JOURNAL OF EMERGING INVESTIGATORS

Analyzing the Membership Ratio of RR Lyrae Variables in Globular Clusters with GAIA Data

Cheng-You Ho[1], Kuan-Fu Huang[1]

[1] Chiayi Senior High School, Chiayi City, Taiwan R.O.C.

Student Authors
Cheng-You Ho: High School

SUMMARY

Globular clusters are dense, low-metallicity star clusters. These objects are home to most RR Lyrae Variables, which are important in determining the evolution of star clusters. Although considerable research has been devoted discussing their period and luminosities, the membership ratio of RR Lyrae stars within globular clusters remains an area of investigation. Using the GAIA Data Release 3 database, the apparent dimensions of the cluster and an estimated number of observed stars in the cluster are calculated. Then, the membership ratio of RR Lyrae stars in 20 globular clusters are analyzed. Quantaties combining the effect of the three parameters are defined. It is concluded that three factors play an essential role in deciding the ratio of RR Lyrae Variables to the number of stars in the home cluster, which are the cluster age, cluster radius and metallicity, listed in order of descending impact. Age and metallicity are negatively correlated with the number density of RR Lyrae variables, while radius is positively correlated. It is hypothesized that the formation of RR Lyrae stars be correlated with tidal forces and properties of the home galaxy. The special cases that deviate from the main trend are discussed.

INTRODUCTION

Star clusters are objects composed of large numbers of stars. They can be classified into globular clusters (GCs) and open clusters, where the former are more densely packed, more easily found near the Milky Way halo, and has a lower metallicity **(1)**. Since globular clusters are formed at the early ages of the universe, research into them can tell us a lot about the evolution of the universe. It has been pointed out that the formation of star clusters is mostly influenced by its mass, with metallicity being the secondary factor **(2)**. However, due to the highly complex chemical evolutions, in addition to processes such as nucleosynthesis and mass-loss, globular clusters still remain an open area of investigation.

Variable stars are stars whose brightness changes over time. RR Lyrae variable stars (RRLs) often appear in globular clusters, and they play a key role in determining the evolution of star clusters. On a Hertzsprung-Russell diagram (HRD), RRLs are located on the instability strip, with a mass of 1 to 2 solar masses. Classification of RRLs based of brightness curves and pulsation periods have been investigated, and it is further known that there is a correlation between metallicity and mass **(3, 4)**.

Due to limited accuracy of past apparatus, the number of stars that could be observed is limited, leaving statistical analyses of variable stars in star clusters an open area of research. In June 2022, the European Space Agency (ESA) released its third data release of the GAIA mission, which boasts photometric data for over 1.8 billion stars **(5)**. Using the newly-released GAIA database, it is possible to calculate the ratio of RRLs in globular clusters. By analyzing such ratio with different cluster parameters, we found that three factors play an essential role in deciding the ratio of RR Lyrae Variables to the number of stars in the home cluster, which are the cluster age, cluster radius and metallicity, listed in order of descending impact. We hypothesized that the formation of RR Lyrae stars be correlated with tidal forces and properties of the home galaxy.

RESULTS

We selected 20 globular clusters mainly from the Messier catalog. For a target cluster, a query of the GAIA database is conducted to find the number of stars around a cluster is conducted, yielding the number of total membership stars for the cluster, along with the dimensions of the cluster. The number of RR Lyrae stars in the cluster is found with another GAIA query. Dividing the latter by the former gives the membership ratio of RR Lyrae stars in globular clusters. This number is expressed in part per million, which will be later referred to as RRppm. Fitting the isochrone curve to the color-magnitude diagram filters out non-member stars.

The RRppm value of the 20 globuar clusters selected ranges from 123 to 7544 **(Figure 1).** Comparing RRppm to various cluster parameters, such as age, metallicity, radius and mass, the relationship between such parameters and the membership ratio of RR Lyrae stars can be deduced **(Table 1).** Plots of RRppm with age, metallicity, radius, mass, right ascension and declination shows that RRppm has negative correlations with age and metallicity, but is positively correlated with radius, and shows no significant relationship with mass **(Figure 2)**.

