

Table 3. A summary of source radio fluxes. Fluxes in the table are measured in mJy. Surveys including the VLA Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014), the NRAO VLA Sky Survey (NVSS; Condon et al. 1998), FIRST (Becker et al. 1995), and VLASS (Lacy et al. 2020) are done by the Karl G. Janksy Very Large Array (VLA; Thompson et al. 1980). Source radio fluxes from the GMRT 150 MHz all-sky radio survey (TGSS ADR1; Intema et al. 2017), the Westerbork Northern Sky Survey at 325 MHz (WENSS; Rengelink et al. 1997) are also measured. We further found literature flux densities from the 7C survey of radio sources at 151 MHz (Waldram et al. 1996), the Texas Survey of Radio Sources at 365 MHz (TXS; Douglas et al. 1996), and the MIT-Green Bank Survey at 5 GHz (MG1,MG2; Bennett et al. 1986; Langston et al. 1990). Source flux densities are calibrated to a common flux scale of Scaife & Heald (2012). The radio luminosity $\log P_{1.4}$ is based on the NVSS images.

| GRG	VLSSr 73.8 MHz	TGSS ADR1 147.5 MHz	7C 151 MHz	WENSS 325 MHz	TXS 365 MHz	NVSS | FIRST 1.4 GHz	VLASS 3 GHz	MIT-Green 5 GHz	$\log P_{1.4}$ (W Hz^{-1})
J0941+3126	1160 ± 202	925 ± 68	930 ± 76	302 ± 87		219 ± 26 | 144 ± 2	126 ± 1		25.73
J1331+2357	2747 ± 676	1214 ± 62	1270 ± 149		901 ± 105	177 ± 22 | 100 ± 1	143 ± 1		25.81
J1402+2442	2728 ± 541	1592 ± 111	1840 ± 140		812 ± 73	381 ± 41 | 263 ± 4	120 ± 1	137	26.15
J1421+1016	2078 ± 495	884 ± 67			701 ± 75	242 ± 30 | 184 ± 4	98 ± 1	106	26.05
J1646+3627	521 ± 160	53 ± 6		41 ± 15		35 ± 6 | 30 ± 1	17 ± 3		25.36
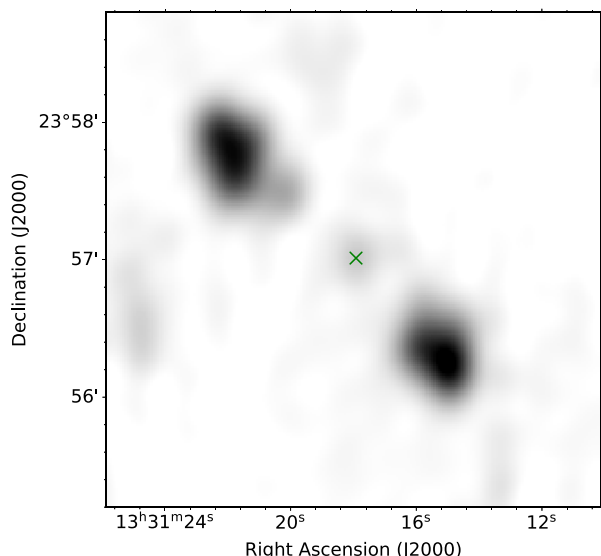
Figure 3. The processed greyscale image of J1331+2357 using the VLA public project AG0635 data, where a faint but visible core is seen at the center of the image. The green cross in the image indicates the host galaxy position given in Table 1.

this respect the RGZ DR1 catalogue represents a well-understood data sample where considerations such as input image size are pre-defined. Hence, although the restricted image size is considered a disadvantage for compiling large catalogues of GRGs, it is potentially an advantage for defining a deep learning training data set with well understood data constraints.

