Mon. Not. R. Astron. Soc. **398,** 1757–1770 (2009) doi:10.1111/j.1365-2966.2009.15242.x

Substructure revealed by RR Lyraes in SDSS Stripe 82

L. L. Watkins,[1] N. W. Evans,[1⋆] V. Belokurov,[1] M. C. Smith,[1] P. C. Hewett,[1] D. M. Bramich,[2] G. F. Gilmore,[1] M. J. Irwin,[1] S. Vidrih,[1,3,4] Ł. Wyrzykowski[1] and D. B. Zucker[1,5,6]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[2]*Isaac Newton Group of Telescopes, Apartado de Correos 321, E-38700 Santa Cruz de la Palma, Canary Islands, Spain*
[3]*Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstrasse 12-14, 69120 Heidelberg, Germany*
[4]*Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 17, 1000 Ljubljana, Slovenia*
[5]*Department of Physics, Macquarie University, North Ryde, NSW 2109, Australia*
[6]*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710, Australia*

Accepted 2009 June 12. Received 2009 June 12; in original form 2009 March 16

ABSTRACT

We present an analysis of the substructure revealed by RR Lyraes in Sloan Digital Sky Survey Stripe 82, which covers 2.°5 in declination on the celestial equator over the right ascension range $\alpha = 20.^{\rm h}7$ to $3.^{\rm h}3$. We use the new public archive of light-motion curves in Stripe 82, published by Bramich et al. in 2008, to identify a set of high-quality RR Lyrae candidates. Period estimates are determined to high accuracy using a string-length method. A subset of 178 RR Lyraes with spectrally derived metallicities are employed to derive metallicity–period–amplitude relations, which are then used, together with archive magnitude data and light-curve Fourier decomposition, to estimate metallicities and hence distances for the entire sample. The RR Lyraes lie 5–115 kpc from the Galactic Centre, with distance estimates accurate to ∼8 per cent. The RR Lyraes are further divided into subsets of 316 RRab types and 91 RRc types based on their period, colour and metallicity.

We fit a smooth density law to the distribution as a simple representation of the data. For Galactocentric radii 5–25 kpc the number density of RR Lyraes falls as $r^{-2.4}$, but beyond 25 kpc, the number density falls much more steeply, as $r^{-4.5}$. However, we stress that in practice the density distribution is not smooth, but dominated by clumps and substructure. Samples of 55 and 237 RR Lyraes associated with the Sagittarius Stream and the Hercules–Aquila Cloud, respectively, are identified. Hence, ∼70 per cent of the RR Lyraes in Stripe 82 belong to known substructure, and the sharp break in the density law reflects the fact that the dominant substructure in Stripe 82 – the Hercules–Aquila Cloud and the Sagittarius Stream – lie within 40 kpc. In fact, almost 60 per cent of all the RR Lyraes in Stripe 82 are associated with the Hercules–Aquila Cloud alone, which emphasizes the cloud's pre-eminence. Additionally, evidence of a new and distant substructure – the *Pisces Overdensity* – is found, consisting of 28 faint RR Lyraes centred on Galactic coordinates ($\ell \approx 80^\circ$, $b \approx -55^\circ$), with distances of ∼80 kpc. The total stellar mass in the Pisces Overdensity is $\sim 10^4\,{\rm M}_\odot$ and its metallicity is [Fe/H] ~ -1.5.

Key words: catalogues – astrometry – stars: variables: other – Galaxy: halo – Galaxy: stellar content – Galaxy: structure.

1 INTRODUCTION

In Aristotle's model of the Universe, the stars were fixed on a rotating sphere and eternally invariable. Stellar variability has been known since Fabricius' discovery of Miras in 1596, although the ancient Chinese and Korean astronomers were already familiar with supernovae or 'guest stars'(Clark & Stephenson 1977). Proper motions were discovered in 1718 by Edmund Halley, who noticed that Sirius, Arcturus and Aldebaran had moved from their fixed positions recorded in Aristotelian cosmology.

Despite this long history, our knowledge of both variable stars and high proper motion sources remains very incomplete. As Paczyński (2000) has emphasized, over 90 per cent of variable stars brighter

⋆E-mail: nwe@ast.cam.ac.uk

than 12 mag have not been discovered. There are still comparatively few large archives of variable sources and, as a consequence, our knowledge of many classes of object, including novae, supernovae, RR Lyraes (the focus of this paper) and high proper motion objects remains limited. Indeed, the variable sky remains one of the most unexplored areas in astronomy, with the exciting possibility that even bright variable objects may correspond to completely unknown astronomical phenomena (see e.g. Paczyński 2001).

The modern era of massive variability searches begins with the microlensing surveys like Massive Compact Halo Object (MACHO; Alcock et al. 1993), EROS (Aubourg et al. 1993) and Optical Gravitational Lensing Experiment (OGLE; Udalski et al. 1992). Typically, these surveys monitored millions of stars down to $V \sim 21$ a few times every night over several years in the directions of the Galactic bulge and the Magellanic Clouds. The resulting huge data bases of light curves yielded information on many rare types of astrophysical variability. They were the first projects that harnessed the power of large format CCD cameras and modern computers to show that the acquisition, processing and archiving of millions of photometric measurements was feasible. The surveys were soon followed by high-redshift supernova surveys [such as the High-Z SN Search (Schmidt et al. 1998) and the Supernova Cosmology project (Perlmutter et al. 1999)], which typically had a lower time resolution and smaller area coverage than the microlensing surveys, although much deeper limiting magnitudes.

The Sloan Digital Sky Survey (SDSS; York et al. 2000) provides deep and homogeneous photometry in five bands in a large area around the North Galactic Cap, but with almost no variability information. The main exception is the compilation of repeat scans of the $\sim$290 deg^2 area – known as Stripe 82 (see e.g. Adelman-McCarthy et al. 2008). The data set has allowed the discovery of many new supernovae, which are publicised and followed up spectroscopically with other telescopes (see e.g. Dilday et al. 2008; Frieman et al. 2008). By averaging repeat observations of unresolved sources in Stripe 82, Ivezić et al. (2007) built a catalogue of one million standard stars with r magnitudes between 14 and 22. Sesar et al. (2007) then carried out the first analysis of $\sim$1.4 million variable stars and quasars using a colour–colour plot to assign variable types.

Recently, Bramich et al. (2008) presented a catalogue of almost four million 'light-motion curves' using the data available in Stripe 82. Objects are matched between the $\sim$30 epochs, taking into account the effects of any proper motion over the 8-yr baseline stretching back to the earliest runs in 1998. Thanks to the high quality of the SDSS imaging, excellent astrometric and photometric precision is attainable. Bramich et al. (2008) also provide a higher level catalogue (HLC) which is a set of 229 derived quantities for each light-motion curve. These quantities describe the mean magnitudes, photometric variability and astrometric motion of a subset of objects whose light-motion curve entries pass certain quality constraints.

Variable stars of particular interest are RR Lyraes, which have often been used to identify Galactic substructure – for example, in studies of the Sagittarius Stream (Ivezić et al. 2000; Vivas et al. 2004; Keller et al. 2008), and the Monoceros and Virgo Overdensities (Vivas & Zinn 2006; Keller, Da Costa & Prior 2009). RR Lyraes are particularly useful for three reasons. First, they are constituents of the old, metal-poor halo, in which substructure is abundant. Secondly, they are standard candles, enabling an estimate of their distances to be made. Thirdly, they are bright enough to be detected out to distances of $\sim$130 kpc in the SDSS data, giving us an insight into the structure of the remote Milky Way halo.

In this paper, we select first the variable objects in Stripe 82 and then the subset of RR Lyraes, using the Bramich et al. (2008) light-motion curve catalogue (LMCC) and HLC. Section 2 discusses the selection of the variable objects and their properties, whilst Section 3 describes the identification of the RR Lyrae stars and properties of the population in some detail. Section 4 discusses substructure seen in the distribution of the RR Lyraes, and we conclude in Section 5.

2 VARIABLE STARS IN STRIPE 82

2.1 Variable selection

The HLC contains 3700 548 objects. Every object in the survey is assigned an object type each time it is observed: 3 if it is galaxy like and 6 if it is star like (see section 3.1 of Bramich et al. 2008). For multi-epoch data, the object type averaged over all the observations provides a relatively reliable indicator of whether the source is star like or galaxy like. The cumulative distribution of mean object type in the left-hand panel of Fig. 1 shows that $\sim$55 per cent of objects in the catalogue are purely star like or purely galaxy like. To extract a sample of stars with essentially zero contamination from galaxies, we require that the mean object type is 5.5 or greater. This results in a 'stellar' sample of 1671 582 objects.

For the stellar sample, a cumulative distribution of reduced χ^2 (that is, χ^2 per degree of freedom) for the g and r bands is shown in the right-hand panel of Fig. 1. The value of reduced χ^2 at which the distributions turn over is $\sim$3, which is taken as the χ^2 value below which stars are assumed to be well modelled by a constant baseline.

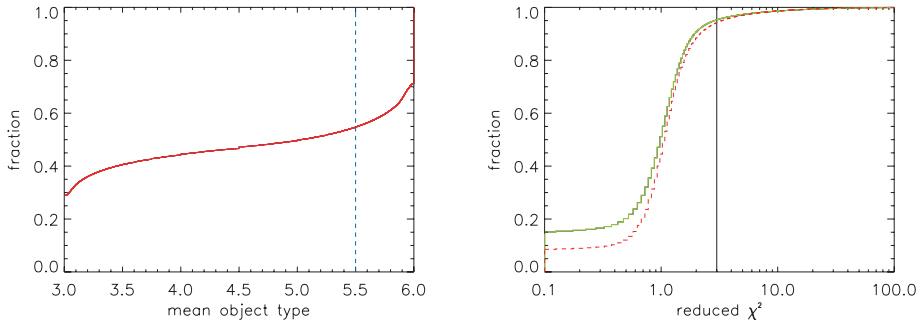

Figure 1. Left: cumulative distribution of mean object type in the HLC. Galaxy-like objects are assigned an object type of 3 and star-like objects are an object type of 6. For multiple observations, the mean indicates whether the object is mostly classified as a star or a galaxy. The cut used to define the stellar sample is shown as a dashed vertical line. Right: cumulative distribution of reduced χ^2 for g (green, solid) and r (red, dashed) bands for all stellar objects. The slope of the distribution turns over at a reduced χ^2 value of $\sim$3, which is taken as the boundary separating constant and variable stars; this value is marked on the graph.