To find a power law relationship between RRppm and various parameters, some variables combining various parameters are defined. The units of age is billion years (byr), the unit of radius is light years, and metallicity is represented with [Fe/H]. To combine the effects of age and metallicity, the $\mathrm{X_{RR}}$ variable is defined as

$$\mathrm{X_{RR}} = \mathrm{a} \cdot \log(\mathrm{Age}) + \mathrm{b} \cdot [\mathrm{Fe/H}]$$

Where a and b are free variables. Using a python program to calculate the $\mathrm{X_{RR}}$ value for each of the 20 clusters with a given a and b, the optimal a and b that yields the highest correlation coefficient with RRppm is determined to be $\mathrm{a} = -6.264$ and $\mathrm{b} = -0.374$. The correlation coefficient of $\mathrm{X_{RR}}$ and RRppm, using the optimal parameters, is -0.501, with a p-value of 0.0244. In a similar manner, the effect of radius could be added by defining a new variable as

$$\mathrm{Y_{RR}} = \mathrm{a} \cdot \log(\mathrm{Age}) + \mathrm{b} \cdot [\mathrm{Fe/H}] + \mathrm{c} \cdot \mathrm{Radius}$$

The values of $\mathrm{a}, \mathrm{b}, \mathrm{c}$ that gives the highest correlation coefficient between $\mathrm{Y_{RR}}$ and RRppm are -71.096, 11.110 and -2.102, with a correlation coefficient of 0.613 and a p-value of 0.004. Most of the 20 data points lie in the 95% confidence interval **(Figure 3)**.

DISCUSSION

The dimensions of the globular clusters are determined so that it is the minimum side length for a square search area that yields almost the same number of stars in the bordering 8 cells of the 3 by 3 heatmap. Instruments with a better resolution would be able to see more stars, hence the dimension it determines would be larger than that determined with a telescope of lower resolution. Most of the dimensions determined in this research are larger than that from the references, shedding light on GAIA's ability to see stars up to a magnitude of 20.7 **(5)**.

Plots of RRppm vs. various cluster parameters show the positive or negative correlations of RRppm with cluster age, metallicity and radius, although the trend is insignificant. Therefore, analyses using the $\mathrm{X_{RR}}$ and $\mathrm{Y_{RR}}$ variables are required. It is shown in the plots that the linear fit is better than simply plotting RRppm vs. individual cluster parameters **(Figure 2, Figure 3)**. The p-values of 0.0244 and 0.004 signify that the null hypothesis could be rejected and the cluster parameters do influence RRppm.

The $\mathrm{Y_{RR}}$ variable has a postive correlation with RRppm, and the coefficients of metallicity and age are negative, while the coefficient of radius is positive. This confirms that metallicity and age are negatively correlated with RRppm while radius is positively correlated, agreeing with the result obtained previously. From the coefficients of various parameters and their average of the 20 globular clusters, it is possible to multiply the absolute value of the coefficient and the average value of a cluster parameter to compare the relative significance of age, metallicity and radius (**Table 2)**. It is found that the importance of these three factors, listed in decreasing order, are age, radius and metallicity.

As stated previously, age and metallicity are the two primary factors governing the evolution of star clusters **(2)**. RR Lyrae stars in older globular clusters may have left the instability strip and moved further into stellar evolution, explaning why RRppm is lower at larger ages. The negative correlation of metallicity with RRppm may be related to how RR Lyrae stars form.

In order to explain why there is likely to be more RR Lyrae stars (higher RRppm) when the radius of the cluster is larger, we proposed a hypotheses. We propose that the formation of RR Lyrae stars, which is still in debate, could be related to tidal forces, which is proportional to the radius of the cluster. This hypothesis is in agreement with past research suggesting that the formation of RR Lyrae stars is correlated with mass loss **(6)**. However, the exact mechanism of RRL formation requires future investigation.

