A single model for the variety of multiple-population formation(s) in globular clusters: a temporal sequence

F. D’Antona,1 E. Vesperini,2 A. D’Ercole,3 P. Ventura,1 A. P. Milone,4 A. F. Marino4 and M. Tailo1

1INAF, Osservatorio Astronomico di Roma, Via Frascati 33, I-00078 Monteporzio Catone (Roma), Italy
2Department of Astronomy, Indiana University, Swain West, 727 E. 3rd Street, Bloomington, IN 47405, USA
3INAF, Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy
4Research School of Astronomy & Astrophysics, Australian National University, Canberra, ACT 2611, Australia

Accepted 2016 February 16. Received 2016 February 16; in original form 2015 December 7

ABSTRACT

We explain the multiple populations recently found in the ‘prototype’ globular cluster (GC) NGC 2808 in the framework of the asymptotic giant branch (AGB) scenario. The chemistry of the five – or more – populations is approximately consistent with a sequence of star formation events, starting after the Type II supernova epoch, lasting approximately until the time when the third dredge-up affects the AGB evolution (age \(\sim 90–120\) Myr), and ending when the Type Ia supernovae begin exploding in the cluster, eventually clearing it from the gas. The formation of the different populations requires episodes of star formation in AGB gas diluted with different amounts of pristine gas. In the nitrogen-rich, helium-normal population identified in NGC 2808 by the UV Legacy Survey of GCs, the nitrogen increase is due to the third dredge-up in the smallest mass AGB ejecta involved in the star formation of this population. The possibly iron-rich small population in NGC 2808 may be a result of contamination by a single Type Ia supernova. The NGC 2808 case is used to build a general framework to understand the variety of ‘second-generation’ stars observed in GCs. Cluster-to-cluster variations are ascribed to differences in the effects of the many processes and gas sources which may be involved in the formation of the second generation. We discuss an evolutionary scheme, based on pollution by delayed Type II supernovae, which accounts for the properties of s-Fe-anomalous clusters.


1 INTRODUCTION

The Hubble Space Telescope (HST) UV Legacy Survey of Galactic globular clusters (GCs; Piotto et al. 2015) is exploiting the sensitivity of UV photometric observations to different molecular bands to disentangle GCs’ multiple stellar populations in 57 clusters. Among the surprises of the UV Survey’s first results, there is the discovery that the multiple-population cluster NGC 2808 not only contains the three populations known to differ in helium content (D’Antona et al. 2005; Piotto et al. 2007) and in light element abundance distribution (Carretta et al. 2006), but it includes at least two more families (Milone et al. 2015a). The five populations identified have been labelled by Milone et al. (2015a) with letters from A to E. The intermediate- and the high-helium main sequences (MS) are easily identified and correspond to populations D and E, respectively. The population with almost standard helium includes at least three stellar groups, namely A, B and C. In particular, group C includes a significant fraction of cluster stars (~25 per cent), does not differ in helium content from group B, but shows the fingerprints of a high surface nitrogen content. Group A is a small population (~6 per cent) for which spectroscopic data are not available and which could be slightly more metal rich than the rest of the cluster stars. Consequently, group B alone is fully compatible with a ‘first generation’ (FG).

Very recently, Carretta (2015) has reanalyzed the large data base of spectroscopic data collected in several years for the cluster giants, and has shown that these five populations can be distinguished also from the spectra, thanks to their clustering in the Mg–Na, O–Mg, Mg–Si and O–Na planes. The groups identified spectroscopically and photometrically do not completely overlap: spectroscopy can distinguish two different groups among the stars of each of the
groups P1, P2, I1 and I2), while the E group is the same (also in photometric groups B, C and D (which include the spectroscopic multiple populations, as it has been one of the first clusters, together with ω Cen, in which helium variations have been postulated and discovered. NGC 2808, however, in addition to populations with milder chemical anomalies similar to those found in most GCs, hosts also a population with more extreme chemical variations, found in only few other clusters. In this work, our goal is to explore the viability and the key ingredients of a unified model for the origin of GC abundance patterns, capable of explaining both the more modest and widespread anomalies and the more extreme ones.

Can the models, built to account for a simpler scheme of multiple populations, be extended to deal with the more complex observational findings emerging from recent studies? In the latest 10–15 years, the observations promoted the formulation of several different theoretical explanations and provided a large number of constraints that have challenged the proposed theories.

We emphasize here that, currently, the strongest observational constraints come from the chemical abundance properties and patterns revealed by spectroscopic and photometric observations. Any effort aimed at identifying the fundamental ingredients in the formation of multiple populations should be driven by these observational constraints which provide fundamental clues on the possible sources of gas out of which second-generation (SG) stars formed. While many factors related to a cluster internal dynamics and its complex interplay with the external galactic environment (particularly in the early stages of galaxy formation) have certainly played a key role in determining the current properties of multiple-population clusters, the first step in assessing the viability of different models should be guided by the observed chemical properties. Although it is extremely important to explore all the implications of the presence of multiple stellar populations for a cluster dynamical history, using other constraints, related to much more uncertain aspects of the formation history of GCs and of their host galaxies, to rule out any model appears premature.

We will exploit the asymptotic giant branch (AGB) scenario (Ventura et al. 2001; D’Ercole et al. 2008; Bekki 2011) to interpret the five populations of NGC 2808. This scenario has been developed with sufficient detail to address the main photometric and spectroscopic constraints. Moreover, we will provide a critical analysis of the other main scenarios for the formation of the distinct stellar populations at the light of the recent observations on NGC 2808.

The AGB scenario could nicely and simply explain helium differences among the stars in this cluster, which had been predicted by analysing the horizontal branch (HB) morphology (D’Antona & Caloi 2004). When the triple MS was discovered (D’Antona et al. 2005; Piotto et al. 2007), parametric modelling of the helium evolution in the cluster became more detailed, as first shown in the hydrodynamical computations in D’Ercole et al. (2008), and subsequently in simple chemical evolution models (D’Ercole et al. 2010, 2012) including other chemical abundance patterns. The same model, slightly extending the time of formation of the multiple population, is here shown to provide a good explanation for the other two populations. The five populations, in the order BEDCA, are shown here to represent the outcome of a clear temporal sequence of star formation (SF) events. Specifically, we show that the different chemical fingerprints are in fact in close relation with the temperatures at the bottom of the convective envelope, $T_{\text{HBB}}$, of the AGB stars providing the gas out of which SG stars formed. The decrease of $T_{\text{HBB}}$ with the evolving initial mass reveals the possible temporal sequence for the formation of the different observed patterns.

In the specific case of NGC 2808, populations C and A simply result from a slight extension in the duration of the SF epoch proposed in the context of the AGB model for populations B, E and D by D’Ercole et al. (2008, 2012).

The AGB scenario is able to trace the star formation history (SFH) of multiple populations in NGC 2808. This model can be easily adapted to any other GC, once we allow for simple-minded differences in the events which modulate SFH in each cluster, based on the same – or similar, depending on the metallicity – temporal evolution of the polluting material. In this respect, this work aims at providing a reasonable baseline to understand this complex problem.

In this paper, we explore the possible chemical evolution of NGC 2808 with a simplified model, discuss the possible events separating the SF epochs of the five (or more) populations, and suggest which parameters (e.g. the duration of the SF SF phase, the timing and intensity of the pristine gas accretion phase, the properties and frequency of the dividing events, the role of delayed binary Type II supernovae (SN II), the sudden or slow onset of SN Ia explosions) may produce different abundance patterns in other clusters.

The outline of the paper is the following.

In Section 2, we summarize the HST spectrophotometric and the spectroscopic data which form the basis of our knowledge about the five (or more) populations hosted in NGC 2808. We show that the magnesium variations are at variance with the results of most models proposed so far, while they can be compatible with the AGB scenario. We anyway describe the difficulties from AGB nucleosynthesis, and the possible solutions.

In Section 3, we describe how the yields of AGB ejecta vary with time, due to the decrease of the hot bottom burning (HBB) temperatures with decreasing initial mass, and discuss that the patterns of abundances may be correctly described only if, at some stage(s), the ejecta are diluted with pristine matter. In Section 4, we produce a parametric model for NGC 2808, able to explain – at least qualitatively – the evolution with time of the composition of the multiple populations. The time order of populations A, B, C, D and E in Milone et al. (2015a) is given by the acronym BEDCA. In this timeline, we propose feasible explanations for the new populations C and A. In Section 5, we show that group C stars are also compatible with their location in the HB.

In Section 6, we present a brief discussion concerning the chemical patterns of clusters of different metallicities and how they are consistent with the trends in the AGB yields as a function of the model metal content.

In Section 7, we extend the model to other clusters, after examining in detail the different epochs during the early life of clusters, in terms of both AGB ejecta composition and of events perturbing the SF. The possible role of delayed SN II and SN Ia in affecting the resulting chemical properties of SG stars is discussed. Finally in Section 8.1 we propose a model to explain the features of s-Fe-anomalous clusters, and we elaborate on the possible iron pollution of both kind of late supernovae. Results are summarized in Table 4 in the conclusions section.

Multiple populations in globular clusters
Table 1. Spectroscopic and spectrophotometric abundances for NGC 2808 populations.

<table>
<thead>
<tr>
<th>C15</th>
<th>M15</th>
<th>[O/Fe]</th>
<th>[Mg/Fe]</th>
<th>[Al/Fe]</th>
<th>[Na/Fe]</th>
<th>[Fe/H]</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>B</td>
<td>0.30</td>
<td>0.38</td>
<td>0.05</td>
<td>0.03</td>
<td>-1.13</td>
</tr>
<tr>
<td>P2, I</td>
<td>L</td>
<td>C</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>I2 D</td>
<td>E</td>
<td>-0.7</td>
<td>-0.3</td>
<td>1.0</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.9</td>
<td>-0.4</td>
<td>1.2</td>
<td>0.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Notes: a Carretta (2015) groups definition.

b Milone et al. (2015a) groups definition.
c Values from Milone et al. (2015a), matched to Carretta et al. (2006) and Carretta (2014).
d All variations with respect to B group values are from Milone et al. (2015a) apart from those in (c).
e From Carretta (2015).