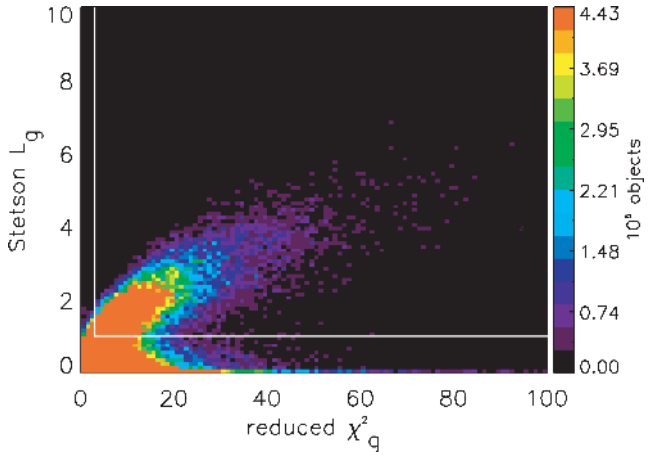

Figure 2. Density plot of reduced χ^2_g against Stetson index L_g for the stellar catalogue. Pixel colour represents the number of objects in each pixel bin. The cuts used to extract the high-quality sample of stellar variables are shown as white lines.

The number of objects that simultaneously satisfy reduced $\chi^2 > 3$ in both g and r is 41 729.

The stars with reduced $\chi^2 > 3$ in both g and r bands are mainly variables, but still contain some artefacts, typically due to one or two outlying photometric measurements. One way to test for true variability is to look for correlations between different bands; a true variable star will usually have changes in brightness that are correlated in all bands whereas a discordant measurement may exist in one band only. Throughout this analysis, the Stetson index (L_g) is used as a measure for correlated variability between the g- and r-band data (see Stetson 1996).

For the stellar sample, a density plot of reduced χ^2 in the g band against the Stetson L_g is shown in Fig. 2, in which two distinct populations can be discerned. The first has an almost linear correlation between reduced χ^2 and L_g, and are almost all true variable stars. The second has a high reduced χ^2, but L_g is low, indicating that the brightness changes which give rise to the high reduced χ^2 values are not correlated between bands. To extract a sample of high-quality variable stars, we impose the simultaneous restrictions $L_g > 1$, reduced $\chi^2_r > 3$ and reduced $\chi^2_g > 3$, together with requiring at least 10 good epochs (see Bramich et al. 2008), leaving 21 939 objects. We present a comparison of the content of our variable catalogue with the earlier catalogue of Sesar et al. (2007) in Appendix A.

To show the quality of the data, a selection of folded light-motion curves is shown in Fig. 3, from which a variety of periodic phenomena such as stellar variability and eclipses are evident. In particular, the two right-most images in the middle row are very likely RR Lyraes. From the periods and the light-curve shapes, we might surmise that the first light curve is an ab-type RR Lyrae and the second is a c-type RR Lyrae.

2.2 Variable properties

Colour–magnitude diagrams (g versus $u - g$) are plotted for both the stellar sample with u-band data and the subset of variable stars, in Fig. 4. For the stellar sample, we see three prominent clumps associated with the thin disc, thick disc and halo on moving redwards in colour. For the variables, there is one prominent clump centred on $u - g \approx 0.2$, which is primarily associated with variable quasars. There are less prominent, but still discernible, peaks associated with variable stars in the thin disc, thick disc and halo.

A crude discrimination between different classes of variable objects is possible in $g - r$ versus $u - g$ space. We find that the variable sample is largely composed of stellar locus stars and low-redshift quasars; stellar locus stars are predominant at bright ($g < 19$) magnitudes while low-redshift quasars dominate at faint ($g < 22$) magnitudes. RR Lyraes also make a significant contribution (see Appendix A for more detail on the statistical properties of the variable sample).

2.3 The proper motions

The HLC offers an improvement over previous variability work in Stripe 82 through the availability of proper motions. The combination of stellar photometry and proper motions has proved to be a powerful way of classifying stars – in particular, members of elusive populations such as white dwarfs, brown dwarfs and wide binaries. Such combined catalogues, drawn from the intersections of SDSS data with USNO-B data (Monet et al. 2003), have been constructed before (Gould & Kollmeier 2004; Munn et al. 2004; Kilić et al. 2006). Compared to such catalogues, the HLC is restricted to Stripe 82 and the proper motion sensitivity is poorer, due to the much shorter time baseline. On the other hand, the Stripe 82 photometric catalogue extends approximately 1.5 mag deeper than the limiting magnitude of USNO-B proper motions ($V \sim 21$).

The reduced proper motion is defined as $H = r + 5 \log \mu + 5$, where r is the apparent magnitude and μ the proper motion in arcsec yr^{-1}. The criteria for inclusion in the reduced proper motion diagram in the top panel of Fig. 5 is that the object lies in our stellar sample, that the proper motion is measured with a signal-to-noise ratio (S/N) > 10 and that the absolute value of the proper motion $|\mu|$ exceeds 2 mas yr^{-1}. The S/N cut has been chosen primarily to ensure easy visibility of structure on the figure, whilst the proper motion cut enables us to excise quasars. We discern three distinct sequences of stars, namely Population I disc dwarfs, Population II main-sequence subdwarfs and disc white dwarfs. Vidrih et al. (2007) have used this reduced proper motion diagram to identify new ultracool and halo white dwarfs in Stripe 82. Similarly, Smith et al. (2009a) and Smith, Evans & An (2009b) have used the same procedure to extract a sample of halo subdwarfs in studies of the velocity ellipsoid and halo substructure. The lower panel of Fig. 5 shows the reduced proper motion diagram just for the subset of variables with proper motion S/N > 5 and $|\mu| > 2$ mas yr^{-1}. The variables are disproportionately drawn from the Population I disc dwarfs, although the other two sequences can still be seen.

3 RR LYRAES

3.1 Identification of RR Lyraes

Here, we develop the tools to extract a high-quality sample of RR Lyraes from the variable catalogue. We begin by selecting the 873 candidates that simultaneously satisfy all the following criteria, which are adapted from Ivezić et al. (2005), namely

$$\begin{aligned} 13.5 < r < 20.7, \quad & 0.98 < u - g < 1.35, \\ -0.16 < r - i < 0.22, \quad & -0.21 < i - z < 0.25, \\ D_{ug}^{\min} < D_{ug} < 0.35, \quad & D_{gr}^{\min} < D_{gr} < 0.55, \end{aligned} \tag{1}$$

where

$$\begin{aligned} D_{ug} &= (u-g) + 0.67(g-r) - 1.07, \\ D_{gr} &= 0.45(u-g) - (g-r) - 0.12. \end{aligned} \tag{2}$$

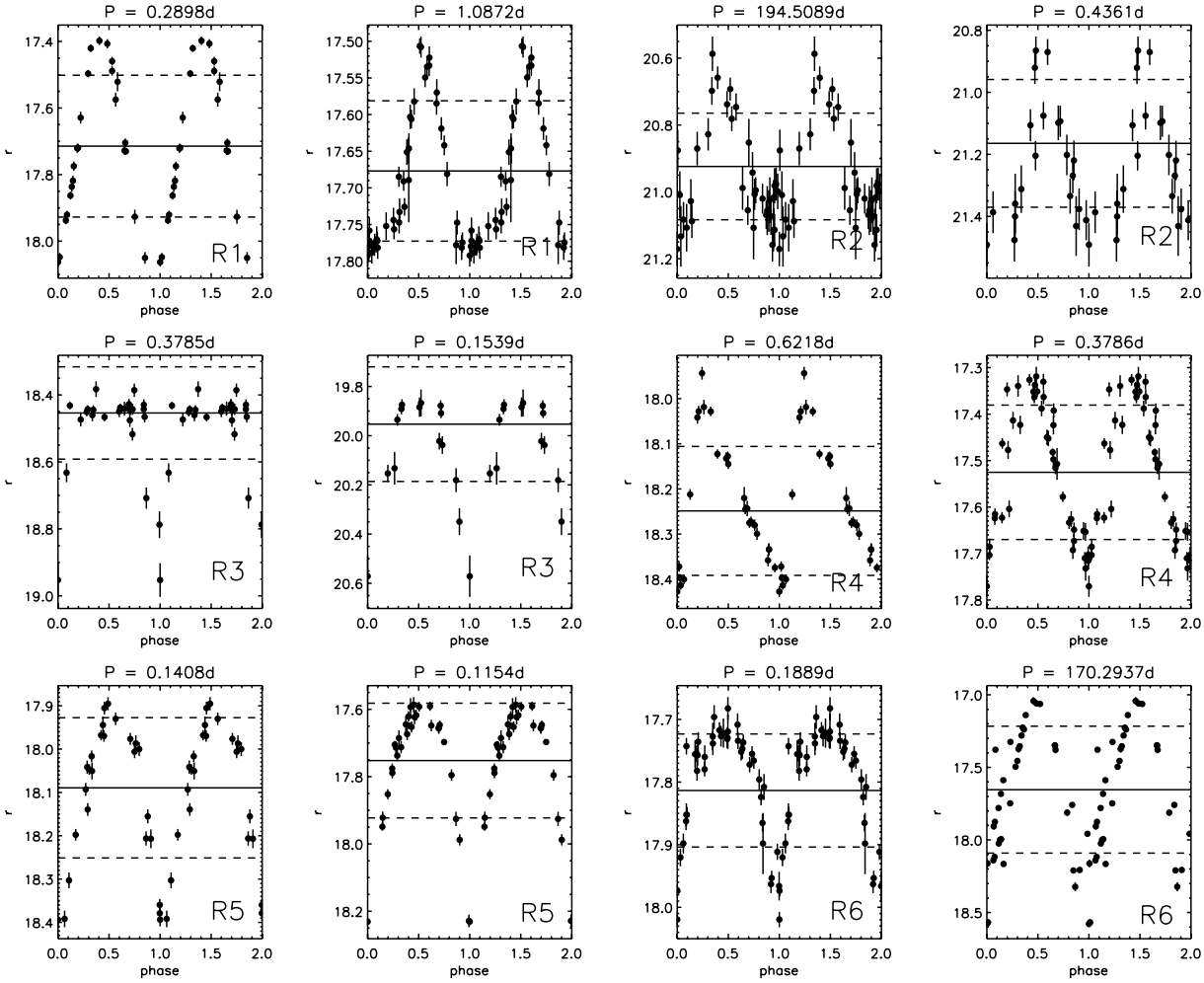

Figure 3. A sample of folded light curves. The period in days is recorded at the top of each plot. The solid line is the mean magnitude, whilst the dashed lines represent 1σ deviations. In the bottom right-hand corner, the number refers to the region in the colour–colour plot in which the light curve lies (see Fig. A1 in Appendix A). The two rightmost light curves in the middle row are probable RR Lyraes.

