On the Evolutionary Connections among OB stars, Wolf – Rayet Stars, Yellow Supergiants and Red Supergiants in the M33 Galaxy

Yordan Darakchiev* and Petko Nedialkov

[1]Department of Astronomy, Sofia University, 5 James Bourchier St., 1164 Sofia, Bulgaria

Abstract. We applied a variety of statistical tests to verify the evolutionary connections between different evolutionary stages of supergiant stars – OB, YSG and RSG on one hand, and their evolved counterparts on the other – WR-stars, Cepheids and supernova remnants. To do so, we compared the spatial distributions in the M33 galaxy among all these types of stars.

We found that there are significant (at 98%) correlations on the scale of 0.7 to 4.75 kpc between several pairs, most notably WR – BSG, BSG – YSG, and pairs involving supernova remnants, which vanish on the galactic scale. We propose that this effect is likely due to a clumpy star formation structure.

Keywords: galaxies: stellar content, galaxies: Local Group, stars: evolution, stars: massive, methods: statistical

1. INTRODUCTION

Massive stars provide key challenges to theoretical models and observations alike. Despite their smaller number relative to lower-mass stars, they greatly influence their environment. These objects release a lot of energy and material via radiation, winds and eventually via explosions as core collapse supernovae. Thus, massive stars facilitate star formation by triggering star formation and enriching their environment with metals (Massey, 2003).

Present-day theoretical models include a variety of parameters which influence stellar evolution. Even though there are many complex scenarios (Maeder, 1996; Massey & Johnson, 1998), they all agree that the mass-loss rate plays a critical role (Maund et al, 2017).

Stellar populations provide a snapshot of the various stages that massive stars can go through. Observational data constrain the importance of factors such as metallicity and binarity, validate the theoretical models and provide constraints for their further development. This is especially true for galaxies in the Local Group where individual stars can be more easily resolved (Maund et al, 2017; Massey et al, 2016).

To explore the evolutionary connections among various evolutionary stages of massive stars, this article poses the question how the spatial distributions of evolved objects (Wolf – Rayet stars, supernovae remnants; WR and SNR, hereafter) are related to the spatial distributions of their possible progenitors – blue and red supergiants (BSG and RSG, respectively). In particular, we are interested in the spatial distributions of Cepheid (CEP) variables and their progenitors – yellow supergiants (YSGs).

We follow in the footsteps of Georgiev & Ivanov (1997) but we consider more classes of massive stars, including SNRs – their non-stellar evolved counterparts. This allows us to compare some *suspected* progenitor – evolved object pairs (for example, BSG – WR, RSG – SNR) with a secure "benchmark" pair YSG – CEP. Furthermore, our sample is significantly larger and we apply a more robust ensemble of statistical tests in our analysis.

2. OBSERVATIONAL DATA

We used three different data sources. First, the Local Group Galaxy Survey (LGGS; Massey et al. 2016) contains a sample of ~150 000 sources; 2290 of them with known spectral types. We selected:

– BSGs: O and B types, luminosity class I; 431 objects.

– YSGs: F and G types, luminosity class I, including all stars labelled "YSG" explicitly; 156 objects.

– RSGs: K and M types, luminosity class I, including all stars labelled "RSG" explicitly; 237 objects.

* Corresponding author's email: jdarakchie@uni-sofia.bg

– WR: WC and WN types. The LGGS does not contain any WOs; 197 objects.

The CEPs were selected from the synoptic survey of Pellerin & Macri (2011). It consists of BVI observations of ~793 000 stars, 663 of them confirmed CEPs. We verified them (even the initially rejected ones) with multi-epoch observations obtained from 2012 to 2015 with VST OmegaCam (Kuijken et al, 2011). We matched objects using nearest neighbors with max matching radius of 0.8". Our final sample contains 607 confirmed CEPs with periods determined by the Lomb – Scargle method.

Finally, we used the *XMM – Newton* supernova remnant survey (Garofali et al, 2017) that contains X-ray observations of SNRs in M33. It consists of deep observations on an eight-field mosaic of M33 with a total exposure of 900 ks and yielded 105 SNRs confirmed at the 3σ-level.