2 MODELS VERSUS OBSERVATIONAL CONSTRAINTS

2.1 Observations

Carretta (2015) summarizes the patterns of abundances of light elements which define the spectroscopic evidence for the presence of multiple populations in NGC 2808. Variations are found in the ‘classical’ elements oxygen, sodium, magnesium and aluminium, but also in the heavier element silicon. Further, variations in potassium (Mucciarelli et al. 2015) must be also considered. In this work, Carretta also compares his spectroscopic analysis with the photometric results, and results in details the specificity of the two different approaches.

One important observational clue is that magnesium decreases by about 0.4 dex, and silicon increases by about 0.2 dex between the P1 (primordial 1) population and the ‘extreme’ E population (Carretta 2015).

The photometric two-colour diagram introduced by Milone et al. (2012c) and used to study the populations of NGC 2808 in Milone et al. (2015a) works on the amplification of the differences in C, N and O abundances in the spectral distribution of stars, which is obtained when using specific combination of the UV and optical HST passbands, e.g. the plane of the colour $y = m_{F336} - m_{F438}$ versus the colour $x = m_{F275} - m_{F814}$. As C, N and O are the main drivers of the clustering of stars in this plane, we will refer to this diagram as ‘the CNO two-colour diagram’. Here, the stars group into at least five main regions, labelled from A to E.

Groups B, D and E are easily identified as the helium-normal, intermediate and extreme groups already known. Two more groups are present: C, nitrogen richer than B, and the small group A. Spectrophotometric differential analysis of the magnitudes in the different bands from Milone et al. (2015a) provides the results in Table 1 for the populations C, D and E with respect to group B assumed to be the FG. The comparison with the Carretta (2015) analysis allows us to include in the table the differential values for Mg and Al. Errors are not listed in the table, as the data come from non-homogeneous measurements, for which a rigorous evaluation is difficult. We use the abundance differences mostly to show the qualitative agreement of their values with the model trends (Figs 3 and 4).

The A group is not listed in the table, as the differential analysis in Milone et al. (2015a) has been done by assuming a helium abundance smaller than in the reference group B by $\delta Y = -0.03$, while in this work we assume that A group stars are more iron rich by $\delta [Fe/H] \sim 0.1$, a value estimated by the slightly cooler location of the A with respect to the B MS (Milone et al. 2015a). Notice that the small colour increase may also be due to an increase in the specific abundances of oxygen, or silicon, or sulphur, as shown in the detailed study by VandenBerg et al. (2012). A conclusive understanding will require abundance measurements in the atmosphere of these stars.

2.2 Constraining viable models with Mg and Si

In NGC 2808, there is evidence of both Mg depletion and Si enhancement. The maximum Mg depletion, about 0.4 dex, is found among the stars of population E in Carretta (2015). This fact alone is sufficient to rule out all proposed models for the formation of multiple populations based on the nucleosynthesis product available from core H-burning in massive stars, namely the fast rotating massive star (FRMS) model (e.g. Decressin et al. 2007), the Massive Interacting Binary model (de Mink et al. 2009) or the accretion model (Bastian et al. 2013) for which the nucleosynthesis products of massive binaries are the source of chemical anomalies.

Inside these stars, the maximum temperature does not exceed $\sim 65$ MK, an extreme value reached only during the latest phases of core H-burning by the most massive models (Decressin et al. 2007), so magnesium cannot be affected by proton captures and silicon cannot be synthesized. The problem is discussed in the Prantzos, Charbonnel & Iliadis (2007) analysis, based on one-zone chemical evolution models at fixed temperature, which are not linked to any specific computation. Notice that this constraint cannot be changed by variations in the details of the stellar models, as it depends on the fundamental properties of stellar structure. By combining the observed mass–radius relation, for stars burning core hydrogen via the CNO cycle, with the hydrostatic equilibrium plus the perfect gas law, the central temperature $T_c$ versus total mass relation is described by a power law with a very shallow exponent $\sim 0.2–0.3$, providing values of $\sim 50$ MK for the interior temperature of stars at $M \sim 100 M_\odot$.

In an attempt to increase the temperature, in the core H-burning phase, up to the canonical 75 MK which allow Mg and other advanced p-capture reactions in the interiors, Denissenkov & Hartwick (2014) have suggested that supermassive stars are responsible for the formation of SG stars in GCs. However, such stars have no observational counterparts.

In conclusion, the only possibility for massive star models to overcome this fundamental problem is to assume a larger cross-section for the $^{24}$Mg($p$,y)$^{25}$Al reaction, possibly through a not yet identified resonance at the needed temperatures. An increase in the cross-section by a factor 1000 is, however, needed (Decressin et al. 2007), and because of such a large discrepancy we think these models can be ruled out based on what is now known from the nuclear evolution point of view. Renzini et al. (2015) further discuss the inability of all these models to produce the observed discreteness and to explain cluster-to-cluster differences in the properties of multiple populations, as well as the reliability of supermassive star models as polluters for the multiple populations, but here we rule them out on the simple basis that they cannot be used to describe the abundances observed by Carretta (2015), et de hoc satis.
2.3 The problems of nucleosynthesis in the AGB scenario

So we claim that p-captures on Mg nuclei and other advanced nucleosynthesis products find a place for the nucleosynthesis in the HBB envelopes of massive AGBs, in the models in which the temperature of HBB ($T_{\text{HBB}}$) is sufficiently high ($\gtrsim 10^8$ K; Prantzos, Charbonnel & Iliadis 2007; Ventura, Carini & D’Antona 2011). Nevertheless, the quantitative result of Mg burning in massive AGBs of metallicity adequate to describe NGC 2808 also faces problems: in the models that will be used here (Ventura et al., 2013), the maximum Mg depletion is $\sim 0.17$ dex, while the data require $\sim 0.4$ dex. We have discussed in specific models (Ventura et al. 2011) that the $T_{\text{HBB}}$’s of the massive AGB are indeed large enough to allow the reaction $^{24}\text{Mg}(p,\gamma)^{25}\text{Al}$ to occur, but longer evolutionary times in AGB are needed to get quantitative agreement in the total depletion. This can be achieved by assuming a smaller mass-loss rate. The essence of the problems with AGB modelling is the following. Smaller mass-loss rates than those adopted in these models allow magnesium and aluminium to be processed more efficiently, but have the counter-effect of reducing the total yield of sodium, disfavoring the O–Na anticorrelation. This trend is also confirmed in the models by Doherty et al. (2014), which show a larger Mg depletion, but also full depletion of sodium in the yields. In fact, the nucleosynthesis in HBB has to face the plain fact that sodium is destroyed at the same time oxygen is destroyed, while the abundances displayed by SG stars in GCs require high sodium abundances at low oxygen abundances.

In the following, we list the three quantitative discrepancies between AGB nucleosynthesis yields and observed abundance patterns that are solved if we allow for smaller mass-loss rates.

1. The oxygen depletion in the most massive super-AGBs is not large enough. To address this problem, we have invoked deep mixing in the giant progenitors of the most extreme population, a mixing which should be favoured by the large helium abundance of these stars (D’Antona & Ventura 2007), and we adopt this solution also here (see Section 4.1). Nevertheless, smaller mass-loss rates in the super-AGB phase would lead to the required large depletion of oxygen and would provide an additional possible solution to this problem.

2. Apart from $^{24}\text{Mg}$ depletion, some other HBB proton captures require a longer time than allowed by the mass-loss rates chosen in our Ventura et al. (2013) paper and in the previous computation. In particular, the silicon increase by proton captures on aluminium is found in Ventura et al. (2013) at the level of $\sim 0.05$ dex, while $\sim 0.15$ dex are required by Carretta (2015) data for NGC 2808. Also in this case, a larger production requires longer evolutionary times.

3. Taking at face value the modest $\delta Y$ implied by the observational width of the MS of some clusters in recent observations, Bastian, Cabrera-Ziri & Salaris (2015) found a discrepancy between the range covered by data in O–Na plane and the range allowed by models. Below, we discuss some limitation of this analysis, and other possible ways to overcome this problem. However, taking this result at face value, also in this case the difficulty is removed if the models have a larger oxygen depletion for a given $T_{\text{HBB}}$, and this can be achieved assuming smaller mass-loss rates during the AGB evolution. If models with smaller $T_{\text{HBB}}$ can describe well the O–Na data, they are progeny of lower initial masses and have smaller $Y$ yield, as required by the comparisons in Bastian et al. (2015), solving the discrepancy.

The mass-loss rates in the AGB models by Ventura et al. (2013) were originally calibrated on a comparison with the luminous lithium-rich stars in the Magellanic Clouds (Ventura, D’Antona & Mazzitelli 2000), but this calibration is also able to provide a positive sodium yield, in spite of the contemporary burning of sodium and oxygen. If we allow for a mass-loss rate smaller, say by a factor of 3 – that is, we triple the time for p-captures in the AGB envelope – we may find a solution for the four problems listed above, at the expenses of burning too much sodium (see e.g. Ventura et al. 2011; Doherty et al. 2014).

The problem of the sodium destruction is examined in detail in Renzini et al. (2015), where it is suggested that a reduction by a factor of 5 of the reaction rate $^{23}\text{Na}(p,\gamma)^{24}\text{Mg}$ would allow sodium to be destroyed at a slower rate, so that its global yield may be in the range of the observed sodium abundances also in the stars in which oxygen is largely depleted. So the discrepancies with observations of AGB nucleosynthesis can be solved by adopting a reduced mass-loss rate, if the proton capture rate on sodium is smaller.