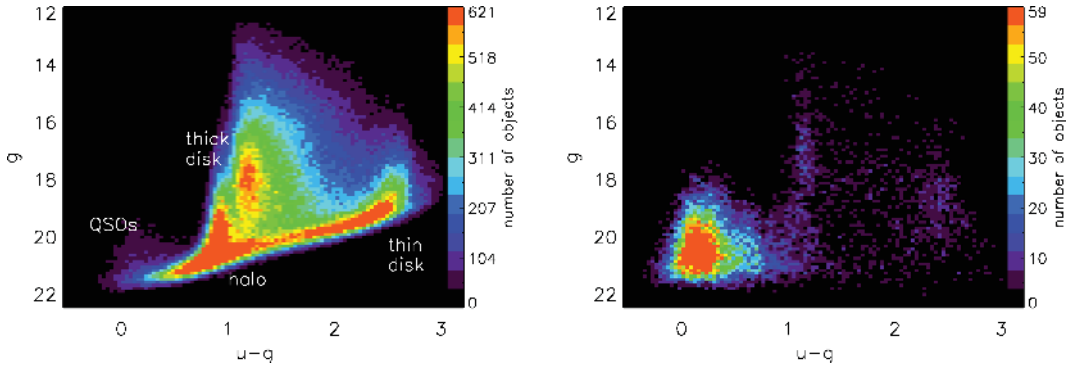

Figure 4. Left: colour–magnitude diagram (g versus $u - g$) for all 527 621 sources in the stellar sample for which high-quality u and g data exist. Right: colour–magnitude diagram (g versus $u - g$) for the subset of 21 939 variables.

The r-band magnitudes correspond to distances ∼5–130 kpc. D_{ug} and D_{gr} represent slopes in the $u - g$ and $g - r$ colour–colour plane. Combined with the cut on $u - g$, the D_{ug} and D_{gr} criteria constrain the RR Lyraes to a hexagonal box in colour–colour space, optimizing the selection of RR Lyraes. The values $D_{ug}^{\min}$ and $D_{gr}^{\min}$ can be altered to adjust the completeness and efficiency of the RR Lyrae selection. We chose to use values $D_{ug}^{\min} = -0.05$ and $D_{gr}^{\min} = 0.06$, which would give a completeness of 100 per cent for the QUEST survey RR Lyraes (Vivas et al. 2004; Ivezić et al. 2005).

3.2 RR Lyrae periods

Determining periods for our RR Lyrae candidates is non-trivial. In general, there are 30 to 40 data points in a light curve, unevenly

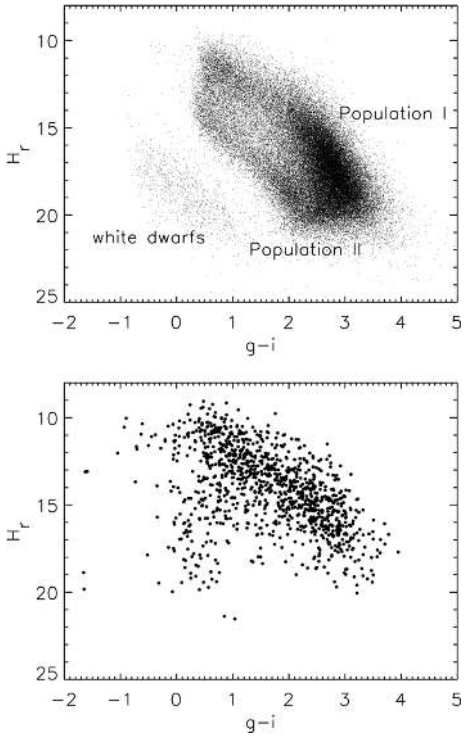

Figure 5. Top: reduced proper motion diagram for all objects in the HLC with proper motion of S/N > 10 and $|\mu| > 2\,\mathrm{mas\,yr^{-1}}$. Bottom: reduced proper motion diagram for all objects in the variable subset with S/N > 5 and $|\mu| > 2\,\mathrm{mas\,yr^{-1}}$.

sampled over an 8-yr baseline. From this sparsely sampled data, we seek a period that is a fraction of a day. The multiband nature of the SDSS survey is an advantage here, as we are able to verify that any period estimate we obtain in one band is consistent with the data in additional bands.

The LMCC contains all the data for a given light curve. Each data point has a flag that is set (or unset) if the data point passes (or fails) certain quality requirements (for more details, see Bramich et al. 2008). For period estimation, it is important that we use only the reliable data to minimize errors. In general, the *g* and *r* bands have smaller errors and fewer outliers and their clean light curves are the best sampled, so these bands are used together to estimate periods for our light curves.

As a first pass, we run a Lomb–Scargle periodogram (see e.g. Press et al. 1989) on each of the *g*- and *r*-band light curves, taking care to ensure that the frequency range extends to the high frequencies (or low periods) expected for RR Lyrae stars and that the sampling rate is detailed enough to discern individual peaks. The resulting period estimates are plotted in Fig. 6. First, we note that there are a number of candidates which have matching period estimates. A number of resonance lines – the locations at which one of the period estimates is a harmonic of the other – are plotted as solid lines in the graph.

The resonance lines are well populated, indicating that the Lomb–Scargle periodogram can return harmonics of the true period as well as the period itself. Hence, we must consider whether the exact period matches are indeed cases where the true period has been recovered, or whether both period estimates are harmonics, and so on. Light curves with periods that do not match or do not lie on resonance lines could be objects that are not periodically variable – quasars, for example – or objects for which the periodogram has failed to recover the true period or a harmonic of the period in one or both cases. Not all of the period estimates lie within the range expected for RR Lyraes (0.2–0.8 d), which is probably a consequence of contamination in our sample. However, it is unclear whether we can simply reject any objects with period estimates larger than those expected for RR Lyraes. Certainly, some of the larger period estimates could be attributed to the sparse sampling of the light curves, which may generate a signal that overwhelms the periodic nature of the RR Lyraes.

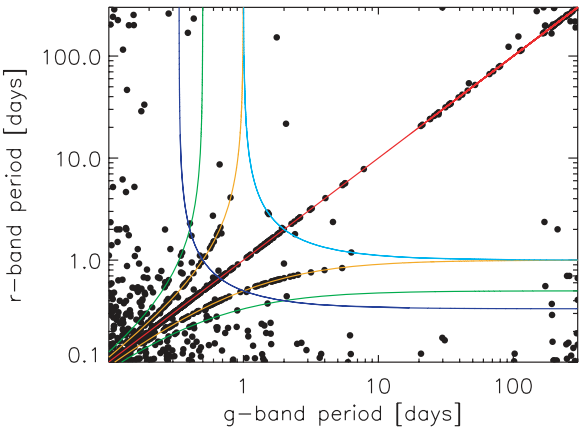

Figure 6. Lomb–Scargle period estimates in the *g* and *r* bands for the candidate RR Lyraes. The coloured lines represent resonance lines along which one period is a harmonic of the other: $P_r = P_g$ (red), $P_r = P_g/(1 \pm P_g)$ (orange), $P_r = (P_g/1 \pm 2P_g)$ (green), $P_r = P_g/(P_g - 1)$ (cyan), $P_r = P_g/(2P_g - 1)$ (blue).

Accordingly, we experimented with a number of alternative methods, including binning, smoothing and phase dispersion minimization, before concluding that a string-length technique is the most effective for our data set. In outline, a string-length method works by phasing a light curve with a trial period and calculating the sum of the straight-line distances between consecutive points. The sum of these distances, or the string-length, will be minimized if the trial period is the true period.

We use a variation of the Lafler & Kinman (1965) string-length technique described in Stetson (1996). For each trial period, the string length is computed for the *g* and *r* bands. Their sum is taken as the overall string length, which is minimized to obtain a period estimate. Running a string-length period finder over a wide and finely-sampled range of trial periods is computationally expensive. However, we can short-cut the process by restricting the string-length period search to a narrow range of periods, centred on a set of the most likely periods identified via the Lomb–Scargle periodogram.

To ensure that the correct RR Lyrae periods are identified, the *g*- and *r*-band Lomb–Scargle periodogram spectra were combined and the highest peaks in each of the ranges $P < 0.2$, $0.2 < P < 0.6$, $0.6 < P < 1.0$ and $P > 1.0$ d were selected. Then, four further peaks were selected from each of the *g*- and *r*-band spectra independently. This was done according to highest amplitude, until four distinct new peaks had been selected in each band. A sample periodgram is shown in Fig. 7 with the period peaks highlighted. Each peak was then considered in turn: the string length was calculated for a narrow range of periods spanning the peak, the period for which the string length was minimized was taken to be the best period in the vicinity of that peak. Finally, the period that returned the shortest string length overall was adopted as the period estimate for the light curve.

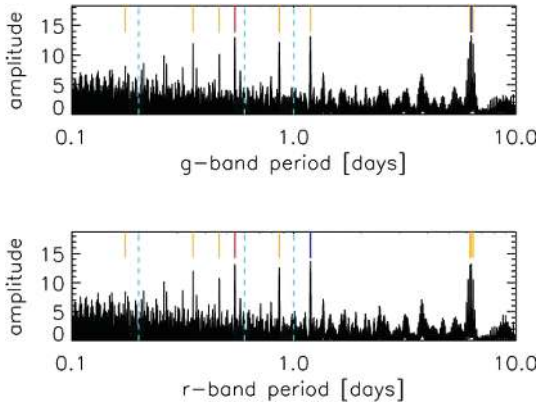

Figure 7. A sample periodogram spectrum. The highest peak in each spectrum is indicated by the blue line; the additional peaks selected by the method described in the text are indicated by the orange lines, with the eventual best-fitting peak marked with a red line. At least one peak was found in each region delineated by the cyan lines.

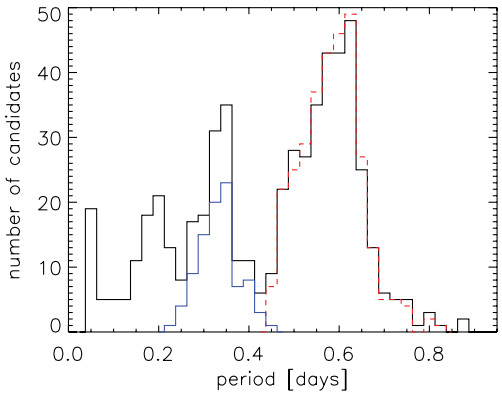

Figure 8. The distribution of periods for all RR Lyrae candidates with periods in the expected range for RR Lyraes (black, solid), for the ab-type RR Lyraes (red, dashed) and the c-type RR Lyraes (blue, solid). The spike at $P < 0.1$ is due in part to δ Scuti and SX Phe stars, and the population with the shortest periods is due mostly to eclipsing variables.

In Fig. 8, the solid black line is the period distribution for all of the RR Lyrae candidates and shows four clear peaks. Moving from high to low periods, these populations are predominantly: ab-type RR Lyraes (peak at ∼0.6 d), c-type RR Lyraes (peak at ∼0.35 d), eclipses (peak at ∼0.18 d) and δ Scuti and SX Phe stars (peak at ∼0.05 d). Also present in this candidate sample are a number of non-periodic variables. The red dashed line is the period distribution for only those stars later determined to be ab-type RR Lyraes and the blue solid line is for those stars we later determine to be c-type RR Lyraes.