4.1 Radio source luminosity

We measured the integrated flux densities for each source using images from the VLSSr, TGSS ADR1, WENSS, NVSS, FIRST, and VLASS surveys. We also retrieved literature integrated source flux densities from the 7C, TXS, 9C, and MIT-Green Bank (MG) surveys using the NASA/IPAC Extragalactic Database,[1] these are listed in Table 3. At low frequencies, all historic data in Table 3 has been re-scaled to match the Scaife & Heald (2012) flux density scale,

---

[1]NED is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

which is consistent with the Perley & Butler (2017) flux scale at higher frequencies. The five GRGs identified in this paper have higher integrated flux densities in the NVSS survey (FWHM = 45 arcsec) than the FIRST survey (FWHM = 5.4 arcsec), which is consistent with a lack of shorter baseline coverage in the FIRST survey compared to that of NVSS. We note that the VLASS measurement should also be treated with caution as all five objects have diffuse emission on angular scales larger than 30 arcsec, which will be poorly recovered by this survey and result in underestimated integrated flux densities (Lacy et al. 2020).

The resulting spectra for all sources are shown in Fig. 4. We find that these GRGs have a range of source spectral indices from $-0.84 < \alpha < -0.62$, with an average spectral index of $\langle\alpha\rangle = -0.75$. This is similar to the mean spectral index, $\langle\alpha\rangle^{1.4}_{0.151} = -0.79$, found for the GRG sample of Dabhade et al. (2020a) and is also consistent with the typical value for radio galaxies more generally (e.g. Kuźmicz et al. 2018). Finally, our result happens to have the same view with Hardcastle et al. (2019) and Shabala et al. (2020) that these long-lived large radio galaxies are the tail of the radio galaxy age distribution.

Since FIRST and VLASS have comparable resolution and flux loss on similar scales, we also compute the spectral index, $\alpha^{3.0}_{1.4}$, of the source core and lobes separately for J1646+3627 where the radio core is visible and has peak flux above $3\,\sigma_{rms}$ in both surveys. We found that the core region of the source has $S_{1.4} = 2.99 \pm 1.21$ mJy and $S_3 = 2.7 \pm 0.3$ mJy, giving a source core spectral index of $\alpha^{3.0}_{1.4} = -0.13$, and $\alpha^{3.0}_{1.4} < -0.69$ for the lobes. This is consistent with other resolved radio systems, where superposition of multiple synchrotron emission components results in a flatter core spectrum.

4.2 GRGs that are also BCGs

The GRG J1646+3627, newly identified here, is also the BCG in the galaxy cluster GMBCG J251.67741 + 36.45295 (Hao et al. 2010). To better understand this emerging population, we performed a literature search for GRGs that are already known but have not previously been identified as a BCG. Following Dabhade et al. (2020a), we cross-matched the Kuźmicz et al. (2018) catalogue with the GMBCG (Hao et al. 2010) and WHL (Wen, Han & Liu 2012) galaxy cluster catalogues.