Reduction of the sodium burning rate also mitigates the problem of the sodium yield in very low metallicity models. Very low metallicity clusters (e.g. M15; Carretta et al. 2009a) show the typical anticorrelation O–Na. As opacities are smaller at lower metallicity, the $T_{\text{HBB}}$’s of the AGB models are larger, so that sodium is burned more efficiently, and it is even more difficult than for clusters of larger metallicity to preserve a large sodium yield (Ventura & D’Antona 2009). This problem might show up particularly in NGC 2419 (Ventura et al. 2012), because this cluster contains a large, extreme, population with very high helium, interpreted as a population born from pure ejecta (Di Criscienzo et al. 2011, 2015). In this cluster, anyway, there are very few measurements of sodium abundances (Cohen & Kirby 2012) and the presence of an O–Na anticorrelation is not clear-cut (Ventura et al. 2012). The truly needed range of reduction in the rate of $^{23}\text{Na}(p,\gamma)^{24}\text{Mg}$, beyond the limits published by Hale et al. (2004),1 should be verified by model computation, together with a check of the best mass-loss rate producing agreement with the whole set of abundances displayed by SG stars. We estimate that the reduction required should indeed be in the range 2–5. This reaction proceeds populating several states. Low-energy resonances in the $^{23}\text{Na}+p$ system have been studied by Zyskind, Rios & Rolfs (1981), Goerres, Wiescher & Rolfs (1989) and Rowland et al. (2002), down to the states of energy of about 200 keV. Below the lowest observed resonance at $E_{\text{res}} = 170$ keV, there are four other states, which could correspond to resonances at 5, 37, 138 and 167 keV. Hale et al. (2004) describe an indirect experiment to determine spectroscopic factors via ($^1\text{He}, d$) for these low-energy states, and concluded that the state at $E_{\text{res}} = 138$ keV may make a significant contribution to the rates of $^{23}\text{Na}(p,\gamma)^{24}\text{Mg}$. Due to the weak experimental constraints, the uncertainty of the reaction rate at the relevant astrophysical energy (70–110 MK) is very large and could be larger than a factor of 2. To shed light on the intensity of this state, several experiments are planned in different laboratories, e.g. LUNA (Costantini et al. 2009), which is concentrating on the gamma channel $^{23}\text{Na}(p,\gamma)^{24}\text{Mg}$.

In the present work, we will mostly consider the trends with time of the yields predicted in AGB models, as we are interested in qualitative or semi-quantitative comparisons. The important point of the comparison is to understand whether these trends, which depend primarily on $T_{\text{HBB}}$, provide a general scheme to understand the different populations in NGC 2808, and if this scheme can be simply adapted to other clusters.

1 The lower limit of the cross-section determined by Hale et al. (2004) is only 20–25 per cent smaller than the recommended rate, at $T \sim 100$ MK.
to the original abundance in the star only for a very small range of spectra (and also here we have to rely on models) is directly linked visible only at the colours of MS stars must be interpreted. The helium lines are Bastian et al. (2015).

A larger depletion of oxygen for each evolving HBB model alleviates or cancels the helium discrepancy presented by Bastian et al. (2015). Although we think that this is a likely solution to all the discrepancies, we further discuss here the helium problem raised by Bastian et al. (2015).

It is important to emphasize that the all the δY are ‘indirect’ determinations from the observations. In particular, the values quoted in Bastian et al. (2015) depend on a number of assumptions on how the colours of MS stars must be interpreted. The helium lines are visible only at \( T_{\text{eff}} \gtrsim 8000 \) K, and the abundance measured in the spectra (and also here we have to rely on models) is directly linked to the original abundance in the star only for a very small range of \( T_{\text{eff}} \) in the HB (Behr 2003; Moni Bidin et al. 2007).

Besides, a careful look at the data adopted by Bastian et al. (2015) is necessary to assess the real importance of the problem. In particular, the O–Na data in Bastian et al. (2015) are mostly taken from the useful collection by Roediger et al. (2014), joining data from different sources. As an example, the data of the cluster 47 Tuc used in the work are the combination of two data sets having different zero-points of abundances.

A most critical point is that the percentage of stars along the predicted dilution curves should be considered in detail. In fact, the δY in Milone et al. (2012b) is based on the average distance in MS colour between the two populations, and is certainly not the maximum allowed value. A longer discussion of this issue is out of the scope of the present work, but it is under detailed scrutiny using the UV Legacy Survey data (Milone et al., in preparation).

Another factor which may affect this possible discrepancy is the following: the helium yield from massive AGBs is the result of the second dredge-up (2DU), occurring between the H-rich envelope of the giant and the H-exhausted core before the model climbs up the AGB (see Section 3.1). In contrast, the p-processing occurs at a later stage, during the HBB in the AGB envelopes, and is highly dependent on entirely different physical inputs [we have discussed mass-loss, but one of the most critical is the convection modelling; see Renzini et al. (2015) for a longer discussion and references]. Thus, should further observations confirm this discrepancy, the solution to this problem might be in models with different modalities of the 2DU, reducing the discrepancy in the δY. In Table 2, we show how a simple difference in the models alters the envelope Y abundance after the 2DU by δY = 0.02 for masses 6 and 5 M⊙.

Table 2. Helium mass fraction at the 2DU as a function of the extra mixing during core-H burning.

<table>
<thead>
<tr>
<th>( M_{\text{AGB}}/M_{\odot} )</th>
<th>( \zeta )</th>
<th>( M_{\text{core}}/M_{\odot} )</th>
<th>( Y_{\text{2DU}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0</td>
<td>0.01</td>
<td>0.988</td>
<td>0.344</td>
</tr>
<tr>
<td>6.0</td>
<td>0.02</td>
<td>1.025</td>
<td>0.349</td>
</tr>
<tr>
<td>6.0</td>
<td>0.03</td>
<td>1.060</td>
<td>0.362</td>
</tr>
<tr>
<td>5.0</td>
<td>0.01</td>
<td>0.904</td>
<td>0.324</td>
</tr>
<tr>
<td>5.0</td>
<td>0.02</td>
<td>0.917</td>
<td>0.332</td>
</tr>
<tr>
<td>5.0</td>
<td>0.03</td>
<td>0.939</td>
<td>0.343</td>
</tr>
</tbody>
</table>

Notes. *Initial mass – Y = 0.25, Z = 0.001 and [α/Fe] = 0.4.
*b-e-folding distance of overshooting (Ventura et al. 1998).
*c-C–O core mass at the AGB.
*d-Helium mass fraction after the 2DU.

2.4 Helium versus p-capture elements

As discussed in the previous section, a reduction of the proton capture rate on sodium, together with the assumption of smaller mass-loss rates, can solve all the quantitative discrepancies between the AGB yields and the observed abundance anomalies; in particular, a larger depletion of oxygen for each evolving HBB model alleviates or cancels the helium discrepancy presented by Bastian et al. (2015). Although we think that this is a likely solution to all the discrepancies, we further discuss here the helium problem raised by Bastian et al. (2015).

It is important to emphasize that the all the δY are ‘indirect’ determinations from the observations. In particular, the values quoted in Bastian et al. (2015) depend on a number of assumptions on how the colours of MS stars must be interpreted. The helium lines are visible only at \( T_{\text{eff}} \gtrsim 8000 \) K, and the abundance measured in the spectra (and also here we have to rely on models) is directly linked to the original abundance in the star only for a very small range of \( T_{\text{eff}} \) in the HB (Behr 2003; Moni Bidin et al. 2007).

Besides, a careful look at the data adopted by Bastian et al. (2015) is necessary to assess the real importance of the problem. In particular, the O–Na data in Bastian et al. (2015) are mostly taken from the useful collection by Roediger et al. (2014), joining data from different sources. As an example, the data of the cluster 47 Tuc used in the work are the combination of two data sets having different zero-points of abundances.

A most critical point is that the percentage of stars along the predicted dilution curves should be considered in detail. In fact, the δY in Milone et al. (2012b) is based on the average distance in MS colour between the two populations, and is certainly not the maximum allowed value. A longer discussion of this issue is out of the scope of the present work, but it is under detailed scrutiny using the UV Legacy Survey data (Milone et al., in preparation).

Another factor which may affect this possible discrepancy is the following: the helium yield from massive AGBs is the result of the second dredge-up (2DU), occurring between the H-rich envelope of the giant and the H-exhausted core before the model climbs up the AGB (see Section 3.1). In contrast, the p-processing occurs at a later stage, during the HBB in the AGB envelopes, and is highly dependent on entirely different physical inputs [we have discussed mass-loss, but one of the most critical is the convection modelling; see Renzini et al. (2015) for a longer discussion and references]. Thus, should further observations confirm this discrepancy, the solution to this problem might be in models with different modalities of the 2DU, reducing the discrepancy in the δY. In Table 2, we show how a simple difference in the models alters the envelope Y abundance after the 2DU by δY = 0.02 for masses 6 and 5 M⊙.

Notice that the C–O core mass is also altered, so a fully consistent modelling is required for a global study of this problem. The second column in Table 2 lists the value of the parameter \( \zeta \) describing the overshooting in our models (Ventura et al. 1998). \( \zeta \) is the e-folding distance – in units of the thickness of the convective region expressed as a fraction of the pressure scaleheight \( H_p \) – over which the model assumes that the convective velocities decay exponentially from the formal convective border. A parameter \( \zeta = 0.02 \) is the value which is generally adopted in the models, based on the comparison with observations of the MS width, but larger or smaller values cannot be excluded. In conclusion, we think that, at this point, the claim by Bastian et al. (2015) that the constraints from O–Na–Y data rule out the AGB model is hasty; until additional theoretical and observational work on this issue is carried out and further light on the actual extent of the problem and its possible solutions is shed, much caution in drawing any strong conclusion should be exercised.

3 INGREDIENTS OF THE AGB MODEL

3.1 The AGB ejecta composition versus time

In the previous sections, we discussed that several modifications to the current AGB models will be needed, and that future observations will be critical in driving the refinement of the AGB models (see also Renzini et al. 2015). Now we focus our attention on whether the fundamental ingredients of the AGB scenario may satisfy the main chemical constraints displayed by multiple populations.

During the thermally pulsing phase, models of large enough mass reach the very high temperatures, \( T_{\text{HBB}} \), necessary to achieve the proton captures which modify the light element abundances previously quoted. Fig. 1 shows the maximum \( T_{\text{HBB}} \) reached, as a function of the initial mass and metallicity, in the evolutionary computations by Ventura et al. (2013). The assumptions behind these results can be found in their paper. Here we simply remind the reader that the relation of \( T_{\text{HBB}} \) versus the initial mass depends on two main parameters.