It is found that two special cases, the clusters M54 and M10, have significantly high and low RRppm values, respectively **(Figure 1)**. Since M54 belongs to the Sagittarius Dwarf Galaxy while others are in the Milky Way, we propose that RRppm may be dependent on properties of the home galaxy. Moreover, the lack of RR Lyrae stars in M10 is in agreement with past research, which suggests that this may have to do with variations on the Red Giant Branch evolution of M10. **(6, 7)**.

In summary, we have found that cluster age, radius and metallicity are the three primary factors influencing the number density of RR Lyrae stars in globular clusters, listed in order of decreasing impact. Age and metallicity are negatively correlated with RRppm, while radius is positively correlated. We propose that the formation of RR Lyrae stars may be dependent on the radius of the cluster and properties of the home galaxy.

MATERIALS AND METHODS

We selected 20 globular clusters, 18 of which have a designation in the Messier catalog with data collected from various sources. **(8, 9, 10, 11)**. 2 other clusters, namely Omega Centauri (NGC 5139) and NGC 6441 are selected. The age, dimension, radius, metallicity, right ascension and declination for all the GCs in this research are given from the references.

For every target cluster, a search around the cluster is first conducted. The region around the cluster is divided into a 3 by 3 square grid (heatmap), where the center of the cluster lies at the middle of the center cell. The number of stars in each grid, retrieved with Python from the GAIA database, is calculated. By subtracting the number of stars in the center cell by the average of that in the 8 surrounding cells, the number of stars in the center cell, filtering out foreground and background stars, could be estimated, assuming that foreground and background stars are homogeneous throughout space. The side length of the 3 by 3 grid is then continuously increased, while calculating the standard deviation of the number of stars in the bordering 8 cells. At the side length which yields the lowest standard deviation, it can be said that the center cell has covered the globular cluster entirely, so such side length is taken as the dimension of the globular cluster **(Figure 4)**.

Python programs are ran on Google Colaboratory. The difference of the visual blue and red bands given by the GAIA database is recorded as (b-r). The isochrone curve from the “isochrone.mist” database is plotted onto the color-magnitude diagram (CMD) of a given globular cluster, where the x-axis is the effective temperature determined with the equation

$$\mathrm{T_{eff}} = 10^{3.939-0.395\times(\mathrm{b}-\mathrm{r})}$$

, and the y-axis is the visual magnitude of the red band **(12)**. The isochrone curve is manually fitted to the CMD, and stars that deviate from the isochrone are removed. After filtering out stars that do not belong to the target cluster, RRppm can be calculated.

Finding optimal coefficients for $\mathrm{X_{RR}}$ and $\mathrm{Y_{RR}}$ is done with Python with 2 and 3 nested loops respectively to find the coefficients to 4 significant decimal places.

REFERENCES

1. Archinal, Brent A., and Steven J. Hynes. “Star Clusters.” *Willmann-Bell*, 2003.
2. Gratton, R., et al. "What is a Globular Cluster? An Observational Perspective." *The Astronomy and Astrophysics Review*, vol. 27, Dec. 2019. https://doi.org/10.1007/s00159-019-0119-3
3. Smith, Horace A. “RR Lyrae Stars”. *Cambridge University Press*, 2004.
4. Stobie, R. "On the Difference Between the Oosterhoff Types I and II Globular Clusters." *Astrophysical Journal*, vol. 168, 1971.
5. Prusti, T., et al. "The Gaia mission." *Astronomy & Astrophysics*, vol. 595, 2016. https://doi.org/10.1051/0004-6361/201629272
6. Ferro, A. A., et al. "The globular cluster M10: reassessment of stellar membership, distance, and age using its variable and HB stars." *Monthly Notices of the Royal Astronomical Society*, vol. 499, no. 3, 2020. https://doi.org/10.1093/mnras/staa2977
7. Braun, K., et al. "Photometry Results for Globular Clusters M10 and M12." *The American Astronomical Society*, vol. 124, no. 4, 2002. https://doi.org/10.1086/342851
8. Goldsbury, Ryan, et al. "The ACS Survey of Galactic Globular Clusters. X. New Determinations of Centers for 65 Clusters." *The Astronomical Journal*, vol. 140, no. 6, 2010. https://doi.org/10.1088/0004-6256/140/6/1830