This search returned 13 new BCG GRG candidates, which are listed in Table 4. Ten out of 13 of these candidates have not been identified as BCGs previously due to a historic lack of availability of large-scale optical galaxy cluster catalogues when they are discovered. The other three candidates are not recognized as finding BCG GRGs was not highlighted in Proctor (2016).
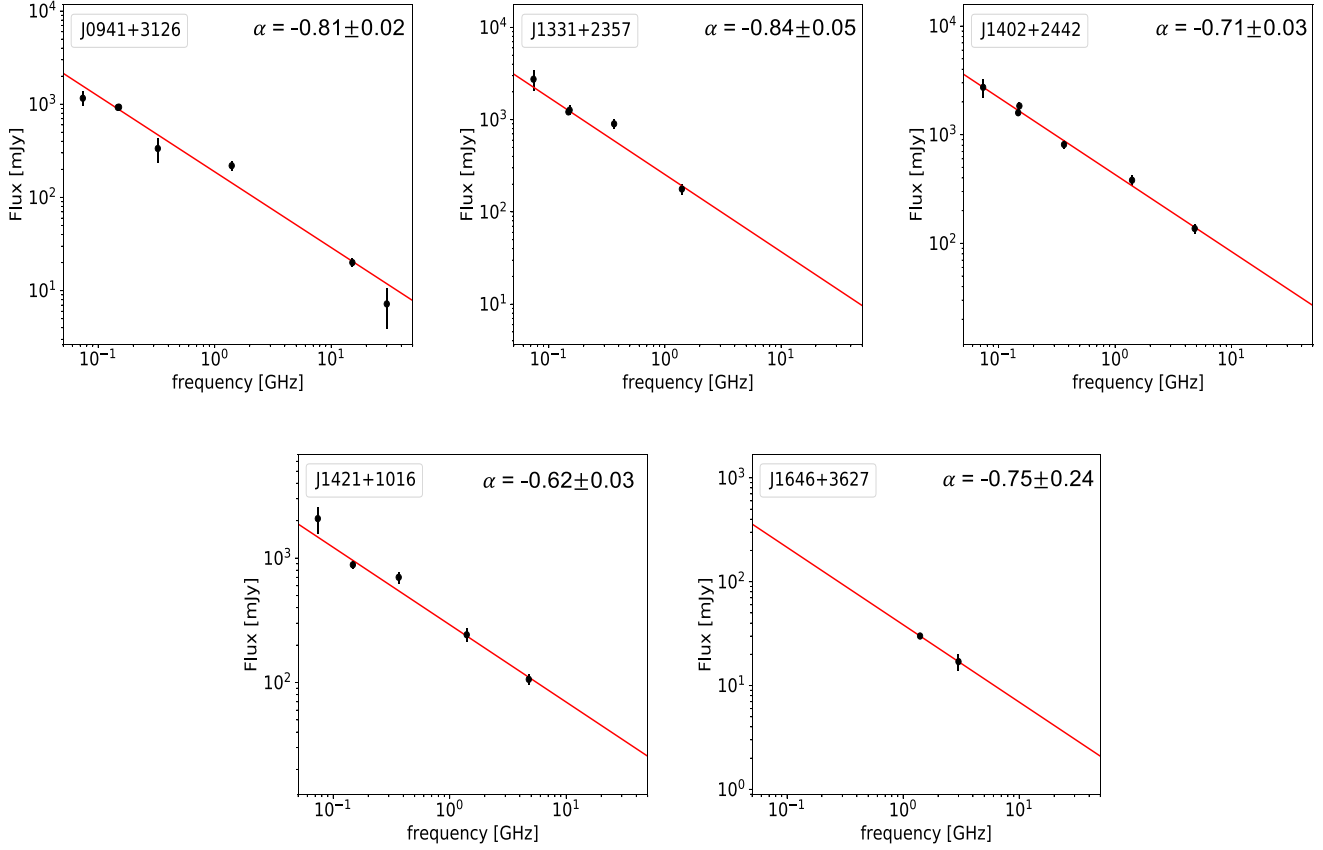
Figure 4. Continuum radio spectra of the GRGs from this work. The solid red lines are linear least-squared fits, where the data points are weighted with their measurement errors when estimating the source spectral indices. Data points used for deriving source spectral indices are from Table 3 and Section 3. Considering the angular resolution differences between surveys, we used data from the NVSS survey at 1.4 GHz and not VLASS data for the top four sources. When deriving the source and core spectral index of J1646+3627, we consider only FIRST and VLASS as they show clear radio core emission and have comparable angular resolution. We were unable to identify clearly visible radio cores in other cited surveys.

Table 4. A summary of the BCG GRG candidates we found from Kuźmicz et al. (2018). RA/DEC, redshift, FR type, and reference number are extracted from Kuźmicz et al. (2018). The galaxy cluster ID are extracted from GMBCG (Hao et al. 2010) and WHL (Wen et al. 2012) galaxy cluster catalogues. R_{200}: the radius of a cluster that its mean density is 200 times of the critical density of the universe; N_{200}: the galaxy number within the R_{200}; R_{L*}: cluster richness; M_{200}: the mass of a cluster that its mean density is 200 times of the critical density of the universe, which is derived from R_{L*} using the equation 2 of Wen et al. (2012).