(1) The assumptions made on core overshooting during the H-core burning, which provides the mass of the H-exhausted core. This He-core mass is modified by the 2DU (Becker & Iben 1979), and has a smaller reduction for stronger core-overshooting assumptions (see Table 2). The mass of the C–O core remnant at He exhaustion (Bressan et al. 1993) is very close to the He-core mass after the 2DU. Larger C–O mass means larger \( T_{\text{HBB}} \) (e.g. Boothroyd, Sackmann & Ahern 1993).

(2) The assumptions made for the efficiency of external convection determine \( T_{\text{HBB}} \) for each core mass (Ventura & D’Antona 2005a).

Different choices for these parameters change the temperature values in Fig. 1. In particular, a smaller efficiency of convection may not allow to reach \( T_{\text{HBB}} \)’s above \( 10^8 \) K, or it may be reached only at very low metallicity, or only for the largest masses. These models would not be able to explain, e.g., the formation of silic.

On addition, C+N+O would increase in the whole mass range of massive AGBs, providing results at variance with observations (Fenner et al. 2004).

While we must take Fig. 1 with caution for what concerns the temperature numerical values, the trends with metallicity and mass have a validity beyond the particular assumptions of specific models. Notice that the Ventura et al. (2013) models follow a line of models computed with very efficient convection (Canuto & Mazzitelli 1991;
Canuto, Goldman & Mazzitelli (1996), so they probably already have the largest $T_{\text{HBB}}$ at fixed core mass that can be reached in 1D models. On the other hand, these models use a moderate overshooting, so a bit larger core mass (and a larger $T_{\text{HBB}}$) is possible for a given initial mass in models with increased overshooting.

The role of metallicity in determining the possible HBB processing is very clear: the smaller the metallicity, the larger is $T_{\text{HBB}}$ (due to the smaller surface opacities). This means that the products of the hottest HBB (e.g. Si production from proton captures on Al) can be found only at low metallicity, and for large initial masses (see also Section 6).

The final nucleosynthesis products depend on $T_{\text{HBB}}$, but also on the total time spent in the AGB phase, and the total time spent in the AGB phase, in turn, ultimately depends on the mass-loss rate. High mass-loss rate may drastically limit the p-captures (Ventura & D’Antona 2005b). In fact, part of the differences between the sodium and oxygen yields provided by Ventura et al. (2013) and Doherty et al. (2014) in the massive AGB star regime are due to the different prescriptions adopted for mass-loss.

This problem is even more evident in the super-AGB computations by Ventura & D’Antona (2011), where the algorithm implemented for the mass-loss rate (Bloecker 1995) leads to smaller oxygen depletion than in the most massive AGB models.\(^2\)

\(^2\) This does not happen in the SAGB models by Siess (2010) and Doherty et al. (2014), which show a very large oxygen depletion, but the same models have the (usual) problem that the sodium is destroyed instead of being enhanced – as discussed in Section 2.3.

Unfortunately, there are no observational tests for constraining the mass-loss in super-AGBs.

Fig. 1 shows the boundaries in $T_{\text{HBB}}$ necessary to process some important elements by proton capture. The boundary for the operation of the full CNO cycle, with an efficient conversion of oxygen into nitrogen, is at about 90–95 MK. At $\sim 100$ MK, the proton captures on $^{24}\text{Mg}$ begin to be efficient.\(^3\) The production of silicon requires even larger $T_{\text{HBB}}$ of $\sim 110$ MK, and the possible potassium production advocated by Ventura et al. (2012) requires $T_{\text{HBB}} > 120$ MK. In addition, adjustments of the relevant cross-sections for these proton captures are required in the models.\(^4\)

\(^3\) This temperature looks much larger than the $T \approx 75$ MK quoted above as temperature for Mg burning. Notice that we are not dealing here with stellar interiors: the densities at the bottom of the convective envelope ($\sim 10-15$ g cm$^{-3}$) are comparable to or larger than the density in massive stars’ interiors, but the time-scale for the entire AGB evolution is shorter.

\(^4\) For intermediate-mass AGBs, the 3DU is actually a spontaneous event (Iben & Renzini 1983), which occurs in the stellar structure without any further assumptions. In smaller mass stars, it is generally achieved through a further parametrization of envelope boundary overshooting.
Dilution is also required to explain the lithium versus sodium anticorrelation in NGC 6397 (D’Antona et al. 2011, and references therein).

The requirement of dilution between ejecta and pristine gas opens a number of interesting and complex questions concerning the gas dynamics during the cluster formation phase and the details of the formation and early dynamical evolution allowing the necessary mixing of pristine gas with AGB ejecta (see e.g. Trenti, Padoan & Jimenez 2015). Here, however, as discussed in the introduction, our attention is focused on the chemistry of SG populations which, at present, provides the strongest observational constraints on the possible sources of polluted material.

It is important to point out that efforts aimed at assessing the viability of the AGB models with dilutions are often based on oversimplified assumptions. Key fingerprints of multiple populations, such as the O–Na anticorrelation, have been often modelled by assuming that all the polluting gas has the same chemical composition, and that the different abundances are due to different ‘dilution’ with pristine gas characterized by the original (FG) composition (e.g. Conroy 2012). This hypothesis can be reasonable only in two cases: (1) if the formation time-scale of the SG is very short, so that the ejecta composition has no time to appreciably change; (2) if the AGB gas accumulates in the cluster for a long time and is well mixed, until re-accretion of pristine matter induces a single SF episode, with stars forming with different degrees of dilution between the AGB ejecta and the accreting gas.

The widespread presence of discrete populations which is likely to be due to multiple bursts of SF is a strong indication that not only different dilutions, but also different AGB ejecta compositions are at play. Thus, for example, the observational result that a single dilution curve cannot explain the Al–Mg data in NGC 6752 (Carretta et al. 2012) and NGC 2808 (Carretta 2014) provides further support to this hypothesis.

### 4 MULTIPLE GENERATIONS: A SEQUENCE IN TIME

Summarizing the result of Fig. 1, multiple populations in the AGB scenario are the result of SF in matter contaminated by proton capture nucleosynthesis, in a variety of physical conditions, which may include – or not – some light nuclei. We may think that the SF proceeds for several tens of Myr, although it must definitely halt before the AGB ejecta become richer in C than in O. D’Ercole et al. (2008) proposed and showed that the end of SF may be linked to the onset of SN Ia explosions, but other mechanisms may be considered too.

In D’Ercole et al. (2008) SF proceeds in a continuous way, as the conditions for SF are met in the cooling flow, in the absence of perturbing events. Looking at the clustering of stars in the Milone et al. (2015a) CNO two-colour plane, or at the distribution of abundances in Carretta (2015), it seems more likely that different populations are formed in separate bursts of formation (D’Ercole et al. 2012), following triggering events which remain to be investigated. Separate bursts may also be thought of as an indeed continuous SF, going on until some perturbing event stops it for a while, and then SF resumes again with different modalities (e.g. the degree of dilution).

The duration and modalities of the formation events will vary from cluster to cluster. In NGC 2808, the process must have lasted long enough to lead to the formation of at least five populations differing in chemical abundances.

We have already proposed some models able to reproduce some of the key properties of multiple populations in NGC 2808 (D’Ercole et al. 2012).
et al. 2008, 2010, 2012). More recent observations confirm the presence of the three main helium groups, which in Milone et al. (2015a) are identified as the FG (B), the ‘extreme’ population born from pure AGB ejecta (E), followed in time by an ‘intermediate’-helium population (D), in which re-accretion of pristine gas plays a role.

Is it possible to extend the model in order to explain populations C and A? We show that this is qualitatively possible. In the following, a full chemical and dynamical model, which is beyond the scope of this paper, will be necessary to reconstruct the whole formation history of NGC 2808, along with the initial mass and structural properties necessary to reproduce all its current observed properties.

### 4.1 A note on the oxygen yields

The discovery that NGC 2808 hosts three MS (D’Antona et al. 2005; Piotto et al. 2007), which could only be interpreted in terms of difference in the initial helium content in the forming gas, was one of the important drivers to interpret the chemical anomalies of GC stars in terms of different stellar generations. Milone et al. (2012a) estimate a helium difference $\Delta Y \sim 0.13$ between the red and the blue MS, using standard stellar models not including any iron variation. This provides $Y \sim 0.38$, if the red MS is located at a standard $Y \sim 0.25$. The helium difference between the MS A and E is estimated again to be $\Delta Y \sim 0.13$ in Milone et al. (2015a) – their table 3 – but the FG is now identified with the B population, and $\Delta Y$ is $\sim 0.10$ between E and B (Table 1, lower panel), giving $Y \sim 0.35$ for the blue MS. The AGB models’ maximum helium yield is $Y \sim 0.36$ in our models (see Fig. 3) very similar to the result by Doherty et al. (2014). Values up to $Y = 0.375$ are found in the super-AGB models by Siess (2010). Therefore, if the E population forms from the most massive pure AGB ejecta, as we will assume, there is good agreement between observations and models.

On the other hand, the Ventura et al. (2013) models have a difficulty in reproducing the most extreme oxygen abundances. This is a long-standing problem. When super-AGB models were not yet available, D’Ercole et al. (2010) assumed semi-empirical yields with low O in their first modellisation. But the yields of super-AGB stars resulted in oxygen abundance larger than those of the most massive AGB (Ventura & D’Antona 2011), so D’Ercole et al. (2012) proposed that the very low oxygen abundances were due to deep mixing in the giants of the extreme population, favoured by their large helium abundance (D’Antona & Ventura 2007). In this work, we make the same assumption but we point out that, as discussed in Section 2.3, alternative solutions are possible: a larger oxygen depletion may be achieved by reducing the mass-loss rate in the models (Ventura et al. 2011), a reduction which must go together with a reduction of the cross-section $^{23}\text{Na}(p,\alpha)^{20}\text{Ne}$.

Fig. 2 shows the O and Na yields computed by Ventura et al. (2013), and the different oxygen values assumed in this work for $M \geq 6 \, M_{\odot}$.

