

oMEGACat. VIII. A Subpopulation Census of ω Centauri

C. CLONTZ ^{1,2}, A. C. SETH ¹, Z. WANG (王梓先) ¹, M. HÄBERLE ^{3,2}, M. S. NITSCHAI ², N. NEUMAYER ²,
P. J. SMITH ^{2,4}, M. LATOUR ⁵, A. FELDMEIER-KRAUSE ^{6,2}, M. LIBRALATO ⁷ AND A. BELLINI ⁸

¹*Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA*

²*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

³*European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching, Germany*

⁴*Department of Physics and Astronomy, University of Heidelberg, Im Neuenheimer Feld 226, D-69120 Heidelberg, Germany*

⁵*Institut für Astrophysik und Geophysik, Georg-August-Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany*

⁶*Department of Astrophysics, University of Vienna, Türkenschanzstrasse 17, 1180 Wien, Austria*

⁷*INAF, Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, I-35122, Italy*

⁸*Space Telescope Science Institute, 3700 San Martin drive, Baltimore, MD, 21218, USA*

ABSTRACT

An understanding of the assembly history of the complex star cluster Omega Centauri has long been sought after, with many studies separating the stars on the color-magnitude diagram into multiple groupings across small magnitude ranges. Utilizing the oMEGACat combined astro-photometric and spectroscopic dataset we parse 14 subpopulations from the upper red-giant branch to below the main-sequence turnoff. We combine our results with previous works to estimate the age and age spread of each population. We find that the chemically enhanced (P2) populations are all ~ 1 Gyr younger (~ 11.6 Gyr old) and have significantly higher intrinsic age spreads (0.6 Gyr) than the primordial (P1) populations (~ 12.6 Gyr old, 0.3 Gyr spread), with the intermediate (Im) populations falling in between the two. Additionally, we connect for the first time the Chromosome Diagram to the two-stream age-metallicity relation, allowing us to link the P1 and P2 stars to the distinct star formation tracks, proposed to be in-situ and ex-situ contributions to the cluster's assembly. Our results are consistent with some suggested formation models and rule out others but no current model can explain all observed features of the subpopulations.

Keywords: globular clusters: general - globular clusters: individual (NGC 5139) - techniques: photometry - techniques: spectroscopy

1. INTRODUCTION

Omega Centauri (ω Cen) is well-known for its many unique features that distinguish it from typical globular clusters, including a large (~ 2 dex) spread in metallicity (Johnson & Pilachowski 2010; Johnson et al. 2020; Nitschai et al. 2023), suggesting it is composed of multiple stellar populations. There exists a multitude of studies aiming to constrain the exact number of subpopulations and their unique characteristics, some focused on the red-giant branch (RGB) (Marino et al. 2012; Milone et al. 2017; Husser et al. 2020), some on the subgiant branch (SGB) (Villanova et al. 2007, 2014), and others

on the main-sequence (MS) (Bellini et al. 2017; Latour et al. 2021). A few studies have focused on reproducing the full color-magnitude diagram (CMD) via fitting of isochrones and/or population synthesis models (Joo & Lee 2013; Tailo et al. 2016). The different techniques for identifying subpopulations provide a varied number of groupings of stars depending on the features the targeted data highlight. While the results have a strong dependence on data type and quality, there are similarities among the findings.

In spectroscopic studies, typically three primary metallicity groups are identified: the Metal-Poor (MP), Metal-Intermediate (MI), and Metal-Rich (MR) (Marino et al. 2011; Tailo et al. 2016; Bellini et al. 2017). The MP population contains around half of the ω Cen stars, exhibits a oxygen-sodium anti-correlation, helium

abundances of up to $Y = 0.40$ (Clontz et al. 2025) and is lanthanum-poor, indicative of r-process nucleosynthesis. The MI population contains 1/3 to 1/2 of the stars, displays a significant internal spread in elemental abundances, exhibits an extreme O-Na anti-correlation, and is lanthanum-rich, suggesting it has substantial s-process contributions to its enrichment. The MR population is a small anomalous group of stars, being significantly more iron rich than the rest of the groups with nearly all stars being Na-rich (~ 1 dex) but having a large range of oxygen values (0 - 0.6 dex) (Marino et al. 2011).

Photometry provides additional ways to parse subpopulations on the color-magnitude diagram (CMD) where varying the analyzed magnitude range (and thus stellar evolutionary phase) can highlight separate characteristic differences between the subpopulations. For example, subpopulations with different helium abundances separate out well on the MS while those with varying nitrogen abundance separate out well on the RGB.

The RGB is the best location on the CMD for parsing individual subpopulations as the intrinsic $[\text{Fe}/\text{H}]$ variations manifest most strongly as color differences here. A common technique to distinguish photometric differences is to construct a pseudo-color-color diagram, referred to here as the “chromosome diagram” (ChD). See Milone et al. (2017) and Nitschai et al. (2023) as well as section 3.2 for details on its construction. The ChD is information rich and has provided the basis for several studies of subpopulations (Milone et al. 2018, 2020; Marino et al. 2019). Beyond individual subpopulations the ChD can also be used to determine light element abundances, such as helium, across metallicity (Clontz et al. 2025).

Additional probes into the formation channels of ω Cen and their respective subpopulations come from ages of individual stars or full populations which are typically constrained along the subgiant branch due to model dependencies being minimized in this region. Villanova et al. (2007) measured relative ages for 80 SGB stars and distinguished 4 groups among them, while Joo & Lee (2013) used synthetic CMDs of 5 subgroups to model the populations, with age spreads up to 1.7 Gyr. A follow-up study by Villanova et al. (2014) using twice as many stars as before was able to separate the SGB stars into 6 groupings based on age, finding an age spread of at least 2 Gyr.

The age-metallicity relation (AMR) has been constrained by several studies (Hilker et al. 2004; Villanova et al. 2007, 2014), and was recently analyzed for a larger sample of stars ($\sim 7k$) by Clontz et al. (2024). The increased resolution clearly showed the two-stream feature suggested in Villanova et al. (2014). These two

discrete features in the AMR implies there are multiple formation pathways contributing to the formation of ω Cen. It is also worth noting the high likelihood that there are additional discrete sub-structures in the age-metallicity diagram which are not yet distinguishable due to the blurring effect of uncertainties on the ages and iron abundances. Additionally, the connection of these age-metallicity streams to the P1, Im, and P2 streams of the ChD is not straightforward, as the SGB ChD is less informative due to the subpopulation color range narrowing in this CMD region, especially in $C_{F275W, F336W, F435W}$.

Earlier studies at focused on fainter magnitudes, unveiling a puzzling split main-sequence (Anderson 2002; Norris 2004; Piotto et al. 2005). Metallicity measurements indicated the blue MS was more metal-rich than its red counterpart, which was later explained by an enhancement in helium in the bMS (Dupree & Avrett 2013; Karakas et al. 2014; Milone et al. 2020; Reddy 2020; Hema et al. 2020). Another study of the MS by Bellini et al. (2017) combined multiple photometric bands to identify distinct sequences. The five core groups they identified (the MSa, rMS, bMS, MSd, and MSe) were further parsed in a color-color space into 15 subpopulations. These groups were found to occupy distinct regions of the ChD of the main sequence.

The presence of multiple stellar populations with a significant spread in ages and diverse chemical properties points to ω Cen being the remnant core of a disrupted dwarf galaxy (Norris et al. 1996; Bekki & Freeman 2003; Hilker et al. 2004; Johnson & Pilachowski 2010; Villanova et al. 2014), especially since the successive starbursts with increasing enrichment observed are best facilitated by the larger host’s ability to retain supernova ejecta. While many studies have linked ω Cen to the Gaia-Enceladus dwarf galaxy merger event (Pfeffer et al. 2021; Callingham et al. 2022; Limberg et al. 2022), the debate remains open as to the cluster’s previous host dwarf galaxy (Majewski et al. 2012; Myeong et al. 2019; Forbes 2020).

In summary, while progress has been made in identifying the primary groups of ω Cen’s subpopulations, the complex variations within these groups necessitate a combination of probes to allow a better-resolved picture of the cluster’s evolutionary history. Subpopulation studies of ω Cen have been confined to small regions on the CMD where populations are separated into groups using a single probe (AMR, ChD). We present, in this work, subpopulations which are separated from the upper RGB, through the SGB, to below the main-sequence-turnoff (MSTO). This allows for the connecting of the ChD with the AMR and thus the connection

of probes of different aspects of the cluster’s assembly history. This work will also facilitate further spatial and kinematic studies and ultimately a clearer picture of the assembly of ω Cen.

In Section 2 we briefly detail the dataset utilized in this study and in Section 3 we outline our subpopulation parsing methodology in four phases: Sample Preparation, Initial Clustering, Iterative Propagation, and Informed Re-iteration and Final Results. In Section 4 we describe our results and cross-match them with the ages catalog of Clontz et al. (2024) where after we highlight the connection of assembly probes. In Section 5 we discuss interpretations of our results and compare our results to the literature. Finally, in Section 6 we summarize our conclusions and outline future works planned.

2. DATA

In this study, we utilize the oMEGACat - the combination of the The Hubble Space Telescope (HST) astro-photometric dataset (Häberle et al. 2024) and the Multi-Unit Spectroscopic Explore (MUSE) spectroscopic dataset (Nitschai et al. 2023) for ω Cen. The HST data were retrieved from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute and are accessible via doi: 10.17909/26qj-g090. This catalog provides photometric data in six broadband HST filters, each with specific corrections applied for the differential reddening and instrumental effects. The MUSE data can be accessed via: IOP Science. In the published MUSE catalog the metallicity is reported as $[M/H]$, as it is measured via full-spectrum fitting. They also apply an atomic diffusion correction which ensures the reported metallicities reflect the birth abundance of each star (Nitschai et al. 2023). More than 300,000 stars in the oMEGACat have measurements in all six photometric bands along with a reliable metallicity. We refer to the above two catalog papers for detailed data reduction procedures and the relevant data correction procedures.