Before proceeding any further, we clean the sample of some eclipsing variable stars, δ Scuti stars, SX Phe stars and non-periodic variable contaminants by adopting a stringent cut on Stetson index $L_g > 2.5$. To perform subsequent analysis, we require that objects have a sufficient number of clean data points; that is, we impose further cuts on the number of clean epochs in the g and r bands: $N_g > 5$ and $N_r > 5$, leaving 604 candidates.

Uncertainties in the period estimates may be attributed to two sources. If we assume that the estimate is indeed close to the true period and not a harmonic, then any error is due to the string-length fitting technique, estimated to be $\sim 10^{-5}$ d, from analysis of the phased RR Lyrae light curves. However, for a small fraction of the sample, where the period estimate is a harmonic of the true period, the error will be ∼0.1–0.5 d.

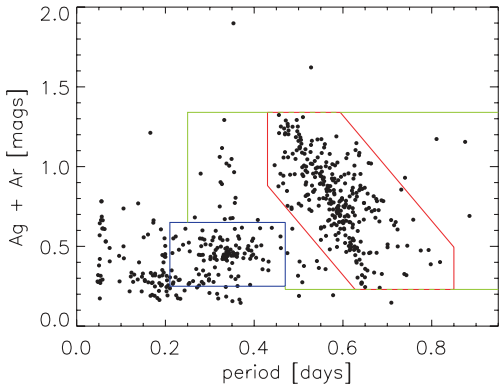

Figure 9. Period versus the combined g- and r-band amplitudes. Three selection boxes are shown: red for candidate RRabs, blue for candidate RRcs and green for candidates warranting further study.

3.3 RR Lyrae classification

Bailey (1902) first divided RR Lyraes into three classes – a, b and c – based on the appearance of their light curves. Further study of Lyraes has altered the classification instead to just ab and c. It is believed that RRab stars are pulsating in the fundamental mode and RRc stars are pulsating in their first overtone (e.g. Smith 1995).

The two classes have somewhat different properties: RRab stars generally have relatively large amplitudes (of order a magnitude) and asymmetric light curves with a steep rising branch and a slow, steady decline. Their periods lie mostly in the range 0.4–1.0 d and studies of RR Lyraes in globular clusters have shown that the amplitude of RRab stars decreases with increasing period. RRc stars have smaller visual amplitudes (around half a magnitude) and their light curves are more symmetrical and nearly sinusoidal. They generally have periods of 0.24–0.5 d.

To separate the two classes, we proceed by plotting the candidates in the plane of (P, A), where P is the period and A is the sum of the amplitudes in the g- and r-band light curves, as shown in Fig. 9. The red box is defined as the region

$$0.43 < P < 0.85, \qquad 0.23 < A < 1.34,$$
$$2.3 < A + 3.3P < 3.3 \tag{3}$$

and includes 296 RRab candidates. The blue box is defined as the region

$$0.21 < P < 0.47, \qquad 0.25 < A < 0.65 \tag{4}$$

and includes 122 RRc candidates.

Finally, the green boxes (not fully shown in Fig. 9 for clarity, but which extend to a period of 2 d) include a number of further candidates, which we do not want to discard without further investigation. Objects with a period of 1 d are almost certainly spurious (the light curves have a sampling period of 1 d) and are removed.

For the 43 objects that remain, there is the possibility that they may lie away from the concentration of RR Lyraes because of an incorrect period estimate. As discussed in Section 3.2, errors in the period estimates are small, but sometimes a harmonic of the true period is obtained. Hence, we revisit the period analysis to determine whether any of the likely period peaks lie within the red RRab box or the blue RRc box. If more than one period peak lies within a box, the period with the minimum string length is used.

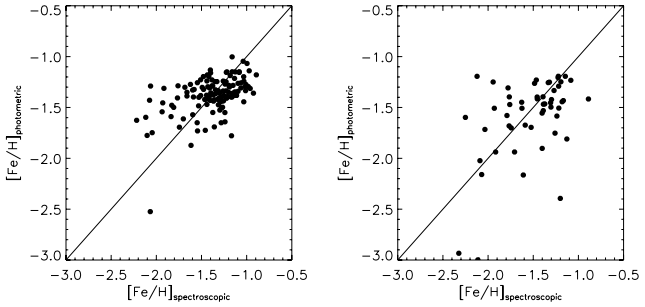

Figure 10. The performance of the photometric metallicity relationships in equations (7) and (9) is shown. The left-hand (right-hand) panel plots the RRab (RRc) candidates with spectroscopic metallicities against the corresponding photometric estimate.

Any objects for which a fitting period can be found are added to the RRab or the RRc candidate set as appropriate. This reclassification results in 330 RRab and 137 RRc candidates.

3.4 The RRab types

126 of our RRab and 52 RRc candidates posses SDSS spectra and have spectroscopic metallicity estimates. The metallicity depends weakly on the phase of the pulsation cycle of the RR Lyrae star at which the spectra was taken (Takeda et al. 2006). We use these objects to calibrate empirical relationships and thence derive photometric metallicities for the entire sample.

Jurcsik & Kovacs (1996) found that the metallicity of an RR Lyrae depends on the period P and the shape of the light curve, which may be parametrized via a Fourier decomposition:

$$f(\theta) = A_0 + \sum_{i=1}^{N} A_i \sin(i\theta + \phi_i). \quad (5)$$

The amplitudes A_i and the phases ϕ_i can then be combined as follows:

$$A_{ij} = \frac{A_i}{A_j}, \qquad \phi_{ij} = j\phi_i - i\phi_j. \quad (6)$$

Inspired by the analogous relation of Jurcsik & Kovacs (1996), we use the spectroscopic metallicities and the light-curve properties to derive the metallicity–period–amplitude-phase relation:

$$[\mathrm{Fe/H}] = 0.845 - 4.487P - 0.187\phi_{31} - 1.454A_{31} + 2.166P\phi_{31} + 1.563PA_{31} - 8.374P^2 - 0.081\phi_{31}^2, \quad (7)$$

which has a typical scatter $\sigma = 0.25$. Its performance is shown in the left-hand panel of Fig. 10. To refine our RRab sample, we insist that

$$-3 < [\mathrm{Fe/H}] < 0, \quad (8)$$

and apply a further restriction to trim the sample of a remaining few eclipsing variables by imposing the selection box in orange (solid lines) in colour–colour space shown in the lower panel of Fig. 11. This leaves us with 325 RRab candidates. Also shown in this figure are the confirmed RRab types as red squares and likely eclipsing variables as green triangles.

3.5 The RRc types

Morgan, Wahl & Wieckhorst (2007) found that the metallicities of c-type RR Lyraes also vary with period and light-curve shape. Owing to the difference in pulsation mode, and hence light-curve shape, between the RRab stars and the RRc stars, a new metallicity relation must be derived for the RRc candidates. The light curves were Fourier decomposed and the 52 RRcs with spectroscopic metallicities were used to derive the following metallicity–period–amplitude relation, the form of which is suggested by Morgan et al. (2007):

$$[\mathrm{Fe/H}] = -10.669 + 60.944P + 4.351A_{21} - 23.418PA_{21} - 95.344P^2 + 1.864A_{21}^2. \quad (9)$$

As the right-hand panel of Fig. 10 shows, there is a somewhat greater scatter of $\sigma = 0.38$ in this relationship than the corresponding one for RRabs.

A metallicity cut $-3.75 < [\mathrm{Fe/H}] < 0$ removes a few obvious outliers. As for the RRabs, we discard eclipsing variables by imposing the selection box in cyan (dashed lines) in colour–colour space shown in the lower panel of Fig. 11. This leaves us with 97 RRc-type candidates. Also shown in this figure are the confirmed RRc types as blue dots and likely eclipsing variables as green triangles.

In the same way that the outlying candidates were re-analysed to determine whether a period could be found that placed the star into the RRab or RRc data clouds, so the RRab and RRc rejects are also re-analysed, just in case they are misclassified RRc and RRab, respectively. The metallicity and colour cuts described above then determine inclusion in the candidate sets. Finally, a visual inspection of the light curves confirms that these are very high quality samples. Judging by light-curve shape, there are at most 21 possible contaminants in the RRab set and only six possible contaminants in the RRc set. The light curves typically possess low S/N and/or are poorly sampled. While many of these objects will be RR Lyrae, to be conservative, they are removed, giving us a final sample of 316 RRab stars and 91 RRc stars, 407 RR Lyraes in all.

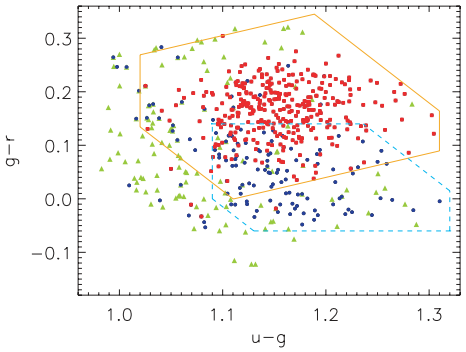

Figure 11. The RR Lyrae candidate selection boxes in the $u - g$ versus $g - r$ plane; the orange (solid) box is used to select RRab types, the cyan (dashed) box for RRc types. The RRab (RRc) candidates that pass the period, amplitude and metallicity cuts are shown as red squares (blue circles). Also shown are suspected eclipsing variables as green triangles.

4 SUBSTRUCTURE REVEALED BY THE RR LYRAES

4.1 RR Lyrae distances

RR Lyraes are 'standard candles' because they have a well-defined absolute magnitude, which, none the less, depends on metallicity. We calculate distances via the distance modulus:

$$\log D = \frac{m_z - M_z + 5}{5}. \quad (10)$$

The apparent magnitudes m_z come directly from the HLC. The absolute magnitudes m_z are obtained via the following relation from

Table 1. The means and dispersions in the properties of the RR Lyrae variables split according to RRab and RRc types. Listed are the amplitudes in the g and r bands, the period, the reduced χ^2 in g and r, the Stetson index, the metallicity, the heliocentric distance D, the Galactocentric distance r, the absolute magnitude M_z and the number of good epochs in g and r.