9. Boyles, J., et al. "Young Radio Pulsars in Galactic Globular Clusters." *The Astrophysical Journal*, vol. 742, no. 1, 2011. https://doi.org/10.1088/0004-637X/742/1/51
10. Koleva, M., et al. "Spectroscopic ages and metallicities of stellar populations: validation of full spectrum fitting." *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 4, 2008. https://doi.org/10.48550/arXiv.0801.0871
11. Marín-Franch, Antonio, et al. "The ACS Survey of Galactic Globular Clusters. VII. Relative Ages." *The Astrophysical Journal*, vol. 694, no. 2, 2009. https://doi.org/10.1088/0004-637X/694/2/1498
12. Casagrande, L., et al. "An absolutely calibrated Teff scale from the Infrared Flux Method: Dwarfs and subgiants." *Astronomy & Astrophysics*, vol. 512, Apr. 2010. https://doi.org/10.1051/0004-6361/200913204

Figures and Figure Captions

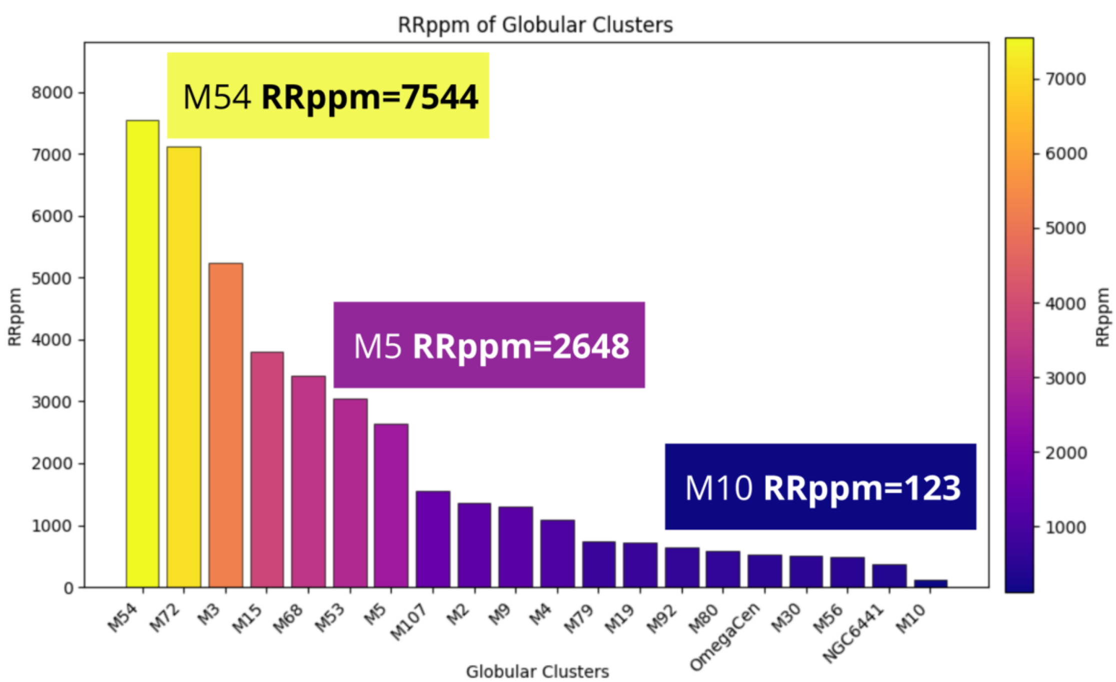

Figure 1. RRppm value of 20 globular clusters. Bar graph lists the part per million membership ratio of RR Lyrae stars in decreasing order. After a query to the GAIA database retrieves the position of individual stars, foreground and background stars are filtered, and RRppm can be calculated.