To make direct comparisons with previous works more straightforward, we convert the $[M/H]$ values to $[Fe/H]$ using Equation 3 from Salaris et al. (1993):

$$[Fe/H] = [M/H] - \log(0.638 \times 10^{[\alpha/Fe]} + 0.362), \quad (1)$$

assuming $[\alpha/Fe] = 0.3$ dex (Norris & Da Costa 1995; Johnson & Pilachowski 2010). The uncertainty in our $[Fe/H]$ values thus depends on the uncertainty (and star-by-star variations) in the abundances of alpha elements, including oxygen. Marino et al. (2012) showed oxygen can increase by ~ 0.2 dex from $[Fe/H] = -1.9$ to $[Fe/H] = -1.05$, which is a source of systematic uncertainty in

our $[Fe/H]$ estimates, although its contribution is similar to the median metallicity error (0.08 dex). Hereafter all mentions of the metallicity of a star or stellar population is a reference to the $[Fe/H]$ value unless otherwise stated.

3. SUBPOPULATION PARSING METHODOLOGY

A schematic of our clustering algorithm is shown in Figure 2. This flowchart outlines each of the steps taken to arrive at the final results; we discuss each of the phases of our procedure in the following subsections. In Appendix A we discuss the various methods which were tested to arrive at our final algorithm and how each choice was informed by the outcome of these tests.

3.1. Phase I: Sample Preparation

This study focuses on parsing individual continuous subpopulations from the upper RGB (F814W $\sim$ 13) to below the MSTO (F814W $\sim$ 20). Phase I: Sample Preparation, indicated by the brown band in Fig. 2, details the preparations taken to ensure clean, high-quality samples on which we perform our subpopulation analysis.

3.1.1. Sample Selection

This step includes all of the photometric and spectroscopic quality cuts applied to our data, as well as additional steps taken to remove outliers.

To select only cluster members we exclude stars $> 3\sigma$ from the cluster velocity in both right-ascension and declination proper motion space (see Häberle et al. (2024) for details). Additionally, we apply radial velocity cuts using $RV_{probability} > 0.9$, given by the Nitschai et al. (2023) catalog.

We apply additional criteria to keep only the most reliable data. These include a high-quality photometry flag as well as an upper cut on the χ^2 weighted errors for each photometric filter we consider (both provided in the Häberle et al. (2024) catalog). Additionally, we select all stars in the range $-2.3 < [Fe/H] < -0.5$ and that have an $[Fe/H]$ uncertainty of < 0.3 dex. Lastly, we require the MUSE signal-to-noise ratio to be > 10 and the Reliability parameter to be > 0.95 (both provided in the Nitschai et al. (2023) catalog). These combined cuts give 106,860 high-quality candidate stars located within the half-light radius of ω Cen.

During the testing phase of our clustering methodology, we found that both blue and red color outliers in our color magnitude diagrams can affect the ability of the clustering algorithm to properly trace subpopulations to fainter magnitudes. This is due to these outliers having a notable effect on the calculation of the centroid of the subpopulation group, which causes the centroid to move,

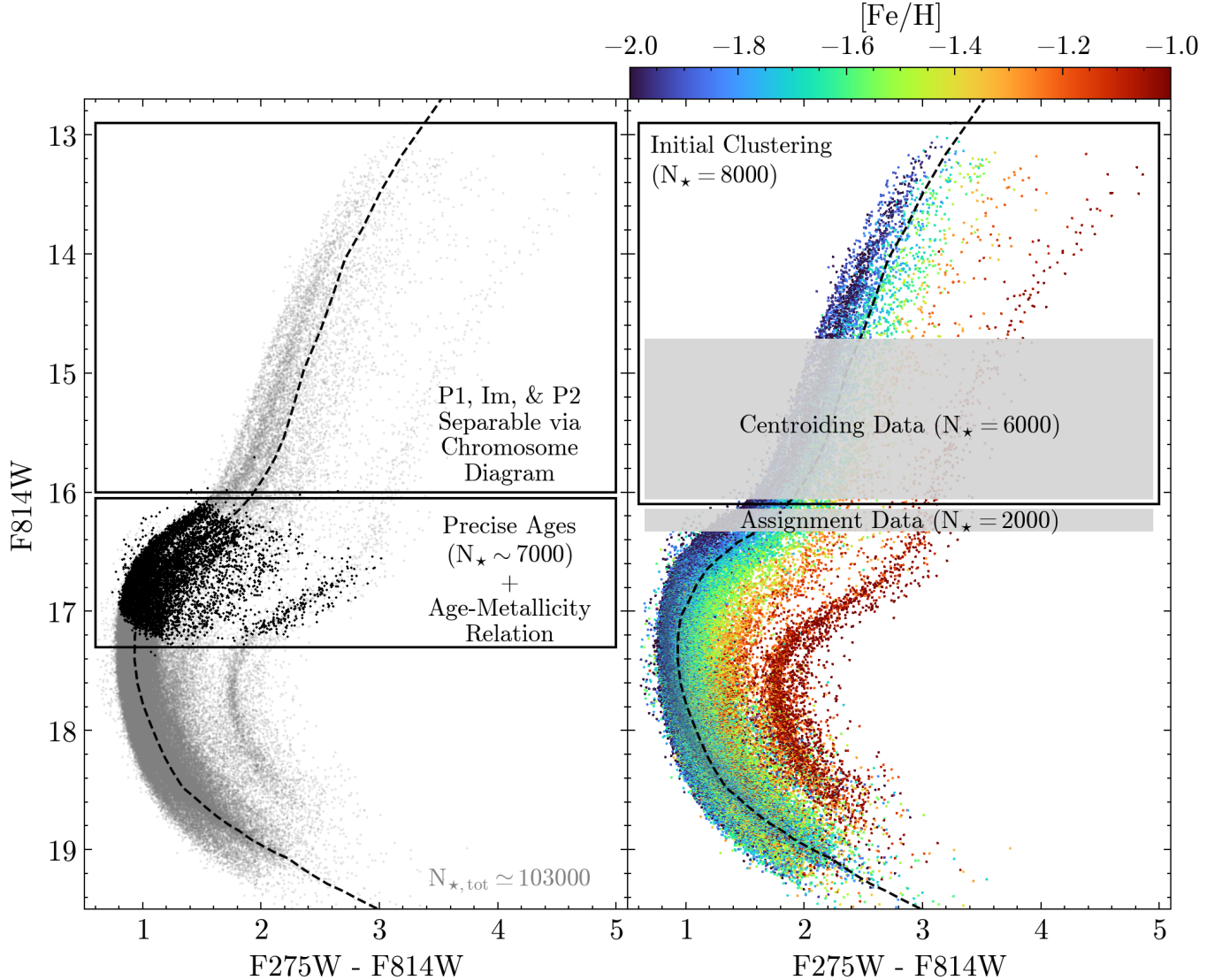


Figure 1. Color-Magnitude Diagrams: (both panels) The black dotted line traces the overdense region of this plot when considering a subset of stars centered around the median metallicity ($[\text{Fe}/\text{H}] \sim -1.7$). (left panel) The full subset of high-quality member stars is plotted with light grey markers. All SGB stars with age determinations are overplotted in black. The upper black box delineates the region where the RGB ChD is generated while the lower black box delineates where the age-metallicity relation is calculated. (right panel) The full subset of high-quality member stars is plotted with individual markers colored by their $[\text{Fe}/\text{H}]$ value, following the colorbar at the top of this panel. The black box delineates where the initial RGB clustering is performed. The upper grey shaded region shows the extent of the first propagation step centroiding sample while the lower grey shaded region shows the extent of the first propagation step assignment data. The scaling data is constituted by the combination of the centroiding and assignment data.

making it difficult to accurately propagate cluster labels to fainter magnitudes (details on this process are outlined in Section 3.3.2). To correct for this, we performed a general cleaning of the $F_{275\text{W}} - F_{814\text{W}}$ vs. $F_{814\text{W}}$ CMD by creating hand-drawn fiducial lines on both the blue and red extremes of the main locus of points, removing blue stragglers and binaries in the process. We also provide a conservative cleaning of the pseudo-color

CMD ($C_{F_{275\text{W}}, F_{336\text{W}}, F_{435\text{W}}}$ vs. $F_{814\text{W}}$). Together these cleanings remove a total of 363 stars.

Through additional tests, we find there is a metal-rich blue straggler population, which also causes the centroids of our subpopulation centroids to migrate as we propagate to fainter magnitudes. These could not be removed via the previous red and blue fiducial line cleaning method given that they fall within the main locus of the more metal-poor MSTO stars. To isolate them,