3. COORDINATE DEPROJECTION

To compare the physical distances within M33, we transformed the equatorial coordinates (α,δ) to orthogonal coordinates (x,y) in kpc, relative to the center of the galaxy. We assume that M33 is a circular disk with the following geometric parameters:

– Coordinates of the center: $\alpha_C = 1^h 33^m 50.904^s$; $\alpha_c = 1^h\, 33^m\, 50.904^s; \delta_c = +30°\, 39'35.79''$ (Skrutskie et al. 2006).
 – Position angle: $P = 23.0°$ (Gil de Pas et al. 2007).
 – Inclination angle: $i = 51.0°$ (de Vaucouleurs 1959).
 – Distance: $d = 840$ kpc (Freedman et al. 1991).

We then converted the rectangular coordinates to galactocentric distances using:

$$r_{gal} = \sqrt{x^2 + y^2} \tag{1}$$

4. DISTRIBUTION FUNCTION

Similar to Georgiev & Ivanov (1997), we will compare the *normalized cumulative distributions* of galactocentric distances:

$$P(r) = N(r)/N_{total} \tag{2}$$

that gives the probability of finding a star from the specified population within a circle with radius r. Here, $N(r)$ is the number of stars from a given population with $r_{gal} \le r$, and N_{total} is the total number of stars in that population.

Figure 1 displays the resulting function for all classes. The maximum distance differs among the classes: for example, there are no WR stars in the outermost regions. However, we found that excluding the outermost objects to compare the distributions over the same spatial region has no effect on our conclusions.

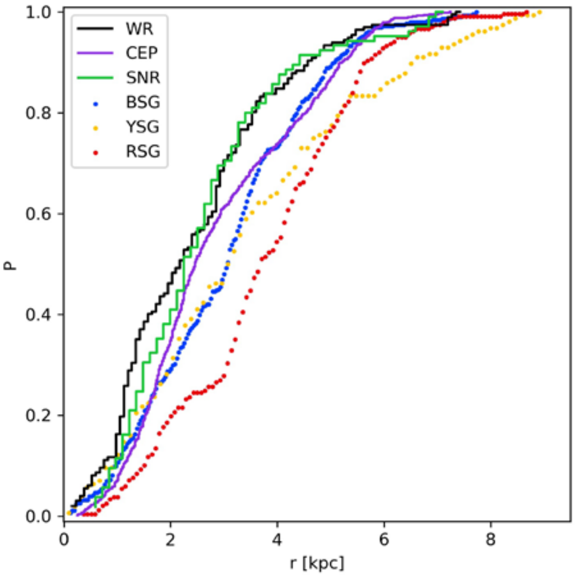

Fig. 1 Distribution functions $P(r)$ for all classes in our sample.

5. STATISTICAL TESTS

First, we formulated the following hypotheses:

H_0: The observed samples come from the same distribution

H_1: The observed samples do not come from the same distribution

Then we compared the spatial distributions, applying on each pair of stellar populations three statistical tests, for robustness. Hereafter we adopted a significance level limit of $\alpha = 0.98$.

5.1. Kolmogorov – Smirnov Test

The two-sample Kolmogorov – Smirnov Test (KST) is a non-parametric hypothesis test which measures the probability that two univariate samples are drawn from the same distribution. It is commonly used in astronomy because it does not depend on the distribution (as long as it is continuous), there are no sample size restrictions and the result is easy to interpret.

The test statistic is defined as

$$D = \max|F_1(x) - F_2(x)| \quad (3)$$

where $F_1(x)$ and $F_2(x)$ are the empirical cumulative distribution functions, with sample sizes n and m respectively.

The null hypothesis is rejected at confidence level α if

$$D > c(\alpha)\sqrt{\frac{n+m}{nm}} \quad (4)$$

where $c(\alpha) = \sqrt{-0.5\ln(\alpha)}$.

Figure 2 displays the obtained p-values. The KST rejects the null hypothesis at the 98% level. These results confirm the initial findings of Georgiev & Ivanov (1997) that the galactic distribution of evolved objects does not simply follow the distribution of their predecessors. In addition, we found that this applies to more pairs of object classes (the original article focuses solely on WR – BSG, and WR – (BSG + RSG) pairs.