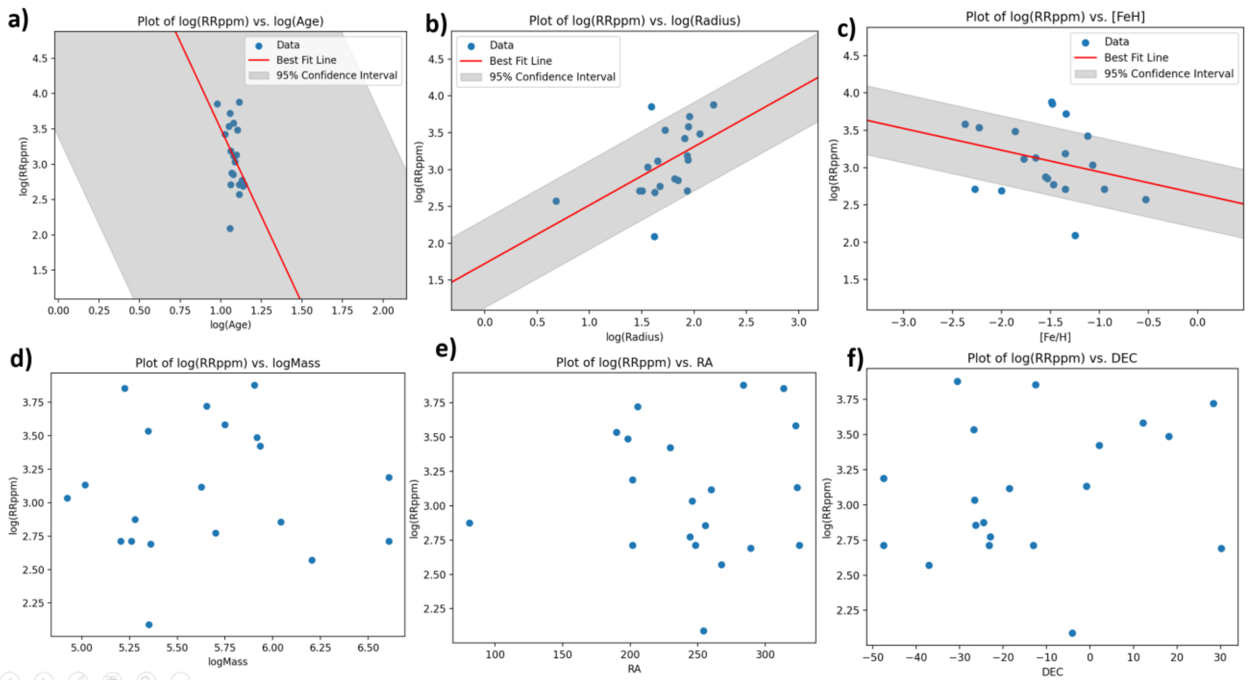

Figure 2. Plots of RRppm versus various cluster parameters. The 20 blue dots shown in each subplot each represent a globular cluster. a) RRppm vs. cluster age shows negative correlation. b) RRppm vs. cluster metallicity also shows negative correlation. c) RRppm vs. cluster radius shows positive correlation. d) RRppm vs. cluster mass. e) RRppm vs. cluster right ascension. f) RRppm vs. cluster declination. Plots d), e) and f) do not present any significant correlation between RRppm and the respective cluster parameters. For the three parameter with significant trends, the best fit line shown in red and the gray 95% confidence interval are superimposed. Data from the Messier catalog and references are plotted using Python, with the color bar indicating metallicity.