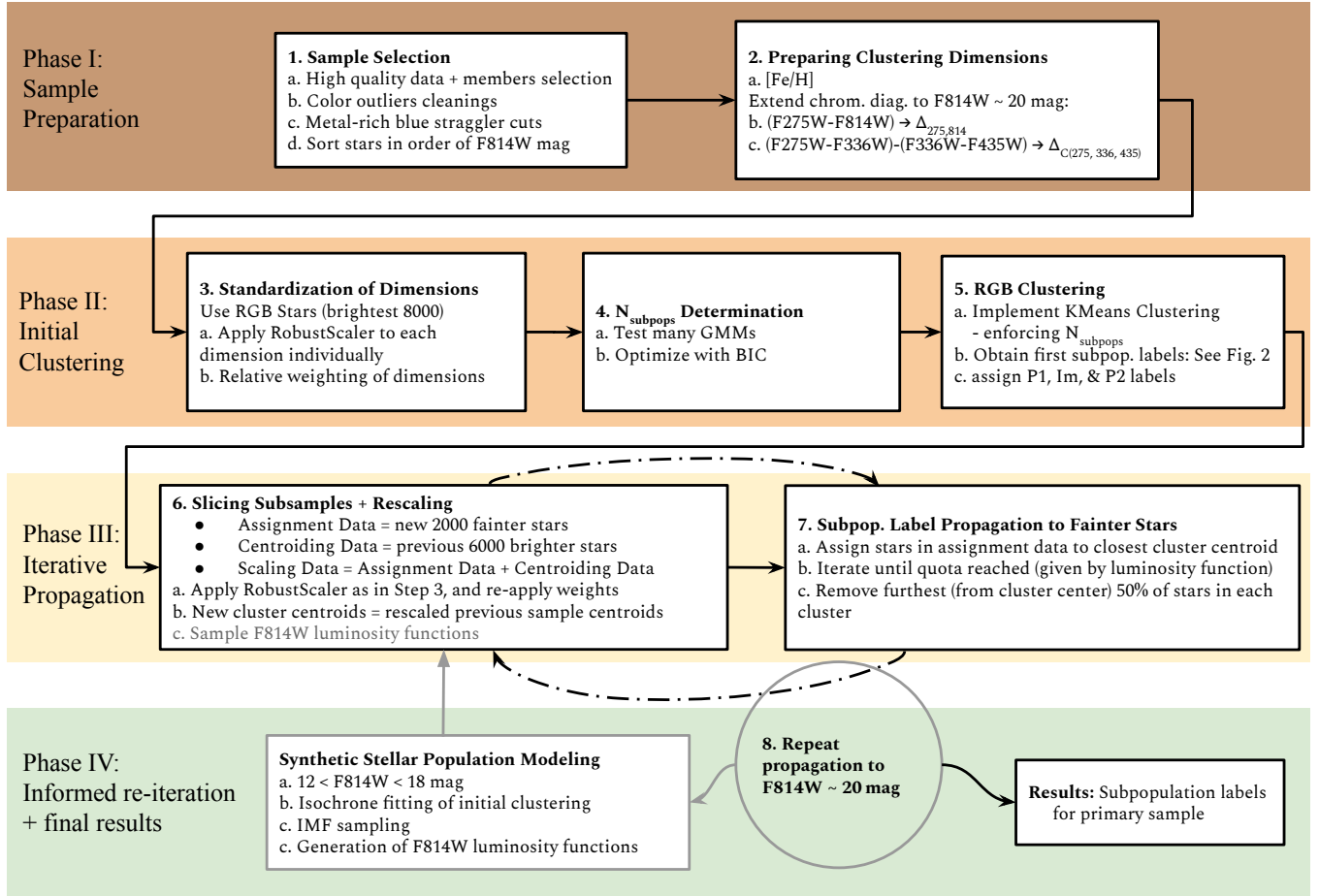


Figure 2. Subpopulation Parsing Algorithm Flowchart: The four phases of our subpopulation parsing algorithm are outlined vertically, top to bottom, with each phase being delineated by a horizontal colored band. Within each white box a bulleted list of grouped steps in the procedure is given. Black solid line arrows connect steps performed in order, and black dash-dot line arrows show steps performed iteratively. The grey lined boxes and arrows indicate steps which are only considered in our second-pass version of the subpopulation parsing algorithm.

we separate the most metal-rich stars into three [Fe/H] bins and further remove the blue straggler populations from each using hand-drawn color-fiducial lines, similar to the process outlined above. This removes an additional 3,038 stars, leaving us with 103,459 well-measured stars for our initial sample. This sample is plotted in grey points in the left panel of Fig. 1 and in the right panel each stars is colored its [Fe/H] value.

3.1.2. Preparing Clustering Dimensions

This step includes the preparation of the dimensions in which the clustering and label propagation is performed. In addition to [Fe/H], we utilize the two dimensions of the ChD.

To track the populations across the full range of magnitude ($12.9 < F814W < 19.7$) we build upon methods used to construct the RGB ChD, outlined in several

previous works (Milone et al. 2017; Nitschai et al. 2023; Clontz et al. 2025). These methods include verticalizing the bulk of the CMD sequences by using fiducial lines to create a pseudo-color-color (chromosome) diagram ($\Delta_{F336W, F814W}$ vs. $\Delta_{C_{F275W, F336W, F435W}}$). Here stellar populations with varying light element abundances separate well. Canonically, the x-axis is $\Delta_{F336W, F814W}$, formed from the verticalization of the F814W vs. F275W-F814W CMD, providing a long color-baseline and thus a proxy for metallicity and helium (Clontz et al. 2025). The y-axis is typically $\Delta_{C_{F275W, F336W, F435W}}$, generated from the verticalization of the pseudo-color $C_{F275W, F336W, F435W}$, which is the difference between two colors: $(F275W-F336W)-(F336W-F435W)$.

This “magic trio” (Milone et al. 2012; Piotto et al. 2015) of filters highlights the separation of stars with

varying abundances, due to the OH band falling in the F275W filter, the NH band falling in the F336W filter and CN and CH falling in the F438W filter. In the RGB ChD of ω Cen (see Fig. 3) there exists two distinct streams of stars extending from the top left to the bottom right, with a less populated middle stream falling in between the two. The lower stream is constituted of primordial (P1) populations which are helium and nitrogen-poor, while the upper stream contains chemically processed (P2) populations which are helium and nitrogen-enhanced. Those falling in the middle stream are considered Im populations and are thought to have intermediate abundances compared to P1 and P2.

For this work, we need to generate the verticalized colors for the ChD axes from the upper RGB down past the MSTO ($12.9 < F814W < 19.7$). To do this we first create fiducial lines by isolating a sample of stars in a narrow metallicity range centered on the median ($[Fe/H] \sim -1.7$). We then hand-draw a line which traces the densest region of the metallicity sample along the CMD, then interpolate this line at each star’s F814W magnitude. We calculate $\Delta_{F275W, F814W}$ and $\Delta_{C_{F275W, F336W, F435W}}$, by taking the difference between the interpolated fiducial line color and the color of each star. We then normalize the delta colors by the difference between the red and blue cleaning fiducial lines. These two delta colors, together with $[Fe/H]$, form the 3-dimensional space in which we perform our subpopulation parsing analysis. We also note that because the SGB slope and MSTO magnitude are highly dependent on helium abundance and metallicity, extending the ChD prescription beyond the RGB in a way that is fair to all of the subpopulation CMD morphologies is quite tricky and results in an imperfect verticalization which we work to compensate for in future steps of our parsing algorithm.

3.2. Phase II: Initial Clustering

Next we need to determine the number of subpopulations identifiable in our data and assign the first subpopulation labels to individual stars. To do so we turn to the RGB, where the clustered overdensities on the ChD, indicative of individual stellar populations, are known to separate best. Phase II, indicated by the orange band in Figure 2, details these steps. We start by selecting the brightest 8,000 stars ($12.9 < F814W < 16.1$), covering the main extent of the RGB, shown by the black box in the right panel of Fig. 1.

3.2.1. Standardization of Dimensions

We then standardize our RGB sample in each of the three utilized dimensions ($\Delta_{F336W, F814W}$, $\Delta_{C_{F275W, F336W, F435W}}$, and $[Fe/H]$). To do this, for each dimension we subtract the mean, which centers the data

at zero. We then scale the 16th and 84th percentiles to unit variance. These steps allow us to equalize the weight each dimension is given and centers the data at the origin, which helps the clustering algorithms perform best. This method was found to be more reliable than other scaling methods due to its insensitivity to outliers.

While scaling gives the same weight to each dimension, some data have more constraining power than others. This constraining power also changes with magnitude. Through empirical testing, we arrive at a set of weights that we find best help the clustering algorithm isolate overdensities and propagate membership labels in the combined ChD and $[Fe/H]$ space. We up-weight the $[Fe/H]$ dimension by scaling the values by 1.7. Similarly, we scale the $\Delta_{C_{F275W, F336W, F435W}}$ values by 1.5. We find that up-weighting the $[Fe/H]$ dimension by more than this causes poor performance on the main-sequence, but weighting it less than this causes poor clustering performance around the SGB and MSTO. Similarly, the $\Delta_{C_{F275W, F336W, F435W}}$ axis becomes critical below the MSTO as the main diagnostic for subpopulation separation. While we did test changing the relative weights of each dimension as a function of magnitude, we found the best results when keeping the scalings consistent along the full CMD.

3.2.2. $N_{subpops}$ Determination

In Step 4, we determine the number of distinguishable subpopulations in our data by passing the scaled and weighted data to a Gaussian Mixture Model (Pedregosa et al. 2011). We test the goodness of models with 1 to 25 clusters and for each we calculate the Bayesian Information Criterion value, which is minimized when the χ^2 and number of model parameters are balanced. We find a minimum BIC at $N_{clusters} = 14$. We chose this as our basis for our subpopulation sample but do not assert that this is the absolute number of subpopulations in ω Cen nor that each grouping is a pristine single stellar population.

3.2.3. RGB Clustering

With the number of subpopulations determined, we next need to calculate the centroids of each cluster in the three-dimensional space and assign subpopulation labels to each RGB star. In Step 5, we implement a K-Means (Pedregosa et al. 2011) clustering algorithm using the RGB sample (8000 brightest stars) from Section 3.2.1 and the 14 subpopulations (determined in Section 3.2.2) as a prior. This routine assigns a number from 0-13 (in no particular order) to each star, designating its subpopulation label. We take the median metallicity

of the RGB stars of each population and re-sort the subpopulation groups by $[\text{Fe}/\text{H}]$ and reassign their labels in ascending order such that the most metal-poor population is Pop. 0 and the most metal-rich is Pop. 13. In the upper panels of Fig. 3 we show the initial clustering for the RGB where each star is colored by its sorted population label. The upper left panel is the classic ChD, illustrating the discreteness of the subpopulations as clustered points in this space. We also show here the dividing lines for the P1, Im, and P2 populations, which follow the underdense regions between the three primary streams. In the upper right panel, the populations are plotted against $[\text{Fe}/\text{H}]$, where they are further distinguished. In both panels, we also overplot the $1\text{-}\sigma$ and $2\text{-}\sigma$ density contours in thick and thin black lines, given by the kernel density estimation of each cluster.