Property	μ_{ab}	σ_{ab}	μ_c	σ_c
A_g (mag)	0.47	0.14	0.26	0.05
A_r (mag)	0.34	0.11	0.19	0.03
P (d)	0.58	0.07	0.33	0.04
χ^2_g	363.8	326.1	146.6	80.0
χ^2_r	250.9	232.3	98.5	53.8
L_g	10.7	5.2	8.4	2.5
[Fe/H]	−1.38	0.16	−1.61	0.42
D (kpc)	28.9	21.9	21.1	17.6
r (kpc)	29.2	21.7	21.5	17.5
M_z	0.44	0.11	0.71	0.15
N_g	29.6	7.4	29.4	8.4
N_r	29.6	7.7	29.6	8.0

Cáceres & Catelan (2008):

$$M_z = 1.3706 + 0.8941 \log Z + 0.1315[\log Z]^2 - (2.6907 + 0.8192 \log Z + 0.0664[\log Z]^2)\ln C_0 + (47.9836 + 31.7879 \log Z + 5.2221[\log Z]^2)[\ln C_0]^2 + (141.7704 + 100.6676 \log Z + 17.4277[\log Z]^2)[\ln C_0]^3 + (0.3286 + 2.0377 \log Z + 0.3882[\log Z]^2) \log P, \quad (11)$$

where P is the fundamental period, $\log Z = [\mathrm{Fe/H}] - 1.5515$ and $C_0 = (u - g)_0 - (g - r)_0$, with the 0 subscript denoting that the colours are unreddened. The intrinsic scatter in this relation is small compared to the errors. Uncertainties on the metallicities [Fe/H] and distances D are computed using standard methods. From this, we find that the distance errors are typically around 8 per cent (this includes the scatter due to the metallicity relationships).

The right ascension, declination, classification, mean magnitudes, amplitude, period and distance of all our 407 RR Lyrae candidates are given in an accompanying electronic table. The means and dispersions for some useful quantities for the RRab and RRc subsamples are given in Table 1.

4.2 The Sagittarius Stream, the Hercules–Aquila Cloud and the Pisces Overdensity

The distribution of RR Lyraes in right ascension and distance is shown in Fig. 12, with the ab-types plotted as red circles and the c-types plotted as blue triangles. There are a number of things to notice. First, there are 296 RR Lyraes at right ascensions $20^{\rm h}.7 < \alpha < 24^{\rm h}$, but only 111 at $0^{\rm h} < \alpha < 3^{\rm h}.3$. The greatest concentration of RR Lyraes is in the fields coincident with the Hercules–Aquila Cloud (Belokurov et al. 2007). Of course, not all these RR Lyraes are necessarily associated with the cloud, as there may be contamination from the underlying smooth population associated with the Galactic spheroid. It is known that the bulge and spheroid harbour a population of RR Lyraes, distributed in a roughly spherical manner around the Galactic Centre, with a metallicity distribution peaked at [Fe/H] ~ -1 (see e.g. Walker & Terndrup 1991; Alcock et al. 1998; Collinge, Sumi & Fabrycky 2006). The plane of

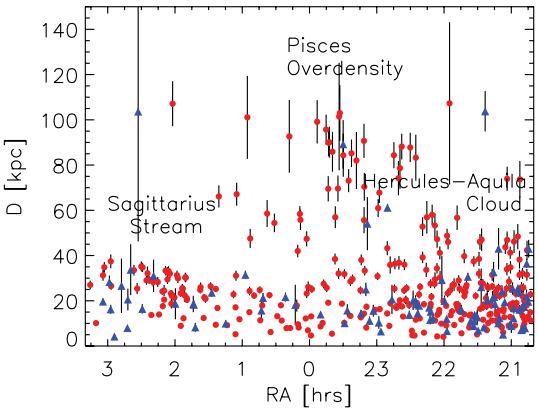

Figure 12. The spatial distribution of the ab-type (red circles) and the c-type (blue triangles) RR Lyraes.

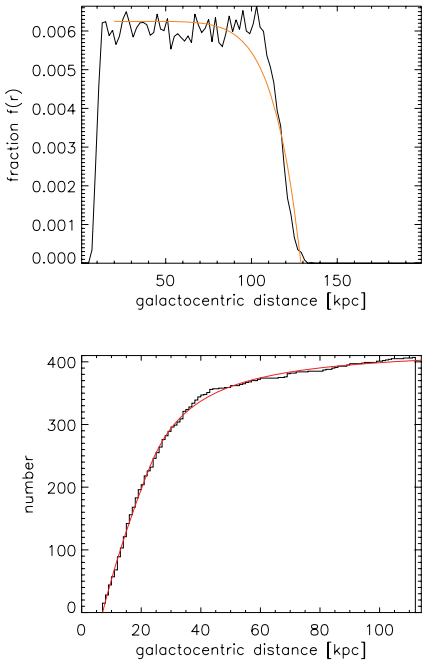

Figure 13. Top: the fraction of Galactic volume probed by the RR Lyraes in Stripe 82 as function of Galactocentric distance r. Bottom: the cumulative number of RR Lyrae within radius r in Stripe 82, together with the fits given in equation (12).

the orbit of the Sagittarius dwarf galaxy crosses Stripe 82, and there is a visible overdensity of RR Lyraes at this location ($\alpha \approx 2^{\rm h}$). Finally, we note that there are few (specifically 47) RR Lyrae at large distances ($D > 50$ kpc), of which 28 lie in a clump at $\alpha \approx 23^{\rm h}.5$. We term the structure the *Pisces Overdensity*. Distance uncertainties are shown as vertical bars for each RR Lyrae – although the error bars for the distant RR Lyrae are large enough to be visible, they cannot be responsible for the overdensity.

The upper panel of Fig. 13 shows the fraction of accessible volume as a function of Galactocentric radius r probed by our survey. The volumes are calculated via Monte Carlo integrations in which the RR Lyrae luminosity function is modelled as a Gaussian with mean and dispersion from Table 1; the magnitude limits used were those defining our RR Lyrae sample. The survey reaches at least $r \sim 100$ kpc before the accessible volume begins to decline. The brightest RR Lyraes in our sample have $M_z = 0.1$ and so are still

detectable within $r \sim 130$ kpc. In classical models of the smooth halo, the RR Lyraes are distributed as a power-law-like $\rho \sim r^{-n}$ with $n \sim 3.1$ (Wetterer & McGraw 1996). With no substructure present in Stripe 82, the right ascension–distance graph of Fig. 12 would look rather different. The fall-off in numbers would be steady, and not as sharp as the drop observed beyond $D \sim 40$ kpc, which is real and cannot be attributed to properties of the survey.

The presence of an edge to the RR Lyrae distribution in the stellar halo at $r \sim 50$ kpc has been proposed before by Ivezić et al. (2000), using a sample of 148 RR Lyraes in SDSS commissioning data. However, the same authors later applied their method to a larger area of the sky and found no break until at least 70 kpc (Ivezić et al. 2004). Vivas & Zinn (2006) also found no break before the limit of their survey at ∼60 kpc. So, 'edge' may be too strong a term, but the number density profile of the RR Lyraes does seem to be best matched by a broken power law, as shown in the lower panel of Fig. 13. Such a parametrization was first advocated by Saha (1985), who noticed that the RR Lyrae density fell off much more rapidly beyond Galactocentric radii of 25 kpc. Adjusting by the fraction of Galactic volume sampled by our survey, and assuming that our efficiency is $\epsilon \approx 1$, we find that the spherically averaged number density of RR Lyrae as

$$n(r) = 2.6 \begin{cases} \left(\frac{23\,\mathrm{kpc}}{r}\right)^{2.4} & \text{if } 5 < r \leq 23\,\mathrm{kpc}, \\ \left(\frac{23\,\mathrm{kpc}}{r}\right)^{4.5} & \text{if } 23 < r < 100\,\mathrm{kpc} \end{cases} \quad (12)$$

out to ∼100 kpc, beyond which our data are highly incomplete, with only bright RR Lyraes detectable (see Fig. 16). Our break radius of 23 kpc is very close to that proposed by Saha (1985).

The inner power-law slope is almost the same as that found by Miceli et al. (2008) – namely $n = -2.43$ – in a very large sample of RR Lyrae stars closer than 30 kpc in the Lowell Observatory Near-Earth Object Search (LONEOS) survey. Of course, formulae such as equation (12) are just a parametrization of the data, as the RR Lyrae distribution is neither spherically symmetric nor smooth, but dominated by the three structures in the Stripe. The break at $r \sim 25$ kpc is really a consequence of the fact that most of the RR Lyraes are in the Hercules–Aquila Cloud and the Sagittarius Stream substructures, which lie within 40 kpc of the Galactic Centre. A similar conclusion regarding the importance of substructure is reached by Sesar et al. (2007), who divide their RR Lyrae distribution into 13 clumps, of which they suggest at least seven correspond to real substructures.

In Fig. 14, the number density of RR Lyraes is plotted in the plane of Galactic longitude versus i-band magnitude. The three substructures show up very clearly, together with some isolated hot pixels that may be indicators of real objects. We define the Sagittarius Stream RR Lyraes via

$$180^\circ > \ell > 135^\circ, \qquad 16.5 < r < 18.5. \quad (13)$$

RR Lyraes associated with the Hercules–Aquila Cloud are extracted via

$$80^\circ > \ell > 45^\circ, \qquad 14.5 < r < 20, \qquad \ell + 15r < 358. \quad (14)$$

For the Pisces Overdensity, we chose the stars satisfying

$$95^\circ > \ell > 60^\circ, \qquad r > 19, \qquad \ell + 15r > 358. \quad (15)$$

These cuts give 55 stars in the Sagittarius Stream, 28 in the Pisces Overdensity and 237 in the Hercules–Aquila Cloud.

Ideally, we would like to separate any contaminating bulge and spheroid RR Lyrae from those of the Hercules–Aquila Cloud, but this is not easy. In particular, Fig. 15 shows the density distribution

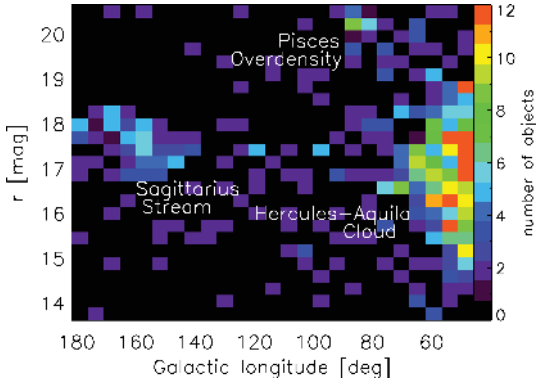

Figure 14. The number density of RR Lyraes in the plane of Galactic longitude versus r magnitude. There are three obvious overdensities corresponding to the Sagittarius Stream, the Hercules–Aquila Cloud and the Pisces Overdensity.