### 4.2 Outline of the BEDCA model

The Milone et al. (2015a) and Carretta (2015) data imply that SF in the cluster occurs in separate bursts. We use the main observational

---

**Figure 3.** Each panel shows the average abundance of nitrogen (top left), helium (bottom left), oxygen (top centre), sodium (bottom centre), magnesium (top right) and carbon (bottom right) versus time in the ejecta of massive AGB stars having $Z = 10^{-3}$ (Ventura et al. 2013). The yield for each initial mass (8, 7.5, 7, 6.5, 6.3, 6.0, 5.5, 5.0 $4.5 \, M_{\odot}$, from left to right) is shown by dots. The abundances of oxygen at $t < 7 \times 10^{7}$ yr have been adjusted as explained in the text. The pure ejecta abundances are shown as a red line with squares. The initial abundances in the AGB stars are represented by the horizontal red lines without dots. Abundances in diluted matter are shown by the brown lines with squares. Starting from the initial FG abundance lines, the AGB gas percentage in each curve is 0.05, 0.1, 0.2, 0.3, . . . , 0.9. In these planes, we show our guess for the location of the five populations defined by Milone et al. (2015a) for the cluster NGC 2808: (1) B (red) is the FG population, having the initial gas abundances (red rectangles, stars born at $t = 0$); (2) E is the most extreme SG, born from pure AGB ejecta, defined by the blue rectangles; (3) D is the ‘intermediate’ SG, in which the AGB ejecta constitute 40–60 per cent of the forming gas (cyan regions); (4) C, the new population discovered in Milone et al. (2015a), with normal helium and high nitrogen: this is obtained by very strong dilution of the AGB ejecta which already have a strong effect of 3DU, with CNO enhancement; (5) A, the small population at the red of all other sequences, is here identified with a remnant SF, mostly from FG gas contaminated by the ejecta of one SN Ia (green circle). The Mg boxes for population E and D (blue and cyan) in the bottom-right panel show with arrows the observed location in the data by Carretta (2015), $\sim 0.15$ dex lower than the models. As explained in the text, a lower mass-loss rate would provide the correct result, but it would reduce the sodium abundances.

---

*Note:* We again have to emphasize that massive star yields do not provide a better fit to the oxygen extreme abundances. The matter ejected from the FRMS can be highly depleted in oxygen in the inner stellar layers processed by core H-burning. But, when formation occurs in the ‘excretion’ disc around the star (Chantereau et al. 2015), the helium abundance in these stars will be much larger (up to $Y = 0.8$) than the values which are attributed to the extreme stars from the MS split. There is no observational evidence for the presence of such stars, and the helium richest MS are clustered, with a small scatter, below $Y \sim 0.40$. 
constraints to build a very simple model and outline the possible epochs during which each of the populations can be born, and the corresponding average values of dilution necessary to reproduce the available abundance patterns (Table 1). As we have discussed, the yields of the models we are using will definitely require further revisions. Anyway, the fundamental properties of the yields, which are linked to the evolution with time of $T_{\text{HBB}}$, emerge clearly, as we are going to show.

The first part of this interpretation was outlined and defined with more complete models by D’Ercole et al. (2008, 2010, 2012), and we start from what emerged from those investigations. The novelty is the possible interpretation of the two new populations (C and A) identified by Milone et al. (2015a), in the context of that same model. The key ingredients of our results are summarized in Fig. 3. The starting abundances of $N$, $He$, $O$, Na, $C$ and Mg in the models by Ventura et al. (2013) are shown as horizontal lines. We also plot the abundances of the same elements in the yields of different masses, for $Z = 0.001$, $\alpha/Fe = 0.4$.

The initial abundances and the yields delimit the region of abundances in ‘diluted’ gas. Starting from the pure yield lines, we plot the lines for mixtures in which the AGB gas constitutes a fraction of $90, 80, 70, 60, 50, 40, 30, 20, 10$ and $5$ per cent. On these schemes, we build our interpretation of the BEDCA sequence summarized below, and explained better in the following sections.

**B:** both the chemical abundances of all light elements (Carretta 2015) and the location in the CNO two-colour diagram show that these are the FG stars. Thus, we represent it by red boxes at the left of the AGB epoch ($t < 30$ Myr), placed on the initial abundances.

**E:** this group has been early recognized as the one in which nuclear processing has been most extreme, with large depletion of oxygen, processing of Mg, production of Si and of K. Thus, in AGB modelling, this requires the largest $T_{\text{HBB}}$’s, and we place it at ages 40–60 Myr. D’Ercole et al. (2008) first suggested to model the blue MS of NGC 2808 as due to SF in the pure ejecta of the most massive AGBs, in which $Y$ is indeed very similar and very large ($Y = 0.35-0.38$ in all published stellar models – Siess 2010; Ventura et al. 2013; Doherty et al. 2014). Chemical modelling by D’Ercole et al. (2010, 2012) showed that also the other chemical abundances could be reproduced under this hypothesis. Thus, we schematically represent group E with the blue boxes, lying on the AGB yield line. The masses involved are the super-AGB-masses from 6–6.5 to $8 \, M_\odot$.

We emphasize again that these values of mass and ages apply to the models we use here; the ranges may be different for other models. The main characteristic of this population is that it is born from the ejecta of the first (most massive) AGBs born after the end of the SN II era. The peculiar carbon yield of super-AGB show in Fig. 3 (bottom-right panel) is discussed in Section 4.4.

**D:** we propose that the discontinuity among group E and D is due to a sudden onset of dilution with pristine gas, at an epoch $\sim 60$ Myr after the cluster formation (D’Ercole et al. 2010, 2012). Fig. 3 suggests SF in a mixture in which the AGB ejecta constitute 40–70 per cent, and extending for $\sim 25$ Myr (cyan boxes). This again is in line with our previous chemical evolution models.

**C:** a discontinuity in the SF must occur at a time 85–90 Myr, followed by an additional SF event in gas which is scarcely contaminated by the AGB ejecta (5–15 per cent). A time of formation from about $t = 90$ to about $t = 110$ Myr is suggested in the figure, and is marked by the yellow boxes. With such a huge dilution, the abundances go down to values very similar to the FG values. $Y$ may be larger than the initial value by $\sim 0.01–0.02$ (an important hint, which we discuss in Section 5). Na is in the range from pristine to $\delta$Na = +0.2 dex, as observed, but, most importantly, nitrogen may be larger than the pristine abundance by 0.4–0.7 dex, as predicted the differential analysis between the spectrophotometry of groups B and C by Milone et al. (2015a). We emphasize that such a population can be in part hidden and overlap with the FG in the Na–O plane (group primordial 2 by Carretta 2015) and can be fully identified as a separate SG population only when studied with spectrophotometric or spectroscopic observations sensitive to the N abundance.

A: we suggest that population A is formed towards the end of the C group formation epoch, from mainly pristine gas polluted by the ejecta of the first SN Ia exploding in the cluster. We are not guided by spectroscopic data, but by considering the following. Although the ‘A’ MS data are not well separated in the CNO two-colour diagram, Milone et al. (2015a) show that it may be slightly redder than the B MS, and suggest that these are stars of the FG affected by a slight iron increase ($\delta$[Fe/H] $\sim 0.1$ dex) attributed to SN II (but see also Section 2.1). Here we revise their interpretation (see Section 4.5).

We identify group A in Fig. 3 with green circles, whose location in time and composition are of course very arbitrary. Actually, their formation could be contemporaneous to group C formation, in a fraction of gas enriched by the first SN Ia. If A stars are the last one to be born in the cluster, they too might have some N enhancement, and this could be the reason for their location in the CNO two-colour diagram. Spectroscopic observations are necessary for this group.

### 4.3 Abundances of other light elements

The abundances of carbon are shown in the bottom-right panel of Fig. 3. Carbon is always very depleted in the ejecta, apart from the most massive super-AGBs, where it reaches $\sim 0.2$ dex. Several computations show that, at low metallicity, the 2DU occurs in the ‘dredge-out’ modality (Iben, Ritossa & García-Berro 1997; Siess 2006; Siess & Pumo 2006). During the inward penetration of the base of the external envelope, a convective zone first forms within the helium-burning, then, after growing in mass, it merges with the surface convective region. This process favours a significant increase in the surface carbon. This kind of event is much less studied than HBB or the 3DU, and so it must be considered uncertain. Nevertheless, the peculiar dredge-out mechanism occurs only in the most massive models, and we can expect an enhanced C-abundance in the (small) fraction of population E born from these ejecta. If this is not found, it is possible that these stars evolve into e-capture supernovae (in this context, see the discussion in Pumo, D’Antona & Ventura 2008).

We show in Fig. 3 the abundances of magnesium (top-right panel). The agreement with the spectroscopic result is only qualitative here. As we mentioned in Section 2.3, Carretta (2015) has shown that also the Mg abundances in NGC 2808 indicate the presence of five groups. Quantitatively, our models fail to reproduce the Mg depletion of 0.4 dex shown by these data for the most extreme stars (fig. 8 in the Carretta paper, and Table 1). The maximum depletion of the models is reached for the mass $6 \, M_\odot$, and amounts to 0.17 dex, which would mean $\leq 0.1$ dex if we operate a dilution by $\sim 50$ per cent with pristine gas. We discussed in Section 3.1 that this problem is due to the mass-loss rates of the models, and these, in turn, depend.

---

Footnote 7: The oxygen yield has been modified according to Fig. 2.
on the necessity of preserving sodium. Therefore, in Fig. 3 we add arrows to the boxes, which represent the group E and D locations from the models. The ends of the arrows indicate where the boxes should be placed to be consistent with the data.

We have not discussed specifically aluminium, which the Carretta (2015) data show to cover a wide range of abundances $\sim$1.1 dex. Also our models display an abundance range of a factor of $\sim$10. A discrepancy with the data is present only for group E. Also our models display an abundance range of a factor of $\sim$10.

Figure 4. Abundances of CNO in the ejecta. The figure adopts the same display as Fig. 3.