To facilitate comparisons between populations that reside in each of these streams, we use the subpopulation centroids on the RGB ChD (upper left panel of Fig. 3) to assign P1 labels to Pops. 0, 3, 4, 8, and 11, Im labels to Pops. 1, 5, and 7, and P2 labels to Pops. 2, 6, 9, 10, 12, and 13. It is important to clarify that our P1, Im, and P2 populations are analogous to "lower stream", "mid-stream", and "upper-stream" definitions in the literature (Marino et al. 2019; Clontz et al. 2025; Dondoglio et al. 2025). Similarly, previous studies of the ChDs of various clusters, such as Milone et al. (2017), denote the lower stream metal poor population (our Pop. 0) as 1G (first-generation) and the mid- and upper-stream metal-poor stars (our Pops. 1 and 2) as 2G (second-generation) stars. Here we do not adopt this nomenclature due to it implying a formation timeline.

3.3. Phase III: Iterative Propagation

In Phase III, highlighted by the yellow band in Fig. 2, we select and scale our next sub-samples of stars without subpopulation labels and propagate the RGB subpopulation labels to fainter magnitude stars using an iterative approach. This is repeated until we reach $F814W \sim 19.7$ mag.

3.3.1. Slicing sub-samples & Rescaling

In Step 6, we create several sub-samples of our labeled data, which we describe in detail below, that are important for aligning the labeled data with the next set of stars to receive population labels.

To begin, we create a sample of 8000 stars (referred to as the "scaling sample"), which contains the faintest 6000 stars from the previous sample (referred to as the "centroiding sample", indicated by the red shaded region in Fig. 1) along with the next 2000 fainter stars needing population labels assigned (referred to as the

"assignment sample", indicated by the blue shaded region in Fig. 1). This means that for the first step we take the 6000 faintest RGB stars as centroiding sample. The choice regarding the number of stars contained within the centroiding sample is made through iterative testing of the performance of our clustering algorithm with the main priority of ensuring good number statistics while at the same time maintaining enough flexibility for the algorithm.

Due to the metallicity dependence of the MSTO magnitude and the varying SGB slope due to helium abundance variations, the verticalization of the $\Delta_{F336W, F814W}$ and $\Delta_{CF275W, F336W, F435W}$ dimensions is not consistent across the full CMD. However, for propagation algorithm to track the same populations to fainter magnitudes it relies on stars that belong to the same population to live in the same region of the ChD. At fainter magnitudes these population even migrate across the three-dimensional clustering space, meaning we need to make sure the algorithm can similarly migrate the cluster centroid for this subpopulation to keep trace of its constituents. To aid in direct mapping, we use the scaling sample to re-normalize the extent of the centroiding and assignment samples. To aid in tracking migrations of clusters at fainter magnitudes we recalculate the new re-scaling at each progressive step in magnitude. To do this, we follow the same process described for the RGB initial clustering, meaning we remove the mean of each dimension, scale the $1\text{-}\sigma$ range to unit variance, and apply relative weights to each dimension.

From the scaling sample, we extract the centroiding data and use it to re-calculate the centroids of each subpopulation in the 3-dimensional space using the newly rescaled dimensions, ensuring the previous cluster centroids have been effectively transformed into the newly rescaled space. These are the cluster centroids we use in Step 3.3.2 to assign labels to the assignment data.

3.3.2. Subpopulation Label Propagation

In the first iteration of Step 7, we allocate cluster labels to the stars in the "assignment" data by finding, for each star, the closest subpopulation cluster centroid from the centroiding data and assigning that tag. After this is complete for all assignment data stars, we move on to Step 8, where we shift our samples down by 2000 stars and follow Steps 6 and 7 again, and then again. We iterate down to $F814W = 19.7$ to arrive at our "first pass" version of our subpopulation tagging. We then take these preliminary clusters as a template for a model of the luminosity functions, a process which is outlined in Phase IV.

3.4. Phase IV: Informed Re-iteration + Final Results

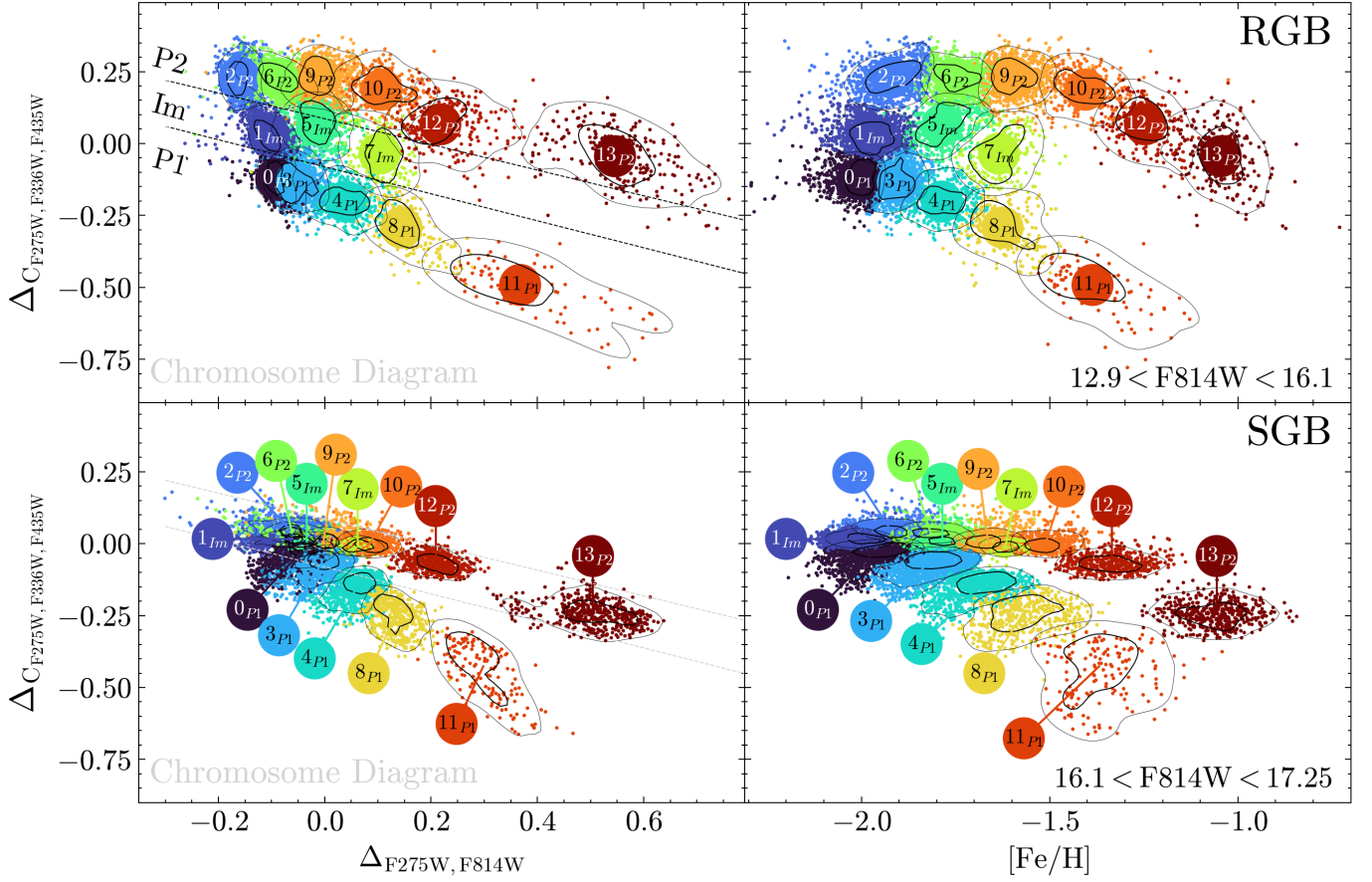


Figure 3. Subpopulation Parsing: (all panels) Each small marker point is a single star, colored by its subpopulation label. The large colored circle marks the associated cluster and within each is the annotation of the cluster number, in order of metallicity, as well as the ChD stream label, as a subscript. The 1- and 2- σ density contours are shown in thick and thin black lines, respectively. (upper left) The RGB chromosome map is constructed from the two delta-colors. The centroid of each subpopulation is given by the large colored markers. (upper right) The pseudo-color vs. $[\text{Fe}/\text{H}]$ space shows the distinction of the subpopulations in this space. (lower panels) Same as upper panels except now for the SGB. Here, the large colored markers are offset to aid visibility and are connected to the centroid of each cluster via a solid line with the same color.

Before beginning the second and final pass of the subpopulation cluster label propagation algorithm, we model the luminosity functions of the first-pass subpopulations to inform our label propagation. This modeling helps us account for the increasingly fainter MSTOs of the more metal-rich and/or older populations, resulting in the relative number of stars contained within each subpopulation not being consistent across various magnitude ranges.

To begin, we take the first-pass subpopulations and select representative isochrones from a grid of models adapted from the Dartmouth Stellar Evolution Database (Dotter et al. 2008) (see also Clontz et al. (2024) for more details). The isochrone grid contains models covering four parameter ranges: $8 < \text{Age} < 14$ Gyr (0.25 increments), $-2.5 < [\text{Fe}/\text{H}] < -0.5$ (0.01 increments), $[\alpha/\text{Fe}] = 0.2, 0.3$, and 0.4 , and $Y = 0.245$ and Y

$= 0.40$ and has a built in consideration of the C+N+O vs. $[\text{Fe}/\text{H}]$ relation for ω Cen.