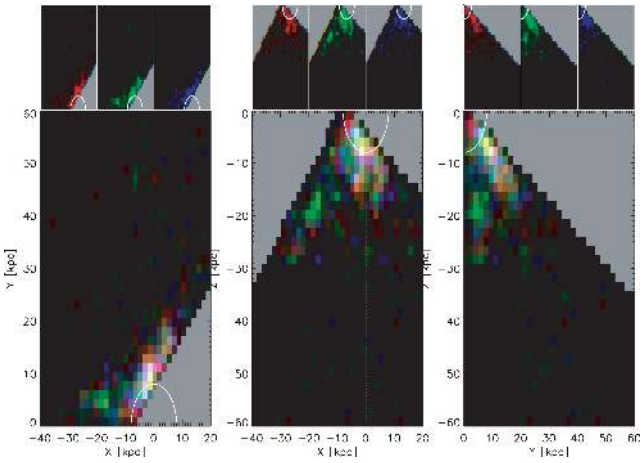

Figure 15. The density distributions of RR Lyrae candidates projected on to the principal planes of Galactocentric (x, y, z) coordinates. Here, red represents the high-metallicity RR Lyrae population ([Fe/H] > −1.3), green medium (−1.5 < [Fe/H] < −1.3) and blue low ([Fe/H] < −1.5). The insets show each distribution plotted separately, together with a white circle centred at the origin. There is no obvious distinction according to metallicity, and the low-, medium- and high-metallicity RR Lyraes are clearly distended and seemingly belong to the Hercules–Aquila Cloud. In particular, none of the populations is distributed in an spherically symmetric manner around the Galactic Centre.

of the RR Lyrae populations colour coded according to metallicity. The comparatively metal-rich RR Lyraes (red in the figure) do not seem to be distributed any differently from the comparatively metal poor (green and blue). In fact, all the distributions are distended and distributed asymmetrically relative to the Galactic Centre, consistent with the bulk of the stars belonging to the Hercules–Aquila Cloud.

The properties of the RR Lyrae in the different substructures are listed in Table 2. Note that very nearly 60 per cent of all the RR Lyraes in Stripe 82 are associated with the Hercules–Aquila Cloud, emphasising the arguments made by Belokurov et al. (2007) as to the importance of this structure. The mean heliocentric distances of the Hercules–Aquila Cloud and the Sagittarius Stream are comparable in Stripe 82, but the Pisces Overdensity is much further away at $D \sim 80$ kpc. The Pisces Overdensity lies within a few degrees of the Magellanic Plane. Although the distance of the overdensity is greater than that of the Large and Small Magellanic

Table 2. The range in Galactic coordinates and the means and dispersions in heliocentric distance, Galactocentric distance, period and metallicity of the RR Lyrae variables in the Hercules–Aquila Cloud, the Sagittarius Stream and the Pisces Overdensity.

Substructure	ℓ (°)	b (°)	D (kpc)	r (kpc)	P (d)	[Fe/H]
Hercules–Aquila Cloud	[45, 79]	[−56, −24]	21.9 ± 12.1	20.1 ± 11.3	0.51 ± 0.12	−1.43 ± 0.36
Sagittarius Stream	[139, 182]	[−62, −46]	26.1 ± 5.6	31.4 ± 5.6	0.54 ± 0.12	−1.43 ± 0.30
Pisces Overdensity	[63, 93]	[−60, −46]	79.9 ± 13.9	79.4 ± 14.1	0.56 ± 0.08	−1.48 ± 0.28

Clouds ($D \sim 55$ kpc), it is possible that the Magellanic stream may be more distant in this area of the sky. Thus, at present, it is unclear whether the Pisces Overdensity is related to Magellanic Cloud debris.

One way to estimate the mass of the Pisces Overdensity is to compare with the Carina dwarf spheroidal, which is at a similar distance (∼100 kpc; Mateo 1998). Carina has a total mass of $\sim 2 \times 10^7$ M$_\odot$ and a mass-to-light ratio of ∼70. At least 75 RR Lyrae stars have been detected by Dall'Ora et al. (2003) using well-sampled multi-epoch data in the B and V bands, although over a small baseline of a few days. Assuming similar stellar populations and similar efficiencies of detection of bright RR Lyrae in the surveys, then we can use a simple scaling argument to estimate the total stellar mass associated with the Pisces Overdensity as $\sim 10^5$ M$_\odot$. We can corroborate this mass estimate by comparison with our data on the Hercules–Aquila Cloud. The calculation using the Hercules–Aquila Cloud has the advantage that the RR Lyrae populations in the two structures have been discovered by the same algorithm, but the disadvantage that the properties of the cloud are also rather uncertain. The absolute magnitude of the cloud is given by Belokurov et al. (2007) as $M_r = -13$, suggesting that its total stellar mass is $\sim 10^7$ M$_\odot$. Of course, the cloud is an enormous structure, covering ∼80° in longitude and probably extending above and below the Galactic plane by 50°. Only a small fraction (∼1 per cent) of the cloud is probed by the Stripe 82 data, suggesting that there must be $\sim 2 \times 10^4$ RR Lyraes associated with the cloud in total. The mass of the cloud covered by the Stripe is $\sim 10^5$ M$_\odot$. Again, assuming similar stellar populations, the mass associated with the Pisces Overdensity is at least $\sim 10^4$ M$_\odot$, a value encouragingly similar to our first estimate.

Of the 28 stars identified as members of the Pisces Overdensity, the intrinsically faintest has $M_z = 0.76$. The stars extend over an area of 55 deg^2 of Stripe 82. Thus, the surface number density of the RR Lyrae is 0.51 deg^{-2}. By comparison, the Hercules–Aquila Cloud has 209 RR Lyraes with an absolute magnitude brighter than $M_z = 0.76$. They cover 95 deg^2 of Stripe 82, and hence the surface number density is 2.20 deg^2. This suggests that the Pisces Overdensity is ∼4.33 times more diffuse than the Hercules–Aquila Cloud.

The distance and metallicity distributions of our RR Lyraes can be used to study the properties of the Hercules–Aquila Cloud, the Sagittarius Stream and the Pisces Overdensity. Plots are shown in Fig. 16. The distribution of heliocentric distances for the cloud has a mean of 22.0 kpc and a standard deviation of 12.1 kpc. One possible interpretation of the cloud is that it is analogous to caustic features like the shells seen around elliptical galaxies. However, the considerable depth of the cloud seen in the RR Lyrae distribution tends to argue against such an interpretation as a caustic structure.

In directions towards Stripe 82, the distances of the arms of the Sagittarius Stream are not well known. Simulations offer a rough guide, but no more than that. The upper panel of fig. 3 of Fellhauer et al. (2006) shows the young leading arm (A), together with parts of the old trailing arm (B), at distances of 15–20 kpc; further, material belonging to parts of B, the old leading arm (C) and the young trailing arm (D) spread out over a swathe of distances 30–60 kpc. The distribution of distances of our Sagittarius RR Lyraes in Fig. 16 does indeed show some evidence of bimodality. It is possible that the peak at distances $D \sim 20$ kpc corresponds to the A and B streams, whilst the peak at $D \sim 35$ kpc corresponds to the other wraps. However, the simulations also suggest that the second peak should be much broader than appears to be the case in the data. Our identifications are tentative and radial velocity data is required to enable a cleaner separation of the wraps, as is evident from the lower panel of fig. 3 of Fellhauer et al. (2006).

From Table 2, we see that the RR Lyrae in the Hercules–Aquila Cloud have a metallicity [Fe/H] = −1.42 ± 0.24, whilst those

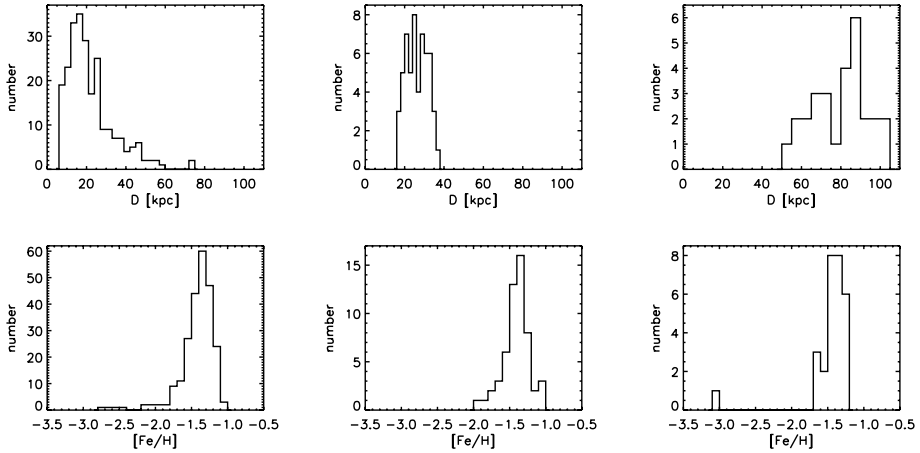

Figure 16. The distributions of distances (top) and metallicities (bottom) of RR Lyraes in the Hercules–Aquila Cloud (left) Sagittarius Stream (centre) and the Pisces Overdensity (right).

associated with the Sagittarius Stream have [Fe/H] $= -1.41 \pm 0.19$. The Pisces Overdensity has a metallicity [Fe/H] $= -1.47 \pm 0.34$, comparable to the Hercules–Aquila Cloud, but richer than the typical populations in the outer halo, which have a metallicity of [Fe/H] ~ -2 (Carollo et al. 2007).

In fact, Vivas, Zinn & Gallart (2005) have already carried out Very Large Telescope (VLT) spectroscopy of 14 RR Lyrae variables that lie in the leading arm of the Sagittarius Stream, finding a metallicity of [Fe/H] $= -1.76 \pm 0.22$. The stars lie well away from Stripe 82 at right ascensions $13^{\rm h} < \alpha < 16^{\rm h}$ and at heliocentric distances of $\sim$50 kpc. Many of the RR Lyraes in our sample will belong to the trailing arm, which may account for some of the difference in the metallicity estimate.

We show a view of Stripe 82 as derived from SDSS main-sequence turn-off (MSTO) stars in Fig. 17. The upper panel gives the number of MSTO stars as a function of right ascension and distance. The one-dimensional histogram plotted in black shows the dependence of number on right ascension alone. Note that the Sagittarius Stream is immediately visible at $\alpha \approx 40^\circ$. There are two density maxima in the black histogram, perhaps hinting that more than one wrap of the stream is detectable in MSTO stars. The Hercules–Aquila Cloud substructure also shows up very clearly, although the fainter Pisces Overdensity is understandably absent. The distance estimates to the substructures derived from MSTO stars agree well with those from RR Lyraes. The SDSS DR6 includes a large number of stellar spectra which have been analysed to provide radial velocities and fundamental stellar atmospheric parameters (Lee et al. 2008). The velocities of all stars with spectra and satisfying $g - i < 1$, to remove most of the thin disc contaminants, are plotted against right ascension along Stripe 82 in the middle panel. Of course, most of the stellar targets are disc stars, and so the curve $v_{\rm GSR} = 190 \cos b \sin \ell \,{\rm km\,s^{-1}}$ is plotted to show the locus of the thick disc in this data set. The Sagittarius Stream stars are clearly offset in velocity from the thin disc at $v_{\rm GSR} \approx -130 \,{\rm km\,s^{-1}}$. The bottom panel shows the same data, but now colour coded according to metallicity so as to highlight different structures. We can detect the kinematically-bifurcated Sagittarius Stream and clearly see the separation of the more metal-rich Galactic disc and bulge stars from the Hercules–Aquila Cloud. The eye can also discern some fainter substructure, the reality of which remains to be established.