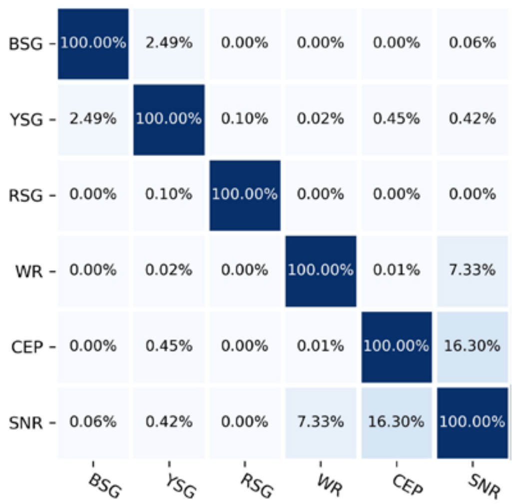

Fig. 2 The KST p-values for all pairs of samples.

5.2. Mann – Whitney *U* Test

The KST has some shortcomings. Most of all, it is highly sensitive to any differences in the distributions. The Mann – Whitney U test (MWUT) is more robust, being sensitive mostly to changes in the sample's median values. Similar to KST, it is also non-parametric. It applies easily to discrete distributions, however this test should only be used when the size for each sample is sufficiently large: 20 objects or more (Harding 1983).

The MWUT statistic is obtained by ranking all data points (for example, smallest to largest). For the two datasets, the sums of ranks R_1, R_2 are computed:

$$U_1 = R_1 - \frac{n(n+1)}{2} \quad (5)$$

$$U_2 = R_2 - \frac{m(m+1)}{2} \quad (6)$$

where n and m are the sample sizes. The test statistic is then given by

$$U = \min(U_1, U_2) \quad (7)$$

U has a known distribution but is commonly approximated by a normal distribution for large enough samples (Harding 1983).

Figure 3 shows the MWUT p-values. Again, the null hypothesis is rejected at the 98% level, but

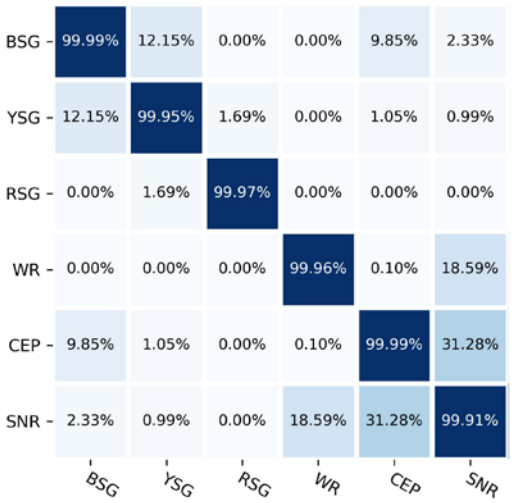

Fig. 3 The MWUT p-values for all pairs of samples.

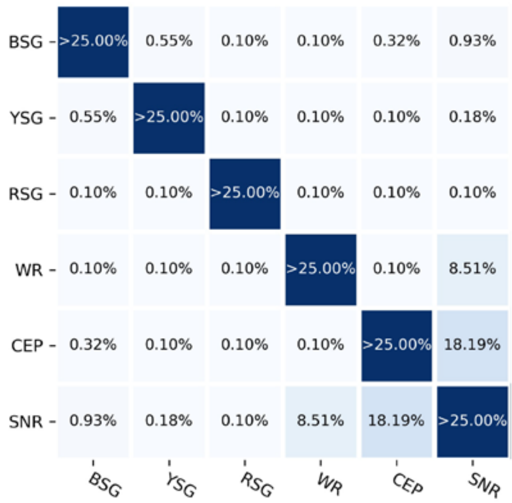

Fig. 4 The ADT p-values for all pairs of samples. The p-values are capped at 25% but none of the samples, except the identical pairs, is close to this value.

this test provides an insight into some correlations. For example, the highest p-value corresponds to the pair SNRs – CEPs. We are going to discuss these in the next section.

5.3. Anderson – Darling Test

The multi-sample Anderson – Daring test (ADT; Scholz & Stephens 1987) is a modification of the KST. While the KST is more sensitive to the center of the distribution, the ADT gives more weight to the tails, therefore it has better statistical power for distributions with strong tails. It is non-parametric as well.

The ADT statistic is based on the sum of squared differences between the empirical sample distribution functions and that of the combined sample. Because of the underlying distribution, this test is capped at 25% significance (e.g. if two distributions yield $p > 25\%$, we are not able to see the exact percentage). Nevertheless, we considered the ADT because other tests reject the null hypothesis at very low p-values. The ADT provides an additional consistency check.