We chose the isochrone model which has the median metallicity of its constituent RGB stars, the median age of its constituent SGB stars (discussed further in Section 4.2), $[\alpha/\text{Fe}] = 0.3$ (as assumed in Clontz et al. (2024)), and a helium mass fractions ($Y = 0.25$ and $Y = 0.40$) based on constraints from Clontz et al. (2025) where P1 populations are all assumed to have $Y = 0.25$ and all P2 populations are helium enhanced, with $Y = 0.40$. The intermediate population is shown to have slight helium enhancements but the coarseness of our helium model grid means we chose to represent these populations with $Y = 0.25$ models. Then, using each subpopulation isochrone model, we finely sample a Kroupa IMF (Kroupa 2001) within the mass range given by the isochrones and create an F814W luminosity function. Because our isochrone models are only calculated down

to F814W ~ 18 , this is the faint-end extent of our luminosity function constraints. Beyond this magnitude, we use the fraction of RGB stars contained in each subpopulation as the constraint for the number of stars assigned to each subpopulation, which is a reasonable approximation on the main-sequence. One caveat is that due to our stringent data quality cuts, we cannot expect the luminosity functions to perfectly follow those modeled from synthetic populations. This is especially true if any of our quality selections preferentially affect stars of a given metallicity.

We then begin a second pass of our clustering algorithm. We keep everything from Steps 1 through Step 6 the same. We start our second pass at Step 7, where we begin to propagate the subpopulation assignments to fainter stars. Now we can determine, from our modeled luminosity functions, how many stars in a given magnitude range should belong to each subpopulation ($N_{\text{exp},i}$, where i is the subpopulation index). Using this information, we take each subpopulation centroid and we find the $N_{\text{exp},i}$ closest stars to it and assign those stars the relevant subpopulation label. Before moving on to the next subpopulation cluster centroid we remove stars which have a distance from their cluster centroid greater than the 50th percentile distance of all stars in that cluster. This helps us prevent outliers from causing our populations to migrate away from their native over-densities and still preserves the relative fraction of subpopulation constituents dictated by the luminosity function samplings in the previous step. We also note that we do not apply this step to the RGB part of our subpopulations, meaning our subpopulation luminosity functions are not fully exemplary of the models on the RGB. After we do this step for one subpopulation, we move to the next subpopulation cluster centroid and repeat until we have assigned $N_{\text{exp},i}/2$ stars to all subpopulations in this magnitude range.

4. SUBPOPULATION RESULTS

We show our subpopulation results across the full F814W range in Fig. 4 where we separate the sequences into the P1, Im, and P2 groupings from the ChD for ease of visualization. Each grouping demonstrates not only the contiguousness of our subpopulations across the CMD, but the necessity for various helium abundance models to explain the spreads seen on the RGB and MS. For example, helium enhanced ($Y = 0.40$) isochrone models tend to have a less steep slope on the MS, a fainter MSTO, and a bluer RGB than their primordial helium ($Y = 0.25$) counterparts. This CMD morphological difference is especially highlighted when comparing the sharp MSTO of the metal-poor P1 populations to

the metal-rich P2 populations. Similarly, varying alpha abundance models can account for spreads seen on the RGB that cannot be fully explained by metallicity differences.

4.1. Verifying Subpopulation Consistency

A critical aspect of this analysis is the connection of the RGB ChD to the SGB AMR. To ensure our connection between these two CMD regions is reliable, we compare their elemental abundance patterns using sodium abundances constrained by Wang et al., (in prep.) where they apply the DD-Payne machine learning algorithm (Ting et al. 2017; Xiang et al. 2019; Wang et al. 2022) to the MUSE spectra from Nitschai et al. (2023) to obtain abundances for all well-measured stars.

Here, we use sodium abundances because they are found to show a large variations between the P1 and P2 populations and have relatively low median uncertainties (RGB: 0.07, SGB: 0.13). We first remove all stars observed with adaptive optics that show no flux in the NaID region of their spectra (as described in Wang et al. (2025)) due to the sodium feature in these spectra being masked. We have a total of 6434 RGB stars and 5654 SGB stars with both Na abundances and cluster labels. We then calculate the median sodium abundance and median [Fe/H] for the RGB and SGB constituents of each population. The results are shown in Fig. 5. As expected, all populations show an increasing [Na/Fe] abundance with increasing metallicity. The P2 populations shown an enhancement in [Na/Fe] compared to P1 and the Im populations fall in between the two. This overall trend persists on the SGB, though the range of [Na/Fe] values is lower due to higher uncertainties in this region. We do see a shifting of populations 5 and 6, driven primarily by a slight change in their median metallicities, but in follow-up we confirm their distinction on the ChD at all magnitudes. This confirms the consistency in our subpopulation identifications across these groups and allows us to connect the RGB and SGB assembly probes with confidence. Lastly, we examined the spread in sodium abundances seen within each subpopulation. On the RGB the median 1- σ range is 0.20 while on the SGB it is 0.26. In both cases this is 0.13 dex higher than the respective median uncertainties.

4.2. Cross-match with SGB Ages

Finally, matching our stars with those from the SGB ages work of Clontz et al. (2025) gives us 2472 stars with both subpopulation cluster labels and reliable ages. The results are shown in Figure 6 where we color each matched star on the SGB with its estimated age, given

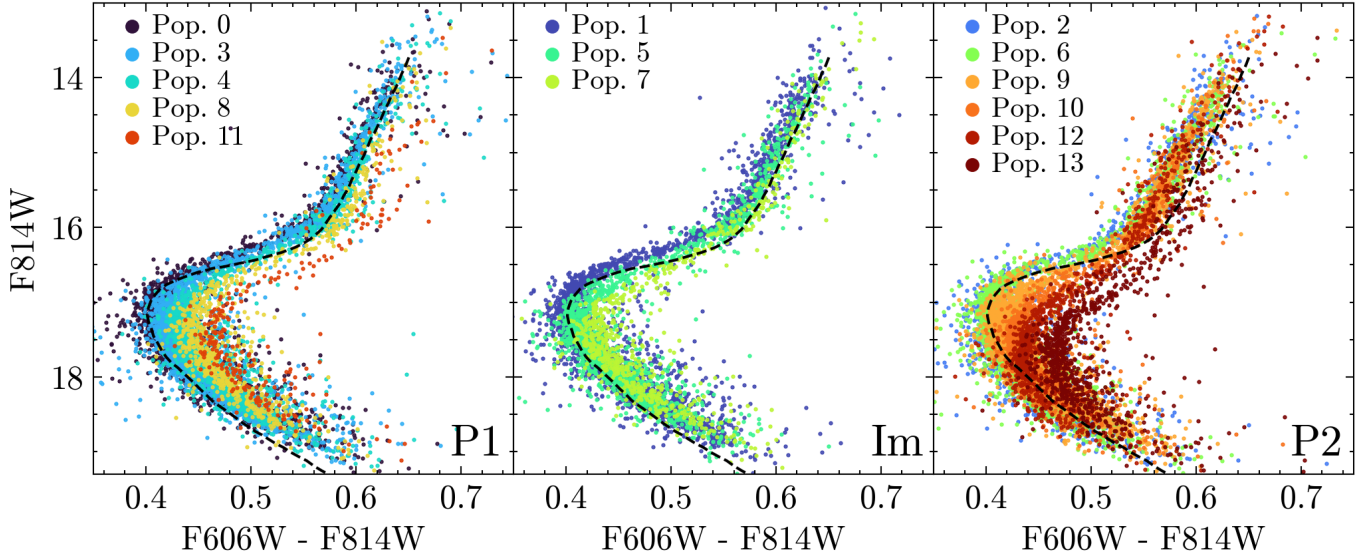


Figure 4. Chromosome Diagram Grouped CMDs: The F814W versus F606W - F814W CMD is shown for each set of populations grouped on the ChD (see Fig.3). The grouping (P1, Im, or P2) is annotation in the bottom right of each panel. Each data point is colored by its subpopulation assignment and indicated in the relevant legends. A fiducial line is given by a black dashed line, which traces the median metallicity population. Through comparison with the fiducial line we can confirm continuous sequences with CMD morphologies reflective of the expected variations in helium and alpha element abundances.

by the color bar along the top. We also plot the age histogram as an inset plot.

Looking at the age histograms, we see that many subpopulations have an overall age spread lower than the median age uncertainty reported in Clontz et al. (2024) ($\sigma \sim 0.7$ Gyr), suggesting our groupings are consistent with single-age/low age-spread stellar populations. Looking at the CMD sequences, we see the variety of MSTO and SGB morphologies across the populations reflect the expected variations in $[\alpha/\text{Fe}]$ and helium abundances, highlighted by the deviations from the fiducial black dashed line drawn in each panel.

We also note in the last panel of Fig.6, a cluster of stars in Pop. 13 which are brighter than the primary SGB sequence. One can also see them sitting slightly blue-ward on the SGB ChD (lower left panel) of Fig.3. It is clear they do not belong to either Pop. 11 nor Pop. 12. For now, we choose to leave them in Pop. 13, though we feel they are a separate population which is not fully distinguishable on the RGB.

4.3. Connecting Assembly Probes

Having a set of SGB stars with age constraints within each of our subpopulations means we can see in which region of the age-metallicity relation each subpopulation lies. To do this, we deconvolve the age histogram with the age uncertainty to calculate the mean age and intrinsic age spread of each subpopulation (following methods outlined in Clontz et al. (2024)). The results are shown in the top panel of Fig. 7, where all stars with ages

from Clontz et al. (2024) are plotted in black points (with their uncertainties given by the black contour in the lower right) and each population is over-plotted with a large medallion marker at its median RGB metallicity and mean SGB age. Each medallion population marker is labeled with the subpop. number and annotated with the relevant P1, Im, or P2 subscript. The median uncertainties in the mean ages are given by the deconvolution (~ 0.03 Gyr) and fall well within the extent of the medallions. For this reason we show the uncertainty on the mean age of the populations as a set of errorbars in the lower right corner. Looking at how the subpopulations occupy the AMR, we see the P1 populations stack vertically in this space, covering a large range in $[\text{Fe}/\text{H}]$ but sharing similar ages of ~ 12.5 Gyr. The same is true for the Im populations, sharing similar ages centered around ~ 12.1 Gyr. The P2 populations are slightly more spread, with ages between 10.7 and 11.9 Gyr. This clearly shows that the P2 populations are all > 0.80 Gyr younger than their similar-metallicity P1 counterparts and that the ages of the Im populations fall between those of P1 and P2 at all metallicities. This also shows that there is not a direct link between the 2-stream AMR sequences and the three-stream feature of the ChD.