4.4 The CNO content of population C

Fig. 4 is complementary to Fig. 3 and displays the time evolution of the C+N+O yields in the Ventura et al. (2013) data, and their dilution curves, as a function of time. The C+N+O increase at small ages is due to the peculiar ‘dredge-out’ of carbon in the most massive super-AGBs, previously discussed in Section 4.3. The increase in CNO at ages $>80$ Myr is an effect of the 3DU, already discussed in Section 3.1. Although CNO increases rapidly with decreasing mass, C is still CN processed in the envelope, so the dominant increase is that of the N abundance, which we have seen in Fig. 3. We claim that the N increase due to the 3DU is allowing us to identify stars belonging to the ‘C’ group, and separate them from the FG stars of group B.

Notice that the processing of N inside the helium inter-shell (at each episode of dredge-up) has two effects:

1. It provides the neutrons for s-process from the chain $^{14}\text{N}(\alpha, \gamma)^{18}\text{F}(\beta, \nu)^{18}\text{O}(\alpha, \gamma)^{22}\text{Ne}(\alpha, \nu)^{25}\text{Mg}$ so that also s-process enhancement begins in the envelope;
2. The primary $^{22}\text{Ne}$ dredged-up will capture protons and further increase Na in the envelope. The sodium increase may be seen in the C group, in spite of the huge dilution of the ejecta. In contrast, we do not expect that the s-process enhancement is seen with such a huge dilution.

In our schematic model, and taking at face values the uncertain CNO increase of Ventura et al. (2013) models, the total CNO in group C should be a factor of 1.1–1.3 larger than the initial value.

5 THE HB MORPHOLOGY IN NGC 2808

5.1 The ‘C’ group location in the HB

The abundance analysis of the HB stars had already pointed out the existence of population C. The B (primordial) and C (late SG) groups can be identified with the two peaks in the [O/Na] spectroscopic distribution by Carretta (2015), at [O/Na] $> 0$. Carretta shows that the high-[O/Na] stars are divided into two subcomponents, with small difference in [O/Fe] and larger differences in Na. He also correlates the presence of these two groups with the O–Na distribution of stars in the red side of the horizontal branch (RHB) of the cluster.

In the first attempts to model the HB of NGC 2808 with multiple populations (D’Antona & Caloi 2004), it was proposed that the RHB should contain only standard helium, FG, stars, and that the blue side of the HB should be populated by He richer stars of SG, down to the extreme HB. The hottest group would contain very He rich star progeny of the blue MS (D’Antona et al. 2005; D’Antona & Caloi 2008), a result also confirmed by Dalessandro et al. (2011). Later on, Gratton et al. (2011) found that an O–Na anticorrelation (although not fully extended) is instead present also among the RHB stars. On this basis, Gratton et al. were the first to suggest that more than three populations were present in the cluster.

A subsequent spectroscopic analysis of RHB stars (Marino et al. 2014) showed a more marked dichotomy in Na, and that the group richer in Na appeared to be slightly more luminous and bluer than the red group. Also this work attributes this feature to the presence of another population hidden among the RHB stars, and explicitly relates it to a group in which the AGB ejecta have suffered more
Figure 5. In the lower part of the figure, the mass–$T_{\text{eff}}$ relations for ZAHB models of $Z = 0.002$ are shown for $Y = 0.25$ (top line, black) and $Y = 0.28$ (red line) for standard C, N and O abundances. The two dash–dotted lines below represent the mass–$T_{\text{eff}}$ location of models in which the total CNO is 1.5 the standard CNO, for $Y = 0.25$ (upper line, blue) and $Y = 0.28$ (lower line, magenta). At the top, the four horizontal lines mark the value of mass in evolution in the RGB at age of 12 Gyr, with different $Y$, as labelled (the evolving mass does not change for different total CNO). The vertical segments represent a constant mass-loss of 0.22 $M_\odot$/yr, and thus the location along the ZAHB obtained for different $Y$. At constant CNO, even a very small helium increase by 0.01 shifts the location from the red clump to beyond the RR Lyr gap. An increase in $T_{\text{eff}}$, we can appreciate that the mass-loss spread in D'Antona & Caloi (2004) and that the scarcely populated RR Lyr gap in NGC 2808 (Clement & Hazen 1989) requires a jump in $Y$ by $\sim 0.01$ to be modelled (D’Antona & Caloi 2004). Thus, if the C group stars had only a slightly enhanced $Y = 0.26$ or even a bit larger, they should fall at the left of the RR Lyr region. An increase in CNO for the same $Y = 0.25$ shifts the model to cooler location (blue dot), contrary to the observations. A small increase in $Y$ in CNO-rich models brings back the model to the $T_{\text{eff}}$ range of B red clump stars (orange dot).

6 CLUSTER-TO-CLUSTER DIFFERENCES EXPECTED AT DIFFERENT METALLICITY

Before we discuss the general case of multiple-population formation, we make few simple considerations on the literature data which already show that an analysis based on AGB ejecta pollution satisfies some basic observations. Fig. 1 shows that the maximum $T_{\text{HBB}}$ reached in AGBs depends on the metallicity of the models. This is a first-order result, which basically depends mainly on the opacities in the envelope: more transparent envelopes (lower $Z$) allow larger $T_{\text{HBB}}$. A very simple consequence is that higher metallicity clusters cannot show the same degree of p-capture processing as the lower metallicity clusters. Of course, there may be low-metallicity clusters which do not show signs of the presence of extreme populations; this depends on the perturbing events, discussed in Section 7 which modulate the SFH. But, for similar conditions in the SF time-scale, that is, when similar masses contribute to the ejecta, we do not expect to find the same extreme compositions in SG of clusters of higher $Z$. 

$^8$ Following the D’Ercole et al. scenario, we suppose that they are the latest stars to have been formed from highly diluted material, such that their abundances in light elements and $Y$ approach the primordial values of the FG. Hence, these stars could have formed after the intermediate SG stars, which show evidence for a higher degree of AGB pollution. If this prediction is correct, the abundance pattern of light elements of these stars is dominated by dilution with pristine gas, and the red MS and the RHB contain the first and last stars formed in the cluster.’ (Marino et al. 2014).
We list a few results lending support to this point.

(i) Potassium, which requires \( T_{\text{HBB}} > 125 \text{ MK} \) to be formed by proton captures on argon (Ventura et al. 2012), is found to vary only in NGC 2419 (Cohen & Kirby 2012) and in NGC 2808 (Mucciarelli et al. 2015). Notice that both clusters have an extreme population (for NGC 2419, see Di Criscienzo et al. 2015), which we interpret as born from pure ejecta, and thus are best qualified to show such p-capture extreme product. None of other examined clusters show variations (Carretta et al. 2013); in particular, no hint of variation is present in 47 Tuc.

(ii) In 47 Tuc (\( \text{[Fe/H]} \sim -0.7 \)), the maximum magnesium processing is by \( 2\text{Mg} = -0.1 \) (Thygesen et al. 2014).

(iii) Aluminium varies by a decade in low- and intermediate-metallicity clusters, but it looks like its variations are much smaller at larger metallicity; see the discussion in Cordero et al. (2015).

7 THE GENERAL FRAMEWORK

We conclude that the populations in NGC 2808 were formed in the order of the acronym BEDCA, and that their characteristics are consistent with the main features of each of these populations, including a good explanation for group C, and a working hypothesis for the role of SN Ia in group A. Thus, the multiple populations in GCs are recognized to be the result of SF events which may be separated by the occurrence of different triggering events, allowing transitions in the SF resulting in chemical differences between the groups. The \( T_{\text{HBB}} \)'s reached by the models are also large enough to allow the contemporary processing of magnesium, and formation of aluminium and silicon, as found in the most extreme populations of NGC 2808. Although the yields of the models adopted do not reproduce the quantitative depletion in the abundance of Mg, or the entire production of Si, we have discussed how this problem can be solved thanks to a smaller mass-loss rate. This can be assumed only if the cross-section \( ^{23}\text{Na}p,\alpha^{20}\text{Ne} \) is smaller, so that sodium can be preserved, at high \( T_{\text{HBB}} \), also with a longer HBB evolution.

It is clear that we can further generalize the scheme shown in Figs 3 and 4, by further subdividing the SF events, as probably required by the finer discreteness of the data. For example, we schematically represented five different boxes of chemical properties for the five groups defined in Milone et al. (2015a) and attributed the separation between the groups to triggering events (onset of dilution, delayed SN II explosions). As already mentioned, within these boxes other minor separations can be present, due to other triggers, so the model can be refined once we know better the amount of discreteness in the data.

We now address two more general questions.

(1) What are the possible events which provide the transition between the different SF epochs?

(2) How can the NGC 2808 scheme be adapted to predict the behaviour of the chemical differences among multiple populations in other clusters?

The specificity of these patterns depends primarily on the metallicity of the cluster stars. The presence of separate bursts depends on other possible parameters, intertwined with each other. We consider specifically

(i) the timing and extent of dilution of nuclearily processed AGB ejecta with pristine gas;
(ii) the role of binary SN II in delaying or halting the formation of the SG;

Figure 6. The mass evolving as a function of the age is shown, from Ventura et al. 2013. Times are dependent on the assumptions made in the models, especially on the core overshooting during the core H-burning phase. The end of the SN II epoch in these models is at \( \sim 40 \text{ Myr} \), and is followed by the epoch of super-AGB evolution which lasts \( \sim 20 \text{ Myr} \). Delayed SN II may explode until the evolutionary age of the primary mass which may provoke the explosion of the companion, by increasing its mass beyond the single SN II minimum mass. The age at which SN Ia begin to explode is uncertain, but it is dependent on the age at the end of the super-AGB phase, which tags the beginning of formation of C–O white dwarfs.

(iii) the role of binary SN II and/or of the first isolated SN Ia explosions in the formation of an iron richer SG;
(iv) the role of SN Ia to end the epoch of SF.

To further study this problem, we show in Fig. 6 the mass versus time evolution of super-AGB and AGB models, together with some limiting epochs for important events. We stress again that the timing of different epochs is based on the evolutionary times of the models we are using throughout this paper, by Ventura et al. (2013), and that the limits may be shifted if other models are adopted, while the general description of the epochs remains qualitatively similar.