Figure 4 shows the ADT p-values. The results agree with the other tests. There are indications for some correlations, e.g. for SNRs – CEPs and for SNRs – WRs.

6. DISCUSSION AND CONCLUSIONS

In agreement with Georgiev & Ivanov (1997), we concluded that *the radial distributions of more evolved stars do not simply follow those of their suspected progenitors* but we explored a larger number of evolutionary stages and our study is based on a richer sample. Importantly, the correlation analysis using all three statistical tests arrived at the same conclusion.

The evolutionary connection between YSGs and CEPs is well known (e.g. Rodgers 1957, Turner 1996), therefore the pair YSG – CEP can be used to verify the reliability of our conclusions. None of the tests indicates a YSG – CEP correlation. The result supports the proposition that evolutionary connections may be traced back to correlations between spatial distributions but this alone cannot explain the majority of the suspected causal relationships.

However, looking back at Figure 1, we can see that even though there are differences in the overall distributions, it seems like they follow each other within some regions. For example, SNRs follow WRs closely in galactocentric distances approx. [2.8; 5.4] kpc. Outside of this

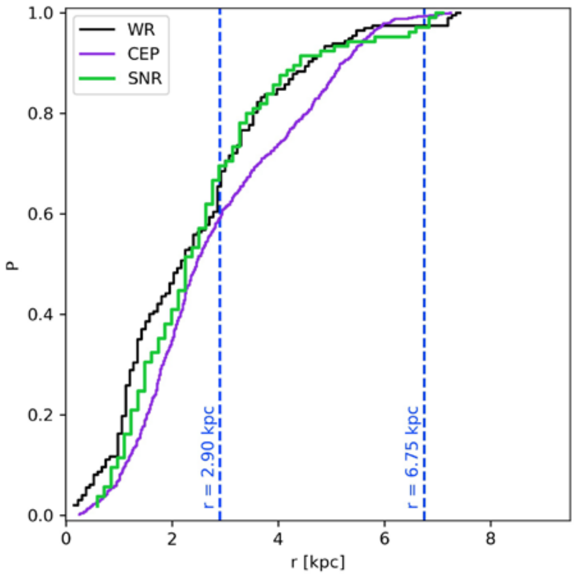

Fig. 5 Local correlation of SNRs and WRs in the intermediate region of M33. p-values: KS = 95.06%; MWUT = 95.68%; ADT > 25.00%

region, SNRs follow CEPs more closely than WRs. This is shown in Figure 5.

Therefore, we can apply the statistical tests to compare the distributions of objects within rings (or circles, if $r_{min} = 0$). The results are presented in Table 1. We have selected only intervals with at least 20 members and $a \geq 98\%$ (the ADT p-values is $> 25\%$ in all presented pairs). We favored longer intervals over higher p-values.

We can see that near the center, SNRs tend to follow Cepheids more closely, and at larger r, their spatial distribution correlates better with WRs.

Contrary to Georgiev & Ivanov (1997), we also found indication of a strong relationship between WRs and BSGs spanning an annulus of ~60% of the galactic radius.

The test presented here failed to find a significant galactic-scale relationship between $P(r)$ for the given pairs of objects. However, we were able to extract such relationships on a scale of 0.7 – 4.75 kpc (~7-50% of the galactic radius). This may be an indication that a clumpy structure dominates stellar formation in M33. As each individual clump has a random age, the correlations between stellar populations vanish on the galactic scale. This proposition requires further study, including comparison by both r_{gal} and polar angle φ, and taking into account the characteristic timescales of clumps to the lifetimes of each stellar population.

Another consideration would be estimating the binary fraction of M33. Binary interactions have a drastic effect on the evolution of massive stars, most notably on WRs (Dorn-Wallenstein & Levesque, 2018).

7. ACKNOWLEDGEMENTS

We would like to thank Dr. Valentin Ivanov of ESO for his unparalleled help in editing and reviewing this article.

This work was supported in part by National Science Programme "Young Scientists and Postdoctoral Students 2019", NSF Bulgaria grant DN18/10-11.12.2017, National RI Roadmap Projects DO1-157/28.08.2018 and DO1-153/28.08.2018.

TABLE 1. The KST an MWUT p-values of object pairs restricted by galactocentric distance between r_{min} and r_{max} in kpc. n and m are the sizes of the first and second subsample respectively.