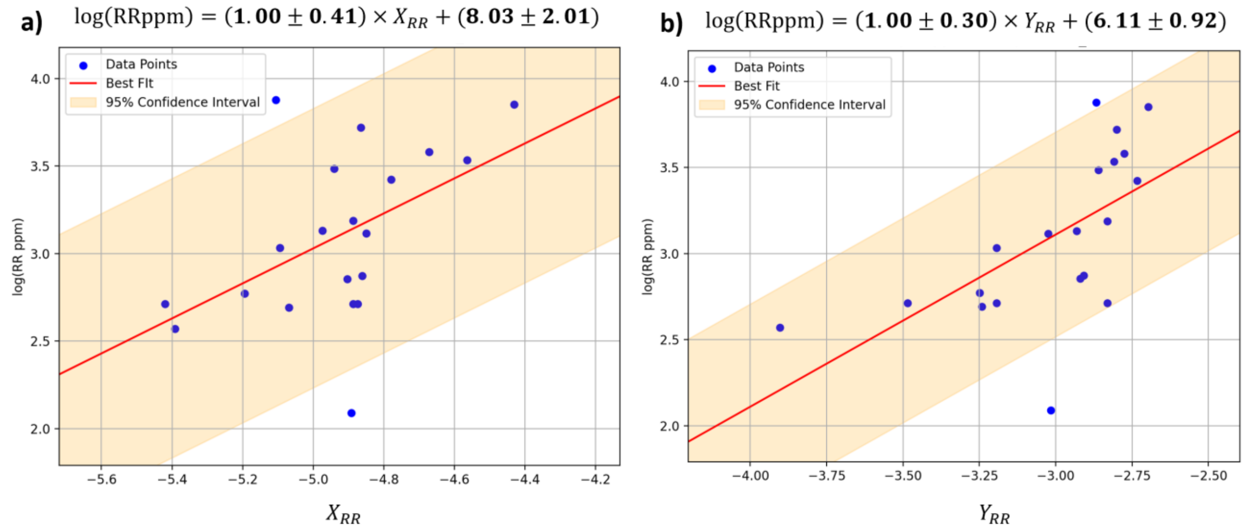

Figure 3. Plots of log(RRppm) versus X_{RR} and Y_{RR}. The 20 blue dots shown each represent a globular cluster. The red line is the regression line with its equation shown on top of each subfigure, and the orange-shaded region is the 95% confidence interval. a) log(RRppm) vs. X_{RR} (p=0.0244). b) log(RRppm) vs. Y_{RR} (p=0.004). Both results suggest that the null hypothesis be rejected. X_{RR} and Y_{RR} are two variables combining the effects of multiple cluster parameters defined previously. Regression and plotting are made using Python.