7.1 The SN II epoch boundary (boundaries)

The first (red) boundary in time in Fig. 6 divides the epoch of single stars’ SN II explosions from the epoch of the quiet AGB evolution. A few further explosions may occur close in time to this limit, if in the most massive super-AGB stars the convective mantle is not fully consumed before e-captures begin inside the core (e.g. Siess & Pumo 2006; Poelarends et al. 2008). No cooling flow is possible earlier than this boundary.

In principle, SN II explosions may continue also at a later time, if the critical mass for explosion (in our example, \( M_{\text{mass}} = 8 \text{ M}_\odot \)) is reached by mass transfer during binary evolution. The existence of this kind of evolutionary path is necessary to explain the presence of young non-recycled pulsars in an eccentric orbit with a
companion white dwarf (WD; van Kerkwijk & Kulkarni 1999). In these systems, the neutron star formation must have occurred after the formation of the WD, through mass accretion on the lighter companion (Portegies Zwart & Yungelson 1999; Tauris & Sennels 2000). D’Antona et al. (2005) suggested that binary SN II may influence the formation of multiple populations in GCs. Here we use the simple-minded assumption made in this latter work to put a strict time limit to binary SN II explosions, by noticing that the primary star evolving in the binary may be assumed, in the best case, to transfer all its envelope, leaving a white dwarf remnant of mass $M_{WD}$. Until the evolving primary $M_1$ is massive enough, mass transfer may push the mass accreting component beyond $M_{\text{mass}}$. The end of such delayed SN II is set by the time of evolution of the minimum possible mass, which, in the best case, is $M_1 > (M_{\text{mass}} + M_{WD})/2$. The limiting time for the possible occurrence of these events in our stellar models is then set by the evolution of the 4.5 $M_{\odot}$ star. This is a very naive schematization; for a more complete description see, e.g., Tauris & Sennels (2000).

An estimate of the impact of binary evolution on late SN II explosion is out of scope of the present work. The primordial binary fraction for intermediate-mass stars of the FG, their orbital period distribution and the different possible paths to achieve $M_2$ (final) $> M_{\text{mass}}$ (Tutukov & Yungelson 1993; Tauris & Sennels 2000) are critical in determining the extent of the role played by late SN II during the SG formation. Also the time span of these possible events is strongly dependent on the precise value of $M_{\text{mass}}$, which not only sets the age of the end of the single SN II era, but also determines the minimum donor mass (and maximum age) which can provide this kind of evolution with mass exchange.

The actual number and frequency of delayed SN II in a cluster strongly depend on many different parameters. In some cases, they may occur already during the evolution of the highest AGB–super-AGB masses, which provide the strongest nuclear processing by $p$-captures. Thus, the abundance patterns of populations E and D in NGC 2808 require that such explosions did not perturb the cooling flow in this cluster, but they may have been important in clusters which show mild O–Na anticorrelations such as M4 (Marino et al. 2011a). In contrast, in NGC 2419, for which the results of Di Criscienzo et al. (2015) and Cohen & Kirby (2012) suggest that the SG is entirely or mostly ‘extreme’ (i.e. made up from undiluted ejecta), it may be possible that, after this extreme population is formed, an intense delayed SN II epoch begins, followed, without interruptions, by the SN Ia epoch. In this case, the intermediate population born from re-accretion of pristine gas diluting the ejecta is not present at all.

We will show that the most interesting case may be an intermediate one: the observed properties of multiple populations of an entire category of clusters may be explained in a simple way, if the delayed SN II explosions indeed have occurred and inhibited SF for a few tens of Myr (Section 8.1).

### 7.2 The super-AGB epoch end and the formation of SN Ia progenitors

The subdividing (blue) line between super-AGB and AGB evolution in Fig. 6 marks the formation of the first C–O WDs. Mass transfer in binaries which contain a mass accreting remnant of super-AGBs, an O–Ne core WD, may cause electron capture supernovae (e.g. Miyaji & Nomoto 1987; Gutierrez et al. 1996), but these events should be much less energetic than SN Ia (Dessart et al. 2006). A very quiet epoch in the cluster life, during which the postulated SF from super-AGB and AGB ejecta may occur, in the absence of any other perturbing energy sources, may plausibly exist in clusters.

The mechanisms proposed for SN Ia explosion all involve one or two C–O WDs. Historically, two models have been discussed, the single-degenerate (e.g. Nomoto 1982) and double-degenerate (DD; e.g. Iben & Tutukov 1984; Webbink 1984) model. In the DD model, the SN Ia is caused by the merger of two C–O WDs, the combined mass of which equals or exceeds the Chandrasekhar mass. The detonation mechanism of DDs has been debated for decades, and today rotation is considered the key ingredient in models of this process. Rotation allows mass accumulation up to the explosive central ignition of carbon, which in turn leads to SN Ia (Piersanti et al. 2003). The DD case occurs earlier during the galaxy evolution.

Consequently, the SN Ia epoch cannot predate the time at which the first massive C–O WDs form,10 plus a further time delay necessary to achieve the mass transfer by which this WD attains the explosion conditions (e.g. Madau, della Valle & Panagia 1998). Cluster-to-cluster variations in the onset of the SN Ia epoch will lead to differences in the SF modalities of the latest multiple populations. In particular, it is even possible that the SG formation stops before the CNO and s-process enhanced ejecta begin to be dominant.

### 7.3 The SN Ia epoch and the (final) end of SG SF

The observational constraints on SN Ia frequency and its probable starting epoch (Mannucci et al. 2005; Totani et al. 2008) are based on observations of different types of galaxies, so it is necessary to understand how these findings may be transferred to the study of GCs.

Recent observations have shown that the peak of frequency of events is, at early times, just somewhat above 10$^9$ yr, in galaxies still subject to SF (Mannucci et al. 2005; Totani et al. 2008). The Totani et al. (2008) study shows the beginning of the SN Ia data at a ‘delay time’ interval 100–250 Myr.11 In this context, the DD model is able to provide large rates at early times (Mennenkens et al. 2010), and this might be particularly true in GCs, where we can expect to have many close binaries with two CO WD components, which can merge after a short delay time, thanks to the spiral-in caused by the emission of gravitational wave radiation.

So the time-scale for the onset of the SN Ia epoch in early GCs is in the same age range of the end of the binary delayed SN II epoch, and in the middle of the CNO- and s-enhanced AGB ejecta. Possible – and probable – variations in the role of these ingredients are likely to affect cluster-to-cluster differences in the properties of multiple populations.

### 7.4 The epoch of C+N+O and s-process enriched AGB ejecta

Fig. 6 shows that the timing of production of CNO- and s-process-rich AGB ejecta, the delayed SN II epoch and the beginning of the

---

10 Recent works have examined the possibility that carbon burning rates are very different (either larger or smaller) from the recommended values (e.g. Chen et al. 2014). This assumption leads to the formation of O–Ne–C cores which would also explode as SN Ia if their mass increases up to the Chandrasekhar mass by mass transfer. This is still only a working hypothesis, so we assume that it does not alter our main conclusion.

11 The delay time here is ‘the delay time from SF’, so it is the total age, if we are dealing with a burst of SF. In this nomenclature, the delay time in Fig. 6 is 10$^9$ yr, while the true delay time between the first CO WD formation and the SN Ia epoch is (100–65) $\approx 35$ Myr.
SN Ia epoch are partially overlapping, so that their roles cannot be examined separately. On the other hand, clusters in which CNO and s-process enhancement is present do exist. In the early years of observations of chemical anomalies in GCs, it appeared that the total C+N+O abundance was constant in the individual clusters examined, and that the C, N and O variations were to be ascribed to the action of the CNO cycle, and in particular of the ON branch, which accounted for the oxygen reduction in the SG stars. Examples of such clusters are M92 (Pilachowski 1988); NGC 288 and NGC 362 (Dickens et al. 1991); M3 and M13 (Smith et al. 1996); M4 (Ivans et al. 1999); NGC 6752 (Yong, Grundahl & Norris 2015).

More recently, the situation became more complex. First, a simple explanation of the splitting found in the subgiant branch (SGB) of the cluster NGC 1851 (Milone et al. 2008) was more easily justified by assuming that its SG (populating the dimmer branch) had a larger CNO than the bright branch, but a similar age (Cassisi et al. 2008; Ventura et al. 2009). Several other clusters have been shown to hold a double SGB (Piotto et al. 2012). The spectroscopic enrichment in total C is established only in NGC 1851 (Yong et al. 2009, 2015) and M22 (Marino et al. 2012a); see also Lim et al. (2015b). In clusters showing a split SGB, this feature is accompanied by an increase in s-process element abundance, which is in line with the 3DU interpretation (Straniero, Cristallo & Piersanti 2014).

Further investigations are necessary, to explore the processes which may explain why some clusters show two main bursts of SF, separated by a few tens of Myr to allow the CNO-, s-enhanced ejecta production.

8 CLUSTERS WITH IRON SPREAD OR BIMODALITY

Recent spectroscopic observations have found a few GCs characterized by an internal [Fe/H] abundance dispersion. As suggested by Da Costa (2015), these systems might be the former nuclear star clusters of now disrupted dwarf galaxies (Willman & Strader 2012; Marino et al. 2015). An iron difference among cluster stars implies that the cluster was able to retain at least some of the supernova ejecta. This may occur either in dwarf galaxies, where possibly dark matter is initially present, or in particularly massive clusters. The spread in iron observed ranges from tiny differences, as in NGC 1851, to the very large differences observed in ω Cen.

In all these clusters, the typical signature of p-capture element variations (the O–Na anticorrelation) are accompanied by metal enrichment, CNO and s-process enrichment together. The category of clusters named ‘s-Fe-anomalous’ described by Marino et al. (2015) includes mainly clusters with split SGB, namely NGC 1851 (Carretta et al. 2011; Gratton et al. 2012), NGC 5286 (Marino et al. 2015), M22 (Marino et al. 2009) and M2 (Yong et al. 2014), plus ω Cen (Johnson et al. 2009; Marino et al. 2011b).