The mean age vs. the intrinsic age spread of each subpopulation is plotted with the same medallion markers in the lower panel of Fig. 7. The uncertainty on the age spread is only visible for Pop. 4, so we chose to represent the median uncertainty on both quantities as

Subpop. Characteristics						Derived Parameters				Literature Comparison			
Pop.	ChD.	AMR	N_*	%	N_{ages}	[Fe/H]	$\sigma_{[\text{Fe}/\text{H}]}$	Age	σ_{age}	B17	B17	L21	D25
0	P1	Tight	6213	11.6	280	-1.95	0.02	12.69	0.28	rMS1	9.46%	-1.68	P1
1	Im	Tight	7877	14.7	405	-2.01	0.04	12.21	0.33	rMS2	9.71%	-1.66	P2 _{mid}
2	P2	Tight	5661	10.6	266	-1.90	0.04	11.87	0.54	bMS1	13.0%	-1.68	P2 _{upper}
3	P1	Diffuse	6007	11.3	295	-1.81	0.03	12.50	0.32	rMS3	14.0%	-1.69	AI
4	P1	Diffuse	3180	6.00	176	-1.69	0.03	12.62	0.25	rMS1	9.46%	-1.68	AI
5	Im	Diffuse	3971	7.40	176	-1.80	0.04	12.07	0.42	MSe1	6.46%	-1.52	AI _{mid}
6	P2	Tight	5595	10.5	284	-1.76	0.04	11.37	0.39	bMS2	9.32%	-1.54	AI _{mid}
7	Im	Diffuse	1525	2.90	93	-1.63	0.05	12.08	0.59	MSe2	6.56%	-1.53	AI _{mid}
8	P1	Diffuse	1751	3.30	89	-1.58	0.05	12.54	0.09	MSe3	1.17%	-1.48	AI
9	P2	Diffuse	5222	9.80	308	-1.65	0.04	11.59	0.65	bMS3	10.0%	-1.43	AI _{upper}
10	P2	Diffuse	2066	3.90	134	-1.51	0.05	11.81	0.63	MSd1,2	3.99%	-1.26	AI _{upper}
11	P1	Metal-rich	582	1.10	29	-1.40	0.10	12.39	0.27	MSe4	1.32%	-1.43	AI
12	P2	Metal-rich	1902	3.50	116	-1.28	0.05	11.28	0.75	MSd2,3	3.18%	-1.18	AI _{upper}
13	P2	Metal-rich	1804	3.40	101	-1.04	0.05	10.70	0.69	Msa1,2	3.53%	-0.93	AI _{upper}

Table 1. Subpopulation Characteristics and Literature Comparison: The Pop. column provides the number assigned to each subpopulation, ordered by metallicity. The ChD column provides the P1, Im, or P2 grouping of each subpopulation based on the ChD. The AMR column provides the Tight, Diffuse, or Metal-Rich grouping of subpopulations based on where they fall on the age-metallicity relation. The N_* column provides the number of stars with a given subpopulation label. The % column indicates the percentages of the total labeled sample contained within a given subpopulation. The N_{ages} column reports for how many stars in a given subpopulation we have age constraints. The [Fe/H] column reports the median RGB metallicity of each subpopulation and the $\sigma_{[\text{Fe}/\text{H}]}$ column gives the intrinsic [Fe/H] spread. Similarly, the Age and σ_{age} columns gives the median SGB ages and intrinsic age spread of each subpopulation. The final four columns show a correspondence of our subpopulations with the literature, more specifically Bellini et al. 2017 (B17), where further details for the same populations are then given by Latour et al. 2021, and Dondoglio et al. 2025(D25) respectively.

a set of errorbars in the lower left of this panel. The P1 populations cluster together at age spreads around 0.25 to 0.30 Gyr while nearly all P2 populations all have age spreads larger than 0.50 Gyr, reaching 0.75 Gyr for Pop 12. Pops. 6 and 8 are clear outliers, and it is notable that Pop. 6 is younger than the next two more metal-rich P2 populations, 9 and 10. The Im populations cover a large range of age spreads, from ~ 0.3 to 0.6 Gyr, increasing with metallicity. Interestingly, they once again fall between P1 and P2 populations, suggesting they may be linked to these star formation events, instead of being brought in from GC inspiral, as suggested by Mason et al. (2025). All of the populations exhibit quite low age spreads overall (< 0.1), consistent with globular clusters. In contrast, ages spreads seen in dwarf galaxy populations tend to be 0.30 - 0.40 Gyr Leaman (2012). This is likely due to the built-in discreteness of our method. As mentioned, some of our subpopulations are likely segments of a larger continuous star formation episode.

5. DISCUSSION

In this section we discuss the implications of the connection of these subpopulations with the AMR in an

effort to support or rule out potential formation mechanisms. We again emphasize that the main contribution this study adds is the age information. This does not allow us to definitively assert a formation model for the cluster but it does provide us with constraints which can support or rule out models which have been suggested in previous studies. We also examine the metallicity distributions of each population and the overall [Fe/H] distributions for the P1, Im, and P2 populations and to further contextualize our subpopulations, we compare their characteristics with those from the relevant literature.

5.1. Subpopulations on the Age-Metallicity Relation

In the upper panel of Fig. 7 we show the connection between our subpopulations and the AMR. In Clontz et al. (2024), we suggest the shape of the lower/tighter AMR sequence is consistent with in-situ star formation, characterized by self-enrichment. This would suggest that Pops. 0, 1, 2, and 6, which all fall along the lower/tight AMR sequence (see Table 1), formed from successive star formation episodes, or perhaps they are segments of a continuous and extended star formation period proceeding in the same environment. The low age spread

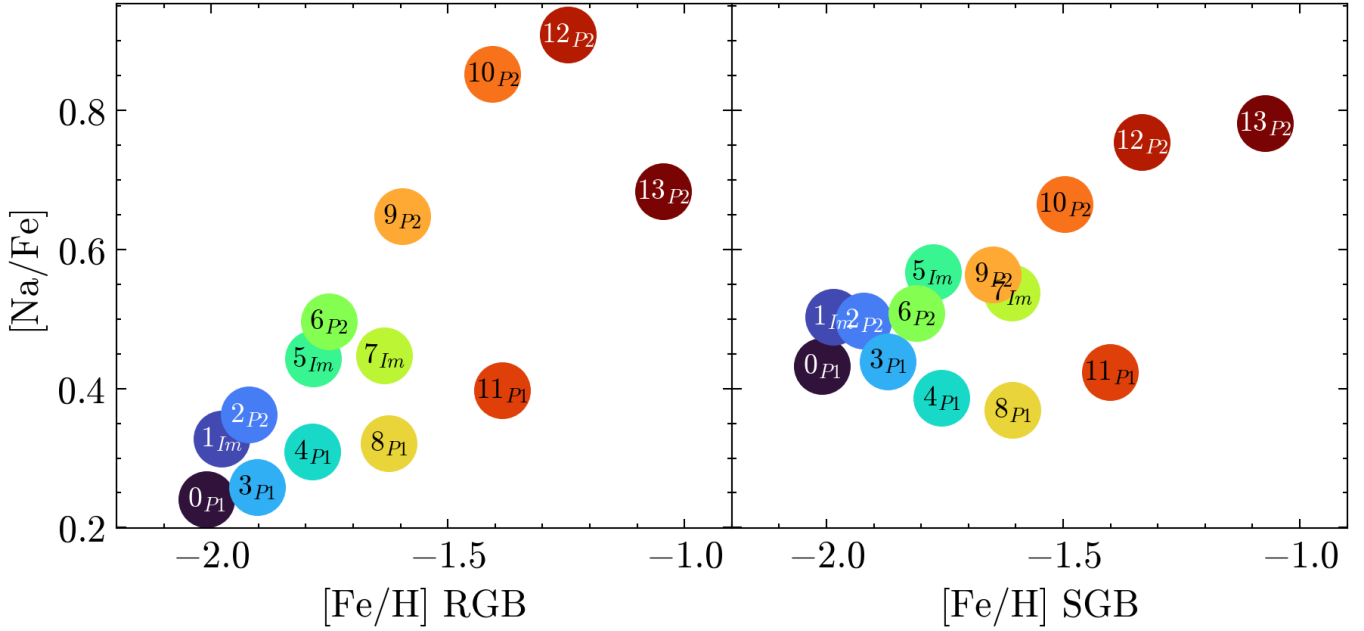


Figure 5. [Na/Fe] Verification of Subpopulation Consistency: (*both panels*) Shown here is the median DD-Payne derived sodium abundance for the each subpopulations versus the respective median iron abundance. The left panel contains values for stars on the RGB ($12 < F814W < 16$ mag) and the right panel contains those for SGB stars ($16 < F814W < 17$ mag). We see consistency in the separation of P1, Im, and P2 across the panels with the overall range of sodium values decreasing on the SGB due to increased uncertainties.

seen in Pop. 6 seems to suggest that population 6 experienced more rapid star formation. Looking at the time between these supposed successive star formation episodes we see that the age difference between Pops. 0 and 1 is 0.48 Gyr and between Pops. 1 and 2 is 0.36 Gyr, while the age difference between Pops. 2 and 6 is more than 1.31 Gyr. If we assume that populations that formed within a single environment follow similar slope tracks on the AMR, then we could theoretically also connect populations 3/5/9 and 4/7/10. Here we see age differences of ~ 0.43 and 0.49 between 3, 5, and 9, respectively. Grouping 4, 7, and 10, we find age differences of 0.54 and 0.27 Gyr. After these three sequences, there exist no further Im populations, making it difficult to assert that there is a connection between P1, Im, and P2 across all metallicities. The age difference between Pops. 8 and 12 is 1.26 Gyr, and between Pops. 11 and 13 is 1.70 Gyr.