GRG ID	Cluster ID	RA (J2000.0) (h:m:s)	DEC (J2000.0) (°:′:″)	z	R_{200} (Mpc)	N_{200}	R_{L*}	M_{200} (10^{14} M$_\odot$)	FR type	Ref.
J1054+0227	GMBCG J163.58817+02.46528	10:54:21.16	+02:27:55.0	0.34	–	–	–	–	II	7
J1400+3019	GMBCG J210.18097+30.32185	14:00:43.43	+30:19:18.7	0.206	–	–	–	–	II	6
J0115+2507	WHL J011557.2+250720	01:15:57.23	+25:07:21.0	0.1836	0.96	15	18.28	1.0	II	7
J0129−0758	WHL J012935.3−075804	01:29:35.26	−07:58:04.3	0.0991^a	1.17	10	28.44	1.6	I/II	4
J0751+4231	WHL J075108.8+423124	07:51:08.80	+42:31:24.2	0.2042	0.98	14	17.87	0.9	II	8
J0902+1737	WHL J090238.4+173751	09:02:38.42	+17:37:51.5	0.1645^a	1.01	14	19.68	1.1	II	7
J0926+6519	WHL J092600.8+651923	09:26:00.82	+65:19:22.7	0.1397	0.84	8	14.41	0.7	I	3
J1108+0202	WHL J110845.5+020241	11:08:45.49	+02:02:40.9	0.1574^a	1.05	26	23.55	1.3	II	4
J1235+2120	WHL J123526.7+212035	12:35:26.67	+21:20:34.8	0.4227	0.79	10	12.03	0.6	II	5
J1418+3746	WHL J141837.7+374625	14:18:37.65	+37:46:24.5	0.1349	1.17	25	28.14	1.6	II	8
J1453+3308	WHL J145302.9+330842	14:53:02.86	+33:08:42.4	0.2482	0.92	14	16.69	0.9	II	8
J1511+0751	WHL J151100.0+075150	15:11:00.01	+07:51:50.0	0.4594	1.09	17	23.20	1.3	II	1
J2306−0930	WHL J230632.2−093020	23:06:32.18	−09:30:20.6	0.1593	1.03	16	20.35	1.1	I	2

Notes. References: 1. Baum & Heckman (1989), 2. Best et al. (2005), 3. Lara et al. (2001b), 4. Machalski et al. (2007), 5. Nilsson (1998), 6. Parma et al. (1996), 7. Proctor (2016), 8. Schoenmakers et al. (2001).
^{a}The cluster redshift and the source redshift have a difference of 0.03–0.04, the cluster membership of these radio sources should be treated with caution.

Prior to this work, 21 BCG GRGs were identified by Dabhade et al. (2017, 2020a). Combining that sample with the 13 we identify from the literature and the one new GRG BCG from RGZ DR1, there are 35 BCG GRGs known in total. From the full sample of 35 BCG GRGs, 28 are in clusters with catalogued R_{200}, the radius where the mean density is at least 200 times the critical density of the Universe, and N_{200}, the local galaxy number within R_{200} (Wen et al. 2012). N_{200} here only counts galaxies with $M_e^r(z) \leq -20.5$,
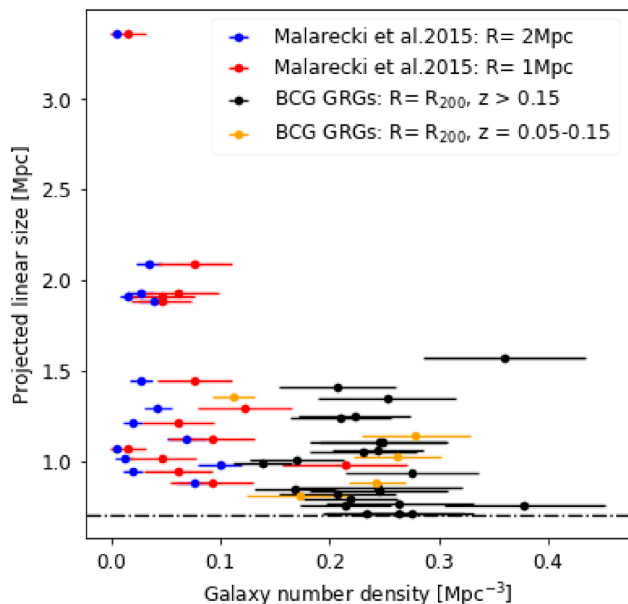
Figure 5. A diagram of galaxy number density versus source projected linear size, comparing samples discussed in Malarecki et al. (2015) and BCG GRGs with R_{200} and N_{200} available from the WHL catalogue. R_{200} and N_{200} of each BCG GRGs in the diagram can be found in table 4 and the table 3 of Dabhade et al. (2020a). The galaxy number density uncertainty of the BCG GRGs is estimated based on equation 1 of Wen et al. (2012) and our cylindrical volume assumption. The galaxy number density uncertainty of Malarecki et al. (2015) samples are extracted from table 4 of their work. The dashed line in the diagram equals 700 kpc.