8.1 Pollution by delayed SN II: the s-Fe-anomalous clusters

Here we propose how a natural extension of the standard GC model for the formation of multiple populations can deal with the s-Fe-anomalous clusters. In these clusters, the O–Na anticorrelation is already present in the fraction of stars having the smaller, and homogeneous, iron content. This anticorrelation is very unlikely to occur, if the iron contamination is due to SN II of the single SN II epoch. So we suggest that the first phases of evolution of these clusters are similar to what happens in GCs which are fully chemically homogeneous in heavy elements, but, at later times, we must account for contamination by further supernova explosions, those from delayed SN II in binaries. Let us assume that

(i) delayed SN II do not explode before the first SG, homogeneous in iron, forms, providing a typical own O–Na anticorrelation;
(ii) afterwards, delayed SN II begin exploding with some regularity in a cluster, destroy the cooling flow, which would have included both AGB ejecta and pristine gas;
(iii) these events (which are much less frequent than the single SN II were) are not able to inject into the gas enough power to definitely push it out of the cluster vicinity.

When, at last, the delayed events become rare, possibly several tens of Myr later, the pristine plus AGB gas will re-accrete and induce a new SF formation burst. In these hypotheses, the pristine gas will now be contaminated by the delayed SN II ejecta, and by the AGB ejecta which were lost during this time span. Further, the contaminating AGBs will be the masses in which the 3DU has been very effective, so this population will have the characteristics of the ‘s-Fe-anomalous’ clusters: larger iron and s-process abundances, and associated C+N+O enhancement.

An appealing feature of this scenario is that it explains the features of the anomalous clusters, without the need for additional hypotheses, such as the merging of two different clusters. Dynamical models are anyway required to test this suggestion, and are under way (D’Ercole et al. 2015, submitted to MNRAS).

The time gap of a few 10$^7$ yr between the formation of the ‘first’, standard SG, and the ‘second’ one, s-Fe and CNO enriched, also justifies another important characteristics of some clusters: the presence of separate SGBs. While this time break is negligible in terms of location of isochrones with identical chemical composition, this short time is sufficient to shift the AGB ejecta composition to the CNO-enriched stage, which, also with the help of the small iron increase, will result in distinct SGBs (for the case of NGC 1851, see Cassisi et al. 2008; Ventura et al. 2009). Finally, notice that the formation epoch of this s-Fe-CNO enriched SG cannot be very extended, as it occurs close to the beginning of the SN Ia era, which will definitely end SF.

8.2 Are there signatures of isolated episodes of delayed SN II explosions?

We remark here that there might be clusters in which one isolated episode or a few separated episodes of delayed SN II explosions occur. In this case, the explosion is able to halt only partially the cooling flow, and some of the gas flowing into the cluster core will be contaminated by the ejecta of this or these supernovae. This is an additional possibility to be taken into account when dealing with the formation of populations with abundance anomalies. Each single SN II cannot alter the iron content of the gas in a detectable way, but it can certainly increase the oxygen content, and also the Mg and Si content. The s-process and the C+N+O in this SF event may result increased or not for events which occur later or earlier during the time span allowed for the delayed SN II events.

It would be important to carry out observational studies aimed at searching the fingerprints of this scenario; signatures of O and Mg increase may be revealed by the combination of the UV and optical HST passbands of the UV Legacy Survey of Galactic GCs (Piotto et al. 2015); studies of the iron abundance will require very
accurate data sets and analysis, such as those of Yong et al. (2013) for the cluster NGC 6752, where a spread of 0.03 dex in [Fe/H] has been detected, while no spread in CNO seems to be present in this cluster (Yong et al. 2015).

8.3 Signatures from the SN Ia first random explosions

The SN Ia, unlike these delayed SN II, play a very different (sometimes double) role. In fact

1. first they might pollute the re-accreting gas, producing a more metal rich last population as we suggest for NGC 2808;
2. when the frequency of explosions becomes stationary (at an age of \( \gtrsim 10^9 \) yr), they are likely to completely halt the SF.

Before reaching a frequency of events large enough to halt the cooling flow and the SG formation, a few initial SN Ia explosions might occur well separated in time each other. D’Ercole et al. (2008) have shown that one SN Ia event, isolated in time, is not able to halt the cooling flow. A few events may be needed, especially if the cluster is suffering strong gas re-accretion. While this problem is to be further studied in more detail, we think that it is likely to lead to a variety of different outcomes. NGC 2808 remains a valid example possibility: we expect that pollution by iron from SN Ia in the pristine re-accreting gas (mixed, or not, with AGB ejecta) will result in the formation of stars which will be more metal rich.

9 A ZERO-ORDER ESTIMATE OF POLLUTION BY ONE SN Ia OR BY DELAYED SN II

For NGC 2808, we have estimated the amount of gas which can be polluted by a single SN Ia, assuming that the average \( \delta [\text{Fe/H}] \) is 0.1 dex larger than the iron content of the rest of cluster stars. This leads to a prediction of 48 000 M\(_{\odot}\) of polluted gas, in very good agreement with the 5.6 per cent of the total mass of the cluster, estimated for the ‘A’ population by Milone et al. (2015a) (about 47 700 M\(_{\odot}\) if we assume 8.5 \( \times 10^3 \) M\(_{\odot}\) for the entire cluster; McLaughlin & van der Marel 2005).

On the same grounds we may try to understand whether the other populations with larger iron content found in several clusters can be explained by pollution by a single SN Ia, assumed to eject 0.8 M\(_{\odot}\) of iron. Notice that the iron mass ejected by a low-mass SN II is about 1/10 of the iron mass of SN Ia (0.07 M\(_{\odot}\), according to Nomoto, Kobayashi & Tominaga 2013), so the schematic reasoning which follows can be applied to pollution either by one SN Ia or by 10 SN II. Anyway, in order to justify that the iron increase in the whole SG considered has a small spread, we must also think that there has been a global mixing of all the SN II ejecta with the pristine gas.

We show in Fig. 7 and Table 3 the result obtained under these simplified assumptions. We plot the lines of expected polluted mass as a function of the [Fe/H] of the standard cluster population, and of the \( \delta [\text{Fe/H}] \) of the anomalous population. We compare the predictions with the ‘observed’ mass, knowing the fraction of involved stars and the total mass of each cluster (Table 3). The two numbers are in reasonable agreement with the observations. In one case (the populations BI+BIi of M2 in Milone et al. 2015a), the expected mass is far larger than observed. In the other cases, it is comparable or modestly larger.

Should we then conclude that all clusters showing the presence of an iron-enhanced population have experienced contamination by this first SN Ia ejecta? Actually, we suggest that this idea is reasonable for population C in M2 and for population A in NGC 2808, for which these groups represent a small fraction of the total cluster mass, and do not show an internal O–Na anticorrelation. In all the other cases examined in Table 3, the presence of a double SGB suggests that we are dealing with a burst of SF after it has been inhibited for a long time, and we find it more natural to attribute the iron increase to delayed binary SN II, as described in Section 8.1. So, for instance, we estimate that the CNO−, s-rich population in NGC 1851 requires contamination by \( \sim 10 \) delayed SN II, rather than by one SN Ia. We point out that, obviously, the number of delayed SN II may have been larger than this because not all the matter polluted by these ejecta will be converted in stars during this final SF event (in particular because this last SF event may be interrupted by the beginning of the SN Ia epoch.)

A small calcium increase in the s-Fe-rich population, found in M2 (Da Costa et al. 2009; Marino et al. 2009) – see also Lee et al. (2009) – in NGC 1851 (Han et al. 2009) – see also Lim et al. (2015a,b) – and possibly in the population BI+BIi of M2 (Yong et al. 2014) is also a signature of SN II contamination, although, in other contexts, a calcium variation may be a sign of SN Ia contamination.
We have shown that the AGB scenario provides a timeline along which the large variety of multiple populations in GCs may find its place, as we have described by reconstructing the SFH of the five or more populations born in NGC 2808. The limitation of this analysis is mostly met when we attempt a quantitative comparison with the abundances of magnesium, silicon and aluminium, but all the qualitative trends and the general reconstruction of the events of SF are adequately met. No other model so far proposed as a basis for the chemical variety of multiple populations can meet such a requirement.

In order to extend the SFH to other clusters, we have examined all the different epochs which may occur in clusters in the first 100–120 Myr of life; in particular we have re-examined the possible role of delayed SN II explosions occurring in binaries in which mass exchange has risen the secondary component of the system above the completeness of delayed SN II explosions. We show that a limited number of explosions lasting for several tens of Myr may stop the SF until these supernovae explode, and lead finally to an intense burst of SF which may explain the characteristics of s-Fe-anomalous clusters.

In Table 4, we summarize schematically our proposal for the different events taking place in different prototype clusters. While further investigations aimed at exploring the occurrence of the events leading to cluster-to-cluster differences are necessary, the scenario we present, although based on imperfect stellar models, can explain the large variety of multiple-population features in clusters.
In Table 4, we do not include o Cen, because, in our scheme, the initial phases of its evolution cannot account for the fast s-process increase in the low-metallicity range (Johnson et al. 2009; D’Antona et al. 2011; Marino et al. 2011b), but we point out that the formation of its metal richer populations may have followed paths similar to those proposed here.

Much work remains to be done, and the scenario has enough details that it can be falsified by further analysis. The AGB nucleosynthesis is still very uncertain and the details of the proposed scheme(s) are not fully settled, but a whole and variegated set of results shows a global consistency with models whose temporal evolution is reasonably in line with what we know about stellar and dynamical evolution in clusters. We claim that, at this stage, the formation of multiple stellar generations in clusters, from matter including massive AGB ejecta, is the only viable option.

ACKNOWLEDGEMENTS

We thank the referee for a useful report which helped to improve the clarity of the work. We also thank Eugenio Carretta for useful comments and Lev Yungelson, Nicki Mennekens and Amedeo Tornambe for discussions on supernova rates. We are indebted to Gianluca Imbriani for enlightenment on the problems of the sodium plus proton reaction channels. FD’A and DE’E acknowledge support from PRIN INAF 2014 (principal investigator S. Cassisi). APM acknowledges support by the Australian Research Council through Discovery Early Career Researcher Award DE150101816. AFM acknowledges support by the Australian Research Council through Discovery Early Career Researcher Award DE160100851. EV acknowledges support from grant NASA-NNX13AF45G.

REFERENCES


MN2016_458_2122.pdf