Pair	r_{min}	r_{max}	n	m	KST	MWUT
BSG – YSG	2.30	5.95	270	75	98.67%	99.53%
BSG – RSG	3.95	5.35	83	68	98.56%	98.96%
BSG – WR	3.20	7.95	198	58	99.16%	99.60%
BSG – CEP	1.00	1.70	63	114	98.09%	99.63%
BSG – SNR	1.60	2.85	99	40	98.60%	98.70%
YSG –RSG	3.20	5.50	52	126	98.74%	99.36%
YSG – SNR	2.95	4.75	42	26	99.52%	99.50%
RSG – CEP	3.35	4.35	57	80	98.62%	98.08%
WR – CEP	3.85	5.15	21	111	99.02%	98.01%
CEP – SNR	0.45	1.75	163	32	97.87%	99.32%

REFERENCES

Abott DC, Conti PS. 1987. *Wolf – Rayet Stars*. ARAA 25 (1987):113-50.

de Vaucouleurs, G. 1959. *Photoelectric Photometry of Messier 33 IN the u, b, v, System*. ApJ 130:728 (1959).

Dorn-Wallenstein TZ, Levesque EM. 2018. *Stellar Population Diagnostics of the Massive Star Binary Fraction*. ApJ (2018) 867, 125.

Freedman WL, Wilson CD, Madore, BF. 1991. *New Cepheid Distances to Nearby Galaxies Based on BVRI CCD Photometry. II. The Local Group Galaxy M33*. ApJ 372:455.

Garofali K, Williams BF, Plucinsky PP, Gaetz TJ, Wold B, Haberl F, Long KS, Blair WP, Pannuti TG, Winkler PF, Gross J. 2017. *Supernova remnants in M33: X-ray properties as observed by XMM-Newton*. MNRAS 472.1 (2017): 308-333.

Georgiev LN, Ivanov GR. 1997. *Observational Evidence for Evolutionary Connections between Wolf – Rayet Stars and Red Supergiants*. RMxAA 33, 117.

Gil de Pas, A et al. 2007. *The GALEX Ultraviolet Atlas of Nearby Galaxies*. ApJS 173.2 (2007): 185-225.

Harding EF. 1983. *An Efficient Minimal Storage Procedure for Calculating the Mann-Whitney U, Generalised U and Similar Distributions*. JRSS: AS (1983) 33: 1.

Humphreys, RM. 2019. *The Complex Upper HR Diagram*.

Ivanov, GR. 1998. *Young Massive Stellar Populations in M33*. A&A 337 (1998): 39-42.

Kuijken, K et al. 2011. *OmegaCAM: ESO's Newest Imager*. The Messenger, 8-11.

Maeder A. 1996. *Stellar Evolution: High Mass*. In: From Stars to Galaxies: The Impact of Stellar Physics on Galaxy Evolution. ASPCS 98, 1996.

Massey P, Johnson O. 1998. *Evolved massive stars in the Local Group. II. A New Survey for Wolf – Rayet stars in M33 and its implication for massive star evolution: evidence of the "Conti Scenario" in action.* AJ 505 793.

Massey P. 2003. *Massive stars in the local group: implications for stellar evolution and star formation*. ARAA. 41, 15–56.

Massey P, Neugent KF, Smart, BM. 2016. *A Spectroscopic Survey of Massive Stars in M31 and M33*. AJ 152.3 (2016): 62.

Maund JR, Crowther PA, Janka H-T, Langer N. 2017. *Bridging the gap: from massive stars to supernovae.* PhTR. A375: 20170025.

Neugent KF, Massey P. 2011. *The Wolf – Rayet Content of M33*. ApJ 733.2 (2011): 123.

Pellerin A, Macri, LM. 2011. *The M33 Synoptic Stellar Survey. I. Cepheid Variables*. ApJS 193.2 (2011): 26.

Rodgers AW. 1957. Radius variation and population type of cepheid variables. MNRAS 117: 85–94.

Scholz FW, Stephens MA. 1987. *K-sample Anderson – Darling Tests*. JASA 82.399: 918-924.

Skrutskie, MF et al. 2006. *The Two Micron All Sky Survey (2MASS)*. AJ 131.1163.

Turner, DG. 1996. *The Progenitors of Classic Cepheid Variables*. JRASC 90:82.