where $M_{e}^{r}(z)$ refers to evolution-corrected absolute magnitude in the r band (Wen et al. 2012): $M_{e}^{r}(z) = M_r(z) + Qz$, for which a passive evolution of $Q = 1.62$ was adopted (Blanton et al. 2003). Using these data, the relationship between local galaxy density and projected linear size for these galaxies is shown in Fig. 5. Of these 28 objects, there are five with host galaxy redshifts $0.05 < z < 0.15$, the same range used by Malarecki et al. (2015) who also investigated the relationship between GRG size and local environment. We re-calculate the galaxy number density of each cluster for these five objects assuming a cylindrical volume with a radius of R_{200} for each source. Consistent with Malarecki et al. (2015), we adopt a physical cylinder length equivalent to $z = 0.1 \pm 0.003$. This returns galaxy number density values from 0.11 to 0.27 Mpc^{-3} with a median galaxy number density of 0.24 Mpc^{-3}. These are shown as yellow data points in Fig. 5. Original data from Malarecki et al. (2015) are shown as blue data points; however, the values of R_{200} for these galaxies are generally closer to 1 Mpc than the cylinder radius of 2 Mpc used by Malarecki et al., consequently we also show the local galaxy density for the sources in Malarecki et al. (2015) re-calculated using a cylinder radius of 1 Mpc. These data are shown as red points in Fig. 5. The maximum galaxy number density of the 23 BCG GRGs with with host galaxy redshifts $z > 0.15$ under the same volume assumption is 0.38 Mpc^{-3}.

From Fig. 5, it can be seen that the BCG GRGs have been growing in generally denser environments than the non-cluster/poor cluster GRGs in the sample of Malarecki et al. (2015). When considering a radius of 1 Mpc, source B 1308-441 from the Malarecki et al. (2015) sample has a comparable local galaxy number density to the BCG GRGs, due to a concentration of galaxies in close proximity. The mean galaxy number density of the Malarecki et al. (2015) GRGs

using a cylinder radius of 1 Mpc is 0.07 Mpc^{-3}, typical for a poor cluster or galaxy group.

For the BCG GRGs, we also compute the cluster mass, M_{200}, in each case. With the exception of WHL J112126.4+534457 with a mass of 4.6×10^{14} M$_{\odot}$, the masses of these clusters lie in the range 0.7–2×10^{14} M$_{\odot}$. Given that the average M_{200} for the WHL catalogue is $\sim 1.12 \times 10^{14}$ M$_{\odot}$, the masses of these particular clusters are unremarkable with respect to the wider catalogue.

5 CONCLUSION

In this work, we have identified five new GRGs from RGZ DR1. These GRGs mostly share an FR II radio morphology and cover the redshift range of $0.28 < z < 0.43$. These GRGs have been identified using a method consistent with the assembly of training data appropriate for a deep learning classifier. We compare the selection of these GRGs to previous studies and suggest that samples defined in this manner are more likely to be representative of future deep learning approaches to GRG identification than previous methods.

We associate one of the newly identified GRGs, J1646+3627 with the BCG in galaxy cluster GMBCG J251.67741+36.45295 (Hao et al. 2010) and using literature data we identify a further 13 previously known GRGs to be BCG candidates. This increases the number of known BCG GRGs by more than 60 per cent. We show that the local galaxy density of these sources is significantly higher than that of non-cluster GRGs, challenging the hypothesis that GRGs are able to grow to such large sizes only due to locally underdense environments.

DATA AVAILABILITY

The data underlying this article were provided by Radio Galaxy Zoo team by permission. Data will be shared on request to the corresponding author with permission of the head of Radio Galaxy Zoo team.