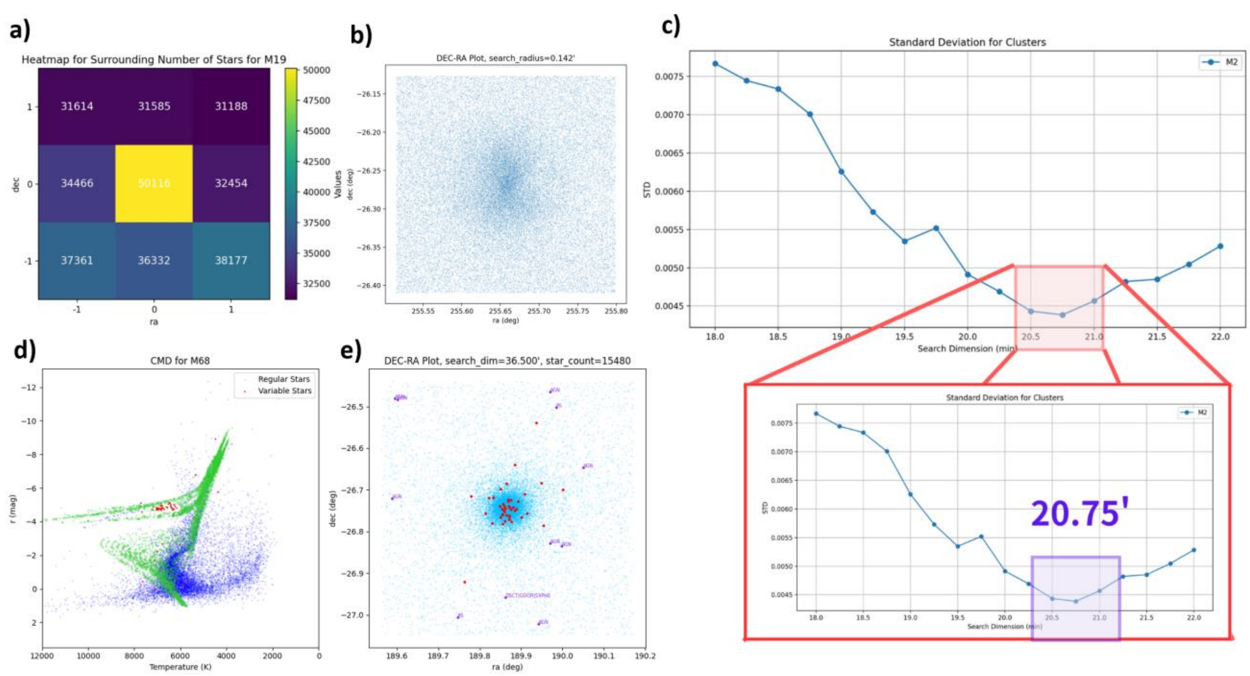

Figure 4. Methodology of finding RRppm. a) The plot of declination vs. right ascension of the cluster M19 depicts what the cluster would look like to a normal eye, with the center of the

cluster significantly denser. b) The plot labeled a) is divided into a 3 by 3 grid and the number of stars in each grid is counted with Python, with the center cell having significantly more stars. c) The dimension of the 3 by 3 grid that yields the lowest standard deviation of the number of stars in the bordering 8 cells is determined with Python. Here data of the cluster M2 is used instead for easier demonstration. d) From the GAIA database, the color magnitude diagram can be plotted, with the blue dots representing normal stars, red stars representing variable stars and the green dots being the fitted isochrone curve. e) RR Lyrae stars, colored in red, and other types of variable stars, colored in purple, can be superimposed onto the plot of declination vs. right ascension with GAIA data.

Tables with Captions

Globular Cluster	Age (byr)	Fe/H	Radius (light years)	Mass (solar mass)	RRppm
M2	12.5	-1.65	87.3	104000	1353.003
M3	11.39	-1.34	90	450000	5243.486
M4	12.2	-1.07	36	84000	1079.763
M5	10.62	-1.12	81	857000	2647.619
M9	12	-1.77	45	422000	1307.190
M10	11.39	-1.25	41.6	225000	122.884
M15	12	-2.37	88	560000	3806.132
M19	11.9	-1.53	70	1100000	714.745
M30	13	-2.27	32	160000	514.304
M53	12.67	-1.86	113	826000	3054.152
M54	13	-1.49	153	800000	7544.141
M56	13.7	-2	42	230000	490.196
M68	11.2	-2.23	53	223000	3423.839
M72	9.5	-1.48	39	168000	7126.314
M79	11.7	-1.55	65	190000	746.130
M80	13.5	-1.47	47	502000	592.467
M92	14.2	-2.32	109	200000	633.767
M107	13.95	-0.95	30	182000	1544.667

OmegaCen	11.52	-1.35	86	4050000	516.193
NGC6441	13	-0.53	4.8	1600000	371.747

Table 1. Age, metallicity, radius, mass and calcualted RRppm of the 20 researched globular clusters. Cluster parameters are collected from the Messier catalog and references. RRppm, the part per million membership ratio of RR Lyrae stars in each globular cluster, is calculated using Excel.

Parameter	Value Range	Coefficient in Y_{RR}	Relative Significance
log (Age (Gyr))	0.978~1.152	-71.096	~1
log (Radius (ly))	0.681~2.185	11.110	~0.22
Metallicity ([Fe/H])	-2.37~-0.95	-2.102	~0.05

Table 2. Estimation of relative significance of cluster age, radius and metallicity. The absolute value of the coefficient of each parameter in Y_{RR} is multiplied by a median value of the 20 globular clusters. The product is normalized so that the most significant parameter, which is age, has a relative significance of 1. The calculation shown here strengthens the fact that age, radius and metallicity are three significant factors influencing RRppm listed in decreasing order of impact. The calculation is made with Excel.