A final view of Stripe 82 is provided in Fig. 18, which shows the density distribution of blue horizontal branch stars (BHBs) and blue stragglers (BSs), selected using the colour cuts of Yanny et al. (2000). The BHB population of course abut the RR Lyrae population in the Hertzsprung–Russel diagram. We might expect to see all three substructures in the BHB density – and so it is reassuring that the Sagittarius Stream, the Hercules–Aquila Cloud and the Pisces Overdensity are all visible. The same substructures are also identifiable in the BS populations with the exception of the Pisces Overdensity which is of course too distant. The Sagittarius Stream is clearly bifurcated in BHBs, although not in BSs, suggesting that the ratio of BHBs to BSs varies along the stream. There is some evidence for bimodality in the BHB distance distribution of the Hercules–Aquila Cloud.

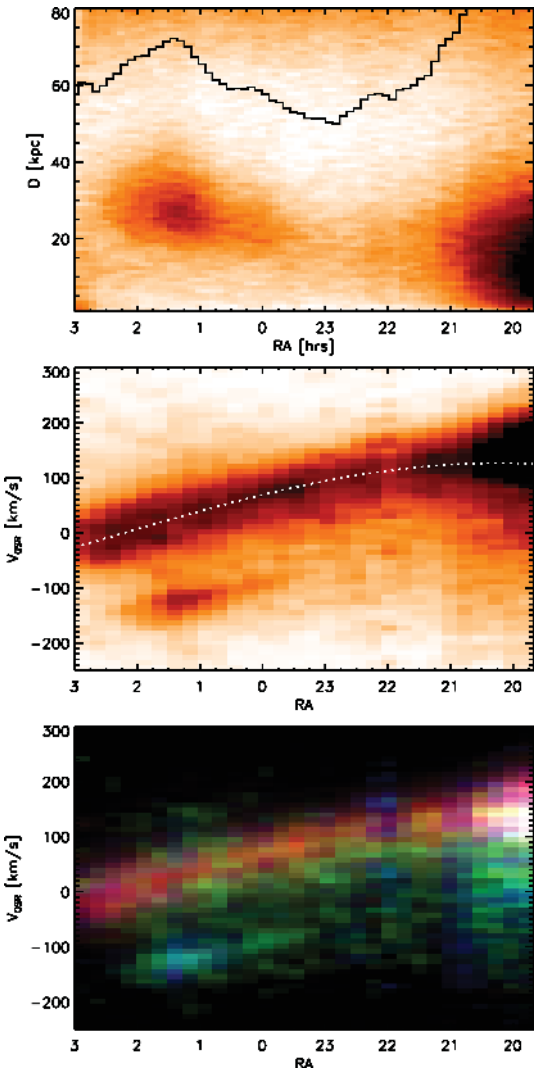

Figure 17. Top: the distributions of MSTO stars in Stripe 82 shown in the plane of right ascension versus heliocentric distance. The black histogram shows the one-dimensional distribution as a function of right ascension only. Middle: the velocity distribution of all stars with $g - i < 1$ with SDSS spectra. The white dashed curve shows the line $190 \sin \ell \cos b \,{\rm km\,s^{-1}}$ and marks the expected locus of stars belonging to the thick disc. Most thin disc stars are excised by the colour cut. Bottom: as middle, but the distribution is now colour coded according to metallicity with red representing [Fe/H] > -1, green $-1.67 <$ [Fe/H] < -1 and blue low [Fe/H] < -1.67. This separates the thin disc stars (reddish) from the older and metal-poor components of the stellar halo, such as the Hercules–Aquila Cloud.

5 CONCLUSIONS

We have constructed a catalogue of 21 939 variable objects in Stripe 82. The catalogue of variables is published in full as an electronic supplement to this paper. We have extracted a sample of RR Lyrae stars, 316 RRab types and RRc types, from the variable catalogue, using a combination of cuts based on colour, period, amplitude and metallicity. The RR Lyraes lie at distances 5–115 kpc from the Galactic Centre and individual distance estimates, accurate to typically 8 per cent, are calculated using the colour, period and metallicity to estimate absolute magnitude.

If the RR Lyrae data are modelled by a smooth density distribution, then a good fit is provided by a broken power law. The number density of RR Lyrae falls with Galactocentric radius r like $n(r) \sim r^{-2.4}$ for $5 < r < 23$ kpc, switching to a much steeper decline, $n(r) \sim r^{-4.5}$ for $23 < r < 100$ kpc. However, smooth, spherically averaged density laws do not tell the whole story, as in reality the RR Lyrae distribution is strongly clumped. In Stripe 82, the distribution of RR Lyraes is dominated by three enormous substructures

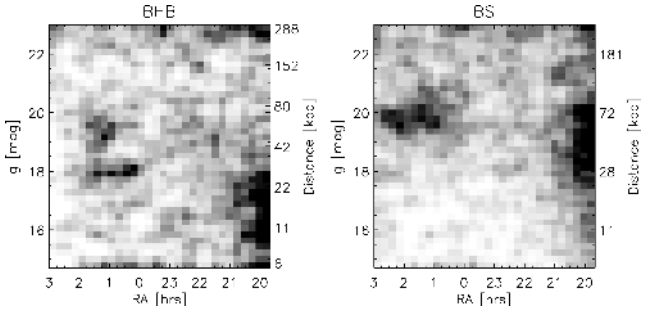

Figure 18. The distributions of BHBs (left) and BSs (right) in Stripe 82 shown in the plane of right ascension versus heliocentric distance.

– namely, the Hercules–Aquila Cloud, the Sagittarius Stream and the Pisces Overdensity. We identified samples of 237 RR Lyraes in the Hercules–Aquila Cloud, 55 stars in the Sagittarius Stream and 28 in the Pisces Overdensity.

RR Lyraes belonging to the Hercules–Aquila Cloud are very numerous, and comprise almost 60 per cent of our entire Stripe 82 sample. Although there may be some contamination from a smooth component of RR Lyraes associated with the Galactic bulge and spheroid, there is no doubt concerning the existence of the structure, supporting the initial identification of Belokurov et al. (2007). We estimate that the total number of RR Lyraes associated with the cloud is 2×10^4. The Hercules–Aquila RR Lyraes lie at distances from the Galactic Centre of 20.2 ± 11.3 kpc, and are metal poor with $[\mathrm{Fe/H}] = -1.42 \pm 0.24$.

Both leading and trailing arms of the Sagittarius Stream also intersect Stripe 82. Simulations predict that the leading wrap is closer in heliocentric distance than the trailing, but the locations of the arms are not accurately known in this region of the sky. The heliocentric distances of our Sagittarius RR Lyraes, which predominantly are associated with the trailing arm, have a mean of 26.2 kpc and a dispersion of 5.5 kpc, whilst their metallicity is $[\mathrm{Fe/H}] = -1.41 \pm 0.19$.

We have also identified a new concentration – the *Pisces Overdensity* – consisting of 28 RR Lyraes centred on Galactic coordinates of ($\ell \approx 80^\circ$, $b \approx -55^\circ$). This is one of the most distant clumps so far found in the halo, as the RR Lyrae lie at distances of $\sim$80 kpc. Although the location is close to the Magellanic Plane, the Pisces Overdensity is much more distant than the Magellanic Clouds and may well be unrelated to any known component of the Galaxy. We have made an order-of-magnitude estimate of the total mass associated with the overdensity as at least $\sim 10^4\,\mathrm{M}_\odot$. The associated RR Lyrae have a metallicity $[\mathrm{Fe/H}] = -1.47 \pm 0.34$, comparable to the Hercules–Aquila Cloud, but richer than the typical populations in the outer halo.

Our investigation has exploited the advantages of RR Lyrae stars for identifying remnants and substructure present in the halo of the Galaxy. Together with earlier SDSS discoveries (Belokurov et al. 2007; Jurić et al. 2008), the patchy and clumpy nature of the RR Lyrae distribution adds support to the picture of an outer halo composed of overdensities and voids, perhaps entirely devoid of any smooth component (e.g. Bell et al. 2008). Further study of the kinematics and metallicities of RR Lyraes in Stripe 82 should lead to a major advance in our understanding of the Galactic halo, albeit that significant observational resources will be required to acquire the necessary follow-up spectroscopy.

ACKNOWLEDGMENTS

We thank an anonymous referee for a thoughtful reading of the paper, as well as Nathan de Lee for generously communicating some of his results in advance of publication. NWE, GFG, PCH and DBZ acknowledge support from the STFC-funded Galaxy Formation and Evolution programme at the Institute of Astronomy. VB is supported by a Royal Society University research Fellowship.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS web site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, Cambridge University, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

Adelman-McCarthy J. et al., 2008, ApJS, 175, 297
Alcock C. et al., 1993, Nat, 365, 621
Alcock C. et al., 1998, ApJ, 492, 190
Aubourg E. et al., 1993, Nat, 365, 623
Bailey S. I., 1902, Ann. Harv. Coll. Obser., 38, 1
Bell E. F. et al., 2008, ApJ, 680, 295
Belokurov V. et al., 2007, ApJ, 657, L89
Bramich D. M. et al., 2008, MNRAS, 386, 887
Cáceres C., Catelan M., 2008, ApJS, 179, 242
Carollo D. et al., 2007, Nat, 450, 1020
Clark D., Stephenson F., 1977, The Historical Supernovae. Pergamon, Oxford
Collinge M. J., Sumi T., Fabrycky D., 2006, ApJ, 651, 197
Dall'Ora M. et al., 2003, AJ, 126, 197
Dilday B. et al., 2008, ApJ, 682, 262
Fellhauer M. et al., 2006, ApJ, 651, 167
Frieman J. A. et al., 2008, AJ, 135, 338
Gould A., Kollmeier J., 2004, ApJS, 152, 103
Ivezić Ž. et al., 2000, AJ, 120, 963
Ivezić Ž. et al., 2004, in Clemens D., Shah R. Y., Brainerd T., eds, ASP Conf. Ser. Vol. 317, Milky Way Surveys: The Structure and Evolution of Our Galaxy. Astron. Soc. Pac., San Francisco, p. 179
Ivezić Ž, Vivas A. K., Lupton R. H., Zinn R., 2005, AJ, 129, 1096
Ivezić Ž. et al., 2007, AJ, 134, 973
Jurcsik J., Kovacs G., 1996, A&A, 312, 111
Jurić M. et al., 2008, ApJ, 673, 864
Keller S. C., Murphy S., Prior S., DaCosta G., Schmidt B., 2008, ApJ, 678, 851
Keller S. C., Da Costa G. S., Prior S. L., MNRAS, 394, 1045
Kilić M. et al., 2006, AJ, 131, 582