REFERENCES

Aguado D. S. et al., 2019, ApJS, 240, 23
Alam S. et al., 2015, ApJS, 219, 12

Alhassan W., Taylor A. R., Vaccari M., 2018, MNRAS, 480, 2085
Amirkhanyan V. R., Afanasiev V. L., Moiseev A. V., 2015, Astrophys. Bull., 70, 45
Aniyan A. K., Thorat K., 2017, ApJS, 230, 20
Banfield J. K. et al., 2015, MNRAS, 453, 2326
Baum S. A., Heckman T., 1989, ApJ, 336, 681
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Bennett C. L., Lawrence C. R., Burke B. F., Hewitt J. N., Mahoney J., 1986, ApJS, 61, 1
Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9
Bilicki M. et al., 2016, ApJS, 225, 5
Blanton M. R. et al., 2003, ApJ, 592, 819
Capetti A., Massaro F., Baldi R. D., 2017a, A&A, 598, A49
Capetti A., Massaro F., Baldi R. D., 2017b, A&A, 601, A81
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
Colla G. et al., 1970, A&AS, 1, 281
Colla G. et al., 1972, A&AS, 7, 1
Colla G. et al., 1973, A&AS, 11, 291
Condon J. J . et al., 1998, AJ, 115, 1693
Cutri R. M. et al., 2013, yCat, II/328
Dabhade P . et al., 2017, MNRAS, 469, 2886
Dabhade P. et al., 2020, A&A, 635, A5
Dabhade P. et al., 2020, preprint (arXiv:2005.03708)
Douglas J. N., Bash F. N., Bozyan F. A., Torrence G. W., Wolfe C., 1996, AJ, 111, 1945
Fanaroff B. L., Riley J. M., 1974, MNRAS, 167, 31P
Fanti C., Fanti R., Ficarra A., Padrielli L., 1974, A&AS, 18, 147
Franzen T. M. O . et al., 2015, MNRAS, 453, 4020
Garon A. F. et al., 2019, AJ, 157, 126
Gawroński M. P. et al., 2010, MNRAS, 406, 1853
Gendre M. A., Wall J. V., 2008, MNRAS, 390, 819
Gendre M. A., Best P. N., Wall J. V., 2010, MNRAS, 404, 1719
Górski K. M . et al., 2005, ApJ, 622, 759
Hao J . et al., 2010, ApJS, 191, 254
Hardcastle M. J. et al., 2019, A&A, 622, A12
Intema H. T., Jagannathan P., Mooley K. P., Frail D. A., 2017, A&A, 598, A78
Ishwara-Chandra C. H., Saikia D. J., 1999, MNRAS, 309, 100
Jarrett T. H . et al., 2017, ApJ, 836, 182
Johnston S. et al., 2008, Exp. Astron., 22, 151
Kaiser N. et al., 2002, in Anthony Tyson J., Wolff S.eds, Survey and Other Telescope Technologies and Discoveries, Proc. SPIE Conf. Ser. Vol. 4836. SPIE, Bellingham
Kaiser N. et al., 2010 , in Stepp L. M., Gilmozzi R., Hall H. J, eds, Ground-based and Airborne Telescopes III, Proc. SPIE Conf. Ser. Vol. 7733. SPIE, Bellingham, p. 77330E
Komberg B. V., Pashchenko I. N., 2009, Astron. Rep., 53, 1086
Kozieł-Wierzbowska D., Stasińska G., 2011, MNRAS, 415, 1013
Kozieł-Wierzbowska D., Goyal A., Żywucka N., 2020, ApJS, 247, 53
Kuźmicz A., Jamrozy M., Bronarska K., Janda-Boczar K., Saikia D. J., 2018, ApJS, 238, 9
Lacy M. et al., 2020, PASP, 132, 035001
Laing R. A., Riley J. M., Longair M. S., 1983, MNRAS, 204, 151
Lane W. M . et al., 2014, MNRAS, 440, 327
Langston G. I., Heflin M. B., Conner S. R., Lehar J., Carilli C. L., Burke B. F., 1990, ApJS, 72, 621
Lara L., Cotton W. D., Feretti L., Giovannini G., Marcaide J. M., Márquez I., Venturi T., 2001a, A&A, 370, 409
Lara L., Márquez I., Cotton W. D., Feretti L., Giovannini G., Marcaide J. M., Venturi T., 2001b, A&A, 378, 826