It is also worth noting that the overall intrinsic age spread vs. $[Fe/H]$ relation seen in the lower panel of Fig. 7 is similar to that shown in Fig. 5 of Clontz et al. (2024), where they constrain the age spread in bins of metallicity and find values of $\sim 0.4 < \sigma < \sim 0.8$ Gyr, with a similar flattening of the relation at the highest metallicities. In Clontz et al. (2024) they also see a dip in overall age spread around $[Fe/H] \sim -1.8$ and -1.6 , where here we see the low age spread contribution of Pops. 6 and 8, respectively, suggesting these low age

spread populations are contributing to the dips seen in their metallicity-binned age spreads.

5.2. Metallicity Distributions of Subpopulations

The top panel of Fig. 8, shows the metallicity distribution functions of the grouped P1, Im, and P2 populations while the lower panel shows the metallicity distribution of each individual subpopulation. The P1 and Im populations show a remarkable similarity in their overall distributions suggesting a common origin or environment for these lower-metallicity subpopulations. P1 does see an additional peak around $[Fe/H] -1.9$, which clearly corresponds to population 3 in the lower panel. The Im population lacks stars above $[Fe/H]$ of -1.5 , while the P1 population has a small tail of stars extending to $[Fe/H] = -1.1$. Peaks are visible in the Im at $[Fe/H]$ of -1.75 corresponding to Pop. 5. The P2 group has considerably fewer stars at low metallicities, compared to the P1 and Im groups. However, it has many more stars above $[Fe/H] = -1.6$, extending all the way to -0.9 . There is considerable structure reflected in peaks corresponding to Pops. 6, 9, and 13. This extended metallicity distribution combined with the discreteness suggests perhaps a series of bursty star formation events within a single helium-enhanced environment; given the offset in metallicity and similarity in the light-element abundances, it seems perhaps this was separate from the environment where the P1 and Im stars formed.

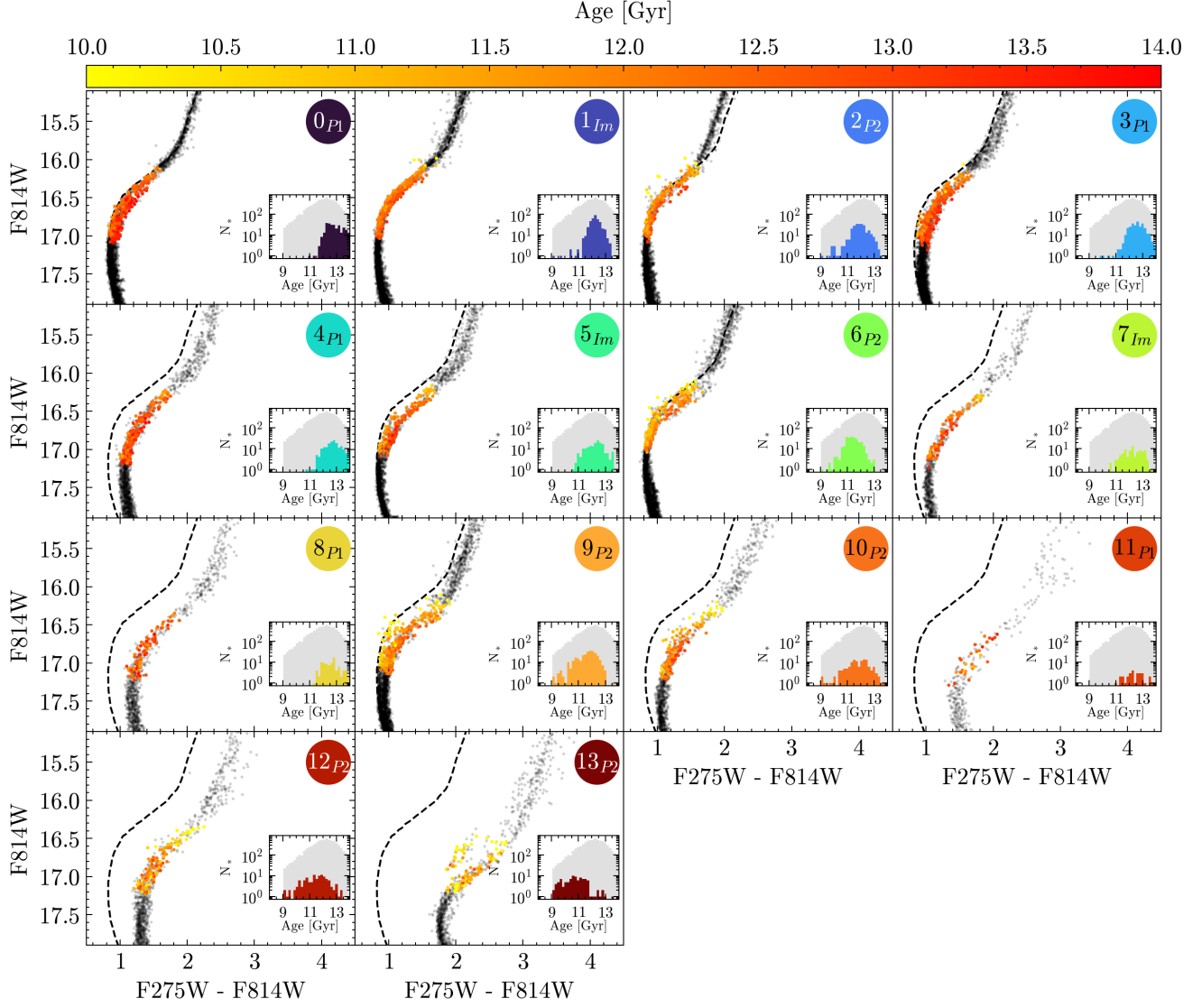


Figure 6. Ages of Subpopulation Constituents: One subpopulation is plotted within each subplot, top to bottom, left to right, from low to high metallicity. Within each panel, individual stars are plotted with black markers and SGB stars with age constraints are colored by their age. In the bottom right of each subplot there is an inset histogram showing in grey the full age distribution for all of our subpopulation labeled data and in the relevant color we shown the age distribution for that subpopulation. We also plot in each panel the same black dotted line which acts as a fiducial line for comparison across subpopulations.

Alternatively, with the addition of the SGB age information from [Clontz et al. \(2024\)](#), the broader categories of P1, Im, and P2 could also be interpreted as the result of similar chemical enrichment in separate environments over a similar span of time as discussed in the previous subsection. If populations 0+1+2, 3+5+9, and 4+7+10 represent these similar enrichment patterns, it would suggest a relatively slow enrichment timescale (~ 1 Gyr) for light elements (helium/nitrogen) taking place in material with different starting metallicities over the same time period.

5.3. Comparison to Literature

To compare with previous studies of ω Cen’s subpopulations, we discuss our 14 populations in terms of their average properties including their metallicity, their position on the ChD (P1, Im, P2), and the two-stream feature of the age-metallicity relation from [Clontz et al. \(2024\)](#). Due to the varying nature of how subpopulations are separated in the literature, we are limited to comparisons based on the relative qualities of the subpopulations.

The only previous subpopulation identification of a comparable number of stars to this work are the 15

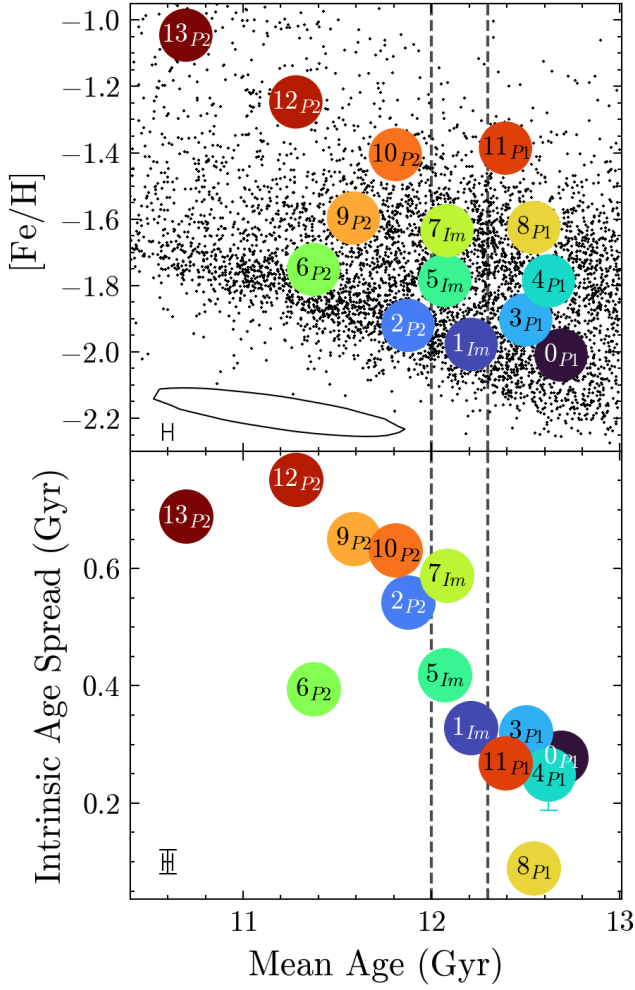


Figure 7. Subpopulation Age Relations: (*all panels*) Each population is represented by a single medallion marker with the color corresponding to the cluster. Inside each the cluster number is annotated along with the relevant ChD stream indication subscript. The vertical black dashed lines show an approximate separation of the P1, Im and P2 populations. (*upper panel*) Black points mark all SGB stars with ages from Clontz et al. (2024). Overplotted in the large colored circles are the mean (deconvolved) SGB age and median RGB metallicity for each subpopulation. The median age and $[\text{Fe}/\text{H}]$ uncertainty for the SGB stars is given by the black contour while the median uncertainty on the mean age of each subpopulation is given by the black errorbars in the lower left. (*lower panel*) Each subpopulation mean age is plotted versus its intrinsic age spread. The median uncertainty on the mean age and intrinsic age spread is represented by a set of errorbars in the lower left corner.