Lafler J., Kinman T. D., 1965, ApJS, 11, 216
Lee Y. S. et al., 2008, AJ, 136, 2022
Mateo M. L., 1998, ARA&A, 36, 435
Miceli A. et al., 2008, ApJ, 678, 865
Monet D. G. et al., 2003, AJ, 125, 984
Morgan S. M., Wahl J. N., Wieckhorst R. M., 2007, MNRAS, 374, 1421
Munn J. A. et al., 2004, AJ, 127, 3034
Paczyński B., 2000, PASP, 112, 1281
Paczyński B., 2001, in Banday A. J., Zaroubi S., Bartelmann M., eds, Proc. MPA/ESO/MPE Workshop, Mining the Sky. Springer-Verlag, Berlin, p. 481
Perlmutter S. et al., 1999, ApJ, 517, 565
Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T., 1989, Numerical Recipes in C: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Saha A., 1985, ApJ, 289, 310
Schmidt B. P. et al., 1998, ApJ, 507, 46
Sesar B. et al., 2007, ApJ, 134, 2236
Smith H., 1995, RR Lyrae Stars. Cambridge Univ. Press, Cambridge
Smith M. et al. 2009a, MNRAS, in press (arXiv:0904.1012)
Smith M., Evans N. W., An J. H., 2009b, ApJ, 698, 1110
Stetson P. B., 1996, PASP, 108, 851
Takeda Y., Honda S., Aoki W., Takada-Hidai M., Zhao G., Chen Y.-Q., Shi J.-R., 2006, PASJ, 58, 389
Udalski A., Szymanski M., Kaluzny J., Kubiak M., Mateo M., 1992, Acta Astron., 42, 253
Vidrih S. et al., 2007, MNRAS, 382, 515
Vivas A. K., Zinn R., 2006, AJ, 132, 714
Vivas A. K. et al., 2004, AJ, 127, 1158
Vivas A. K., Zinn R., Gallart C., 2005, AJ, 129, 189
Walker A. R., Terndrup D. M., 1991, ApJ, 378, 119
Wetterer C. J., McGraw J. T., 1996, AJ, 112, 1046
Yanny B. et al., 2000, ApJ, 540, 825
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: COMPARISON WITH OTHER VARIABILITY SURVEYS IN STRIPE 82

Ivezić et al. (2007) constructed a catalogue of one million standard stars with r magnitudes 14.0–22.0, by averaging repeated observations of unresolved sources in the 290 deg^2 area of Stripe 82. Sesar et al. (2007) used the catalogue to carry out the first analysis of variability in Stripe 82. In particular, they applied cuts $\chi^2_r > 3$ and $\chi^2_g > 3$, followed by the requirement that the root-mean-square scatter exceeded 0.05 mag, to identify variability, obtaining a catalogue of 20 533 variable sources.

Bramich et al.'s (2008) LMCC is based on observations of Stripe 82 restricted to a smaller area of 249 deg^2, extending in right ascension from $20^h.7 < \alpha < 3^h.3$ with a width $2^\circ.52$ in declination from $\delta = -1^\circ.26$ to $1^\circ.26$. We extracted a sample of high-quality variable stars from the LMCC by imposing the restrictions that (i) $\chi^2_r > 3$ and $\chi^2_g > 3$, (ii) a cut on the Stetson index $L_g > 1$, (iii) at least 10 good epochs are retained, giving a catalogue of 21 939 variable objects. Applying Sesar et al.'s cuts to our catalogue gives 22 483 objects, with ≈80 per cent in common with our subsample based on Stetson index cuts. Even though Sesar et al.'s cuts give more candidates, the additional objects possess variability in different passbands that is not well correlated.

Sesar and coworkers used a colour–colour plot to discriminate between different classes of variable objects. In Fig. A1, our stellar subsample is plotted in $g - r$ versus $u - g$ and $r - i$ versus $g - r$. Here, following Sesar et al. (2007), the $g - r$ versus $u - g$ plot is divided into six regions, and labelled according to possible occupants: white dwarfs (the red-coloured, region 1), low-redshift quasars (the orange-coloured, region 2), M dwarf/white dwarf binaries (the green-coloured, region 3), RR Lyraes (the cyan-coloured, region 4), stellar locus stars (the blue-coloured,

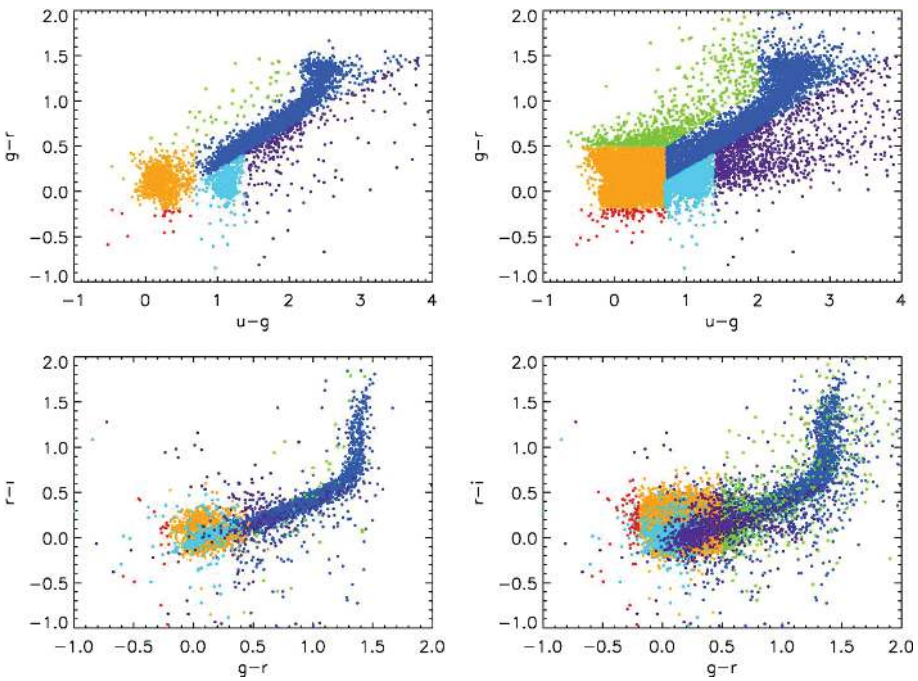

Figure A1. Colour–colour plots for 4648 variables brighter than $g = 19.0$ (left) and 21 789 variables brighter than $g = 22.0$ (right). The upper panels are $g - r$ versus $u - g$, the lower panels are $r - i$ versus $g - r$. Sesar et al. (2007) label these regions as white dwarfs (region 1, red), low-redshift quasars (region 2, orange), M dwarf/white dwarf binaries (region 3, green), RR Lyraes (region 4, cyan), main stellar locus (region 5, blue) and high-redshift quasars (region 6, purple).

Table A1. The distribution of candidate variable sources in the $g - r$ versus $u - g$ diagram. The columns list the fraction of the whole sample and the variable subsample lying in the six regions of the colour–colour plot.

		$g < 19$		$g < 20.5$		$g < 22$	
Region	Name	Per cent all	Per cent var	Per cent all	Per cent var	Per cent all	Per cent var
1 (red)	White dwarfs	0.09	0.43	0.16	0.33	0.20	0.36
2 (orange)	Low-redshift quasars	0.29	21.32	1.20	52.96	5.31	61.05
3 (green)	dM/WD pairs	5.95	1.01	9.51	1.59	11.77	3.65
4 (cyan)	RR Lyrae stars	3.27	15.23	3.62	9.31	3.47	7.85
5 (blue)	Stellar locus stars	75.93	56.41	74.62	31.44	69.65	22.48
6 (purple)	High-redshift quasars	14.47	5.59	10.89	4.37	9.61	4.62
	Total count	283 899	4648	447 800	12 788	518 058	21 789

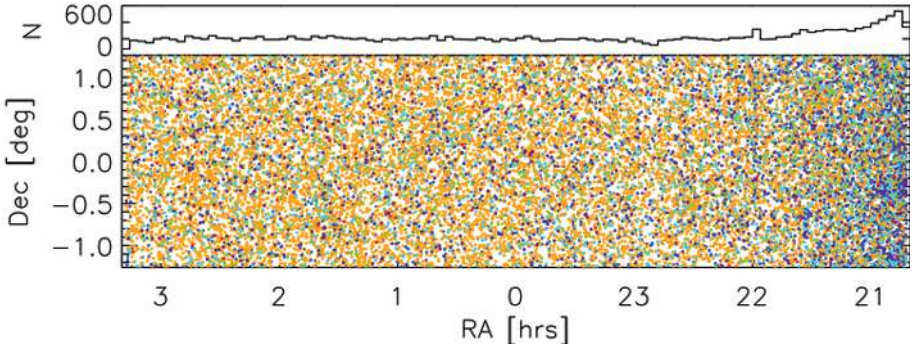

Figure A2. The spatial distribution of the variable subsample in Stripe 82. Objects are colour coded according to the regions of the colour–colour plot in which they lie (see Fig. A1). The upper panel shows the number of all the variable objects versus right ascension.

region 5) and high-redshift quasars (the purple-coloured, region 6). The colour–space divisions provide only very rough classifications. In some cases (such as region 1), the label does not even describe the typical population, and we merely use the labels as a point of comparison to Sesar's work.

The percentages of the variable subsample and the whole sample lying in the regions of the colour–colour plots are given in Table A1, whilst sample light curves have already been shown in Fig. 3. Almost all (>93 per cent) of the variable objects lie in three regions – namely, low-redshift quasars (53 per cent of the catalogue), stellar locus stars (31.4 per cent) and RR Lyrae stars (9.3 per cent). When split according to magnitude, the bright ($g < 19.0$) variable sky is dominated by stellar locus stars, but the faint ($g < 22.0$) variable sky is dominated by quasars. We can compare our results to table 1 of Sesar et al. (2007), which shows the same quantities for their variable subsample. Our variability criteria pick out more variable objects, and in particular more denizens of the main stellar locus.

The spatial distribution of variable objects in Stripe 82 is shown in Fig. A2. The equatorial stripe reaches down to low Galactic latitudes beyond $\alpha \approx 18^{\rm h}$ (see e.g. fig. 1 of Belokurov et al. 2007). Variables belonging to region 4 (RR Lyraes) and region 5 (the main stellar locus) dominate here, whereas variables belonging to the other regions are more uniformly dispersed in right ascension.

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