Large M. I., Mills B. Y., Little A. G., Crawford D. F., Sutton J. M., 1981, MNRAS, 194, 693
Large M. I., Cram L. E., Burgess A. M., 1991, The Observatory, 111, 72
Lonsdale C. J . et al., 2003, PASP, 115, 897
Lukic V . et al., 2018, MNRAS, 476, 246
Lukic V . et al., 2019, MNRAS, 487, 1729
Ma Z . et al., 2019, ApJS, 240, 34
Machalski J., Jamrozy M., Zola S., 2001, A&A, 371, 445
Machalski J., Koziel-Wierzbowska D., Jamrozy M., 2007, Acta Astron., 57, 227
Malarecki J. M., Jones D. H., Saripalli L., Staveley-Smith L., Subrahmanyan R., 2015, MNRAS, 449, 955
Mingo B. et al., 2019, MNRAS, 488, 2701
Nilsson K., 1998, A&AS, 132, 31
Ochsenbein F., Bauer P., Marcout J., 2000, A&AS, 143, 23
Ortega-Minakata R. A., Torres-Papaqui J. P., Andernach H., 2013, in Torres Martinez M., ed., Supercomputing in Mexico: A Navigation Through Science and Technology. Univ. de Guadalajara, p. 47
Parma P., de Ruiter H. R., Mack K.-H., van Breugel W., Dey A., Fanti R., Klein U., 1996, A&A, 311, 49
Peng B., Chen R. R., Strom R., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 109
Perley R. A., Butler B. J., 2017, ApJS, 230, 7
Planck Collaboration XIII, 2016, A&A, 594, A13
Proctor D. D., 2016, ApJS, 224, 18
Rengelink R. B., Tang Y., de Bruyn A. G., Miley G. K., Bremer M. N., Roettgering H. J. A., Bremer M. A. R., 1997, A&AS, 124, 259
Safouris V., Subrahmanyan R., Bicknell G. V., Saripalli L., 2009, MNRAS, 393, 2
Sahr B., Hunt G., Cornwell T., 2002, ASPC, 281, 160
Saripalli L., Hunstead R. W., Subrahmanyan R., Boyce E., 2005, AJ, 130, 896
Saxena A. et al., 2018, MNRAS, 475, 5041
Scaife A. M. M., Heald G. H., 2012, MNRAS, 423, L30
Schoenmakers A. P., de Bruyn A. G., Röttgering H. J. A., van der Laan H., 2001, A&A, 374, 861
Seymour N. et al., 2020, PASA, 37, e013
Shabala S. S., Godfrey L. E. H., 2013, ApJ, 769, 129
Shabala S. S. et al., 2020, MNRAS, 496, 1706
Shimwell T. W . et al., 2017, A&A, 598, A104
Shimwell T. W . et al., 2019, A&A, 622, A1
Solovyov D. I., Verkhodanov O. V., 2011, Astrophys. Bull., 66, 416
Subrahmanyan R., Saripalli L., Hunstead R. W., 1996, MNRAS, 279, 257
Thompson A. R., Clark B. G., Wade C. M., Napier P. J., 1980, ApJS, 44, 151
van Haarlem M. P. et al., 2013, A&A, 556, 2
Waldram E. M., Yates J. A., Riley J. M., Warner P. J., 1996, MNRAS, 282, 779
Waldram E. M., Pooley G. G., Davies M. L., Grainge K. J. B., Scott P. F., 2010, MNRAS, 404, 1005
Wen Z. L., Han J. L., Liu F. S., 2012, ApJS, 199, 34
Wiita P. J., Rosen A., Gopal-Krishna, Saripalli L., 1989, LNP, 327, 173
Williams W. L. et al., 2019, A&A, 622, A2
Willis A. G., Strom R. G., Wilson A. S., 1974, Nature, 250, 625
Wright E. L. et al., 2010, AJ, 140, 1868
Zou H., Gao J., Zhou X., Kong X., 2019, ApJS, 242, 8

This paper has been typeset from a TEX/LATEX file prepared by the author.