main sequence subpopulations sorted by Bellini et al. (2017) using an iterative selection on a variety of color-magnitude and color-color diagrams. Their work focused on the lower main-sequence and, therefore, the magnitude range of our classifications does not overlap

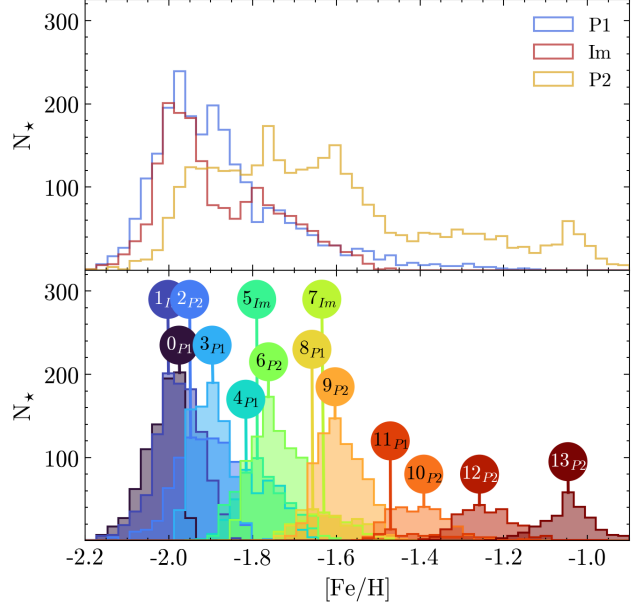


Figure 8. Metallicity Distributions: (*upper panel*) The metallicity distribution functions of the full P1 (blue), Im (crimson), and P2 (gold) groups are shown. While the P1 and Im groups have similar distributions, that of P2 differs significantly. (*lower panel*) The metallicity distribution functions of each of the individual subpopulations are shown, colored by their cluster number and labeled with the relevant medallion population markers at their peak value.

enough for a direct comparison. However, we can connect our subpopulations through the chromosome diagram morphologies.

We can compare our RGB ChD to the MS ChD under the assumption that abundance variations (e.g. Nitrogen enhancement) have similar impacts on these the color of stars in this space. An inversion (flipping of the y-axis) of panel f of Fig. 10 in Bellini et al. (2017) indeed shows a similar morphology to that of our RGB ChD. In this comparison, the MSa connection with our subpopulation 13 is clear. The three bMS subpopulations: bMS1, bMS2, and bMS3 seem to correspond to our populations 2, 6, and 9, while the three MSd populations correspond to our populations 10 and 12. The three rMS populations cover both P1 and Im populations 0, 1, 3, and 4. The four MSe populations also cover P1 and Im populations 5, 7, 8, and 11. We also tabulate this comparison in Table 1. Another way to compare the populations is their fractional composition, which we also outline in Table 1. Lastly, Latour et al. (2021) analysed MUSE observations for 4200 stars on the main-sequence and reported mean metallicities for each of the Bellini et al. (2017) subpopulations. We tabulate these as well. Where one of our populations is represented by

a mixture of two populations from Bellini et al. (2017) we list the average.

Recently, Mason et al. (2025) used spectroscopic abundances to separate populations by applying Gaussian Mixture modeling to a multi-dimensional abundance space using ~ 1500 APOGEE VAC stars belonging to ω Cen and were able to identify three groupings. These roughly translate to the P1, Im, and P2 populations on the ChD (see their Fig. 8). Their claim that the Im populations have a separate chemical enrichment consistent with globular clusters is well supported by the abundance patterns seen, especially in $[\text{Al}/\text{Fe}]$ vs. $[\text{Mg}/\text{Fe}]$ space (see their Fig. 5). Based on the interpretation of the population histories in the Mason et al. (2025) paper, this would suggest that our Pops. 1, 5, and 7 are NSC assembly contributions from GC in-spiral. While certainly possible, we note that the Im populations fall between P1 and P2 in every probe we examine (age, age spread, and sodium abundance), which would require a perhaps unlikely coincidence between these quantities in different formation environments. They also run chemical evolution models from which they conclude P1 formed first, in the center of ω Cen’s host galaxy and that P2 formed later, perhaps with GC origin. Our age constraints are not inconsistent with these models. They again note the mass budget problem as an added complexity to constraining formation models for these groups.

The recent publication by Dondoglio et al. (2025) discusses subpopulation in ω Cen in the context of the ChD and APOGEE abundances covering out to 5 half-light radii. They separate their RGB stars into three streams in the ChD and then further distinguish subgroupings in each stream. In the lower stream, they identify a first-generation (1P) population and an anomalous population (AI). For several of their analyses the middle and upper stream are combined, but are also separated into a metal-poor second-generation (2P) and additional more metal-rich anomalous population (AII). The correspondence with the subpopulations from this work is outlined in Table 1. In their Table C.1 they outline their formation scenario for each grouping and how they fit together. Their primary finding is that models can, in the case of aluminum abundance trends with metallicity, explain the formation of the middle stream stars as having formed from 2P ejecta diluted with 1P-like-gas. When comparing with this work, the main difference is our age constraints, which show the middle stream stars (Im populations, here) formed between the lower and upper stream stars in time. This argues against their dilution formation mechanism for the Im population. Another interesting finding is that the 2P stars are more centrally concentrated than the 1P stars and this

is mirrored in the AI and AII populations respectively. This supports the idea that the P2/2P stars formed from gas that was enriched by P1/1P ejecta that then sank to the center of the cluster before forming P2. This formation model implies all of ω Cen’s populations formed through self-enrichment processes, a model which struggles to reproduce the differences in metallicity distributions between P1 and P2 groupings.

All of these studies represented several significant steps forward in understanding the formation of ω Cen. Additional theoretical work on polluter mechanisms and the mass budget problem will surely soon uncover the complexity of its assembly.

5.4. Numbers of Subpopulations, their Discreteness, and their Origin

In this study we find 14 subpopulations that exist across the RGB, SGB, and MSTO. The fact that we detect a similar number to what was previously found by Bellini et al. (2017) on the lower MS, using different photometric analyses, is strong evidence for the number of identifiable subpopulations to be around 15. While undoubtedly some of these individual subpopulations or combinations of these subpopulations correspond to discrete evolutionary events or locations, some may also be part of a single formation channel spread out over time, as we suspect for the lower stream of the age-metallicity relationship. For example, the low age spread in population 6 and 8, and the very distinct age and metallicity of population 13 may correspond to a discrete formation event or environment separate from the rest of the cluster. It is clear that as more and better data become available, our ability to resolve the origin of these subpopulations will also improve. In particular, high-precision abundances across these subpopulations can enable determinations of the discreteness or continuity of star formation events among subpopulations.

6. CONCLUSIONS

This study combines multi-band photometry with precise spectroscopic metallicities that together provide, for the first time, the constraining power necessary to parse the multiple subpopulations of ω Cen into individual sequences from the upper RGB to the upper MS.

This analysis immediately provides information on the number of identifiable populations and their relative number densities. In addition, it allows, for the first time, the connection between the generations of stars (identifiable on the ChD) and the multiple streams seen in the AMR. This shows definitively that the two-stream feature of the AMR does not map directly to the three-stream feature of the ChD. Exactly how to group the

populations on the AMR is still unclear. The combination of these results with the DD-Payne abundances from Wang et al. (in prep.) will provide additional insight into the star formation history and assembly of ω Cen.

Thus, from the analyses outlined above we have concluded upon the following key points.

- In ω Cen we identify ~ 14 subpopulation groupings on the RGB down to below the MSTO. This likely does not represent the full set of distinct stellar populations that constitute this cluster but serves as a benchmark for further studies.
- The populations that fall into each of the P1, Im, and P2 groupings have similar mean ages across all metallicities, with the P1 populations being oldest at ~ 12.6 Gyr, the Im populations having mean ages about 0.4 Gyr younger than the P1 populations, and the P2 populations all having mean ages around 11.6 Gyr. The median age of the P2 population is 1.1 Gyr younger than the median age of P1.
- The P1 populations show lower intrinsic age spreads ($\langle \sigma_{intrinsic} \rangle = 0.24$ Gyr) compared to the P2 populations ($\langle \sigma_{intrinsic} \rangle = 0.61$ Gyr). The Im population shows a large range of age spreads, falling in between the P1 and P2 groupings.
- One interpretation of population occupation of the AMR is that sets of P1, Im, and P2 populations trace diagonal lines, most notably within the distinct lower stream found by Clontz et al. (2024). This could suggest parallel evolution and light element enrichment in multiple environments at the same time. Another interpretation is that there are three relatively vertical sequences in the AMR, traced by P1, Im, and P2 separately. In this model the connection with the structure of the AMR complicated.

Despite this wealth of information, a fundamental puzzle of the origin of ω Cen and its unique stellar pop-

ulations remains. Formation must have taken place in discrete environments, but how the evolution of the different subpopulations relate to one other remains uncertain. Several ongoing and planned works will continue to add detail to the model of ω Cen’s assembly. Follow up high-resolution spectra observed earlier this year with FLAMES/GIRAFFE@VLT will soon provide additional constraints on the different formation mechanisms for the two sequences of the AMR and ongoing work studying the abundance differences between the most-metal-rich P1 and P2 populations will help us understand the very anomalous Population 13. Additionally, updated isochrone models that can reproduce the F336W magnitudes of the nitrogen enhanced stars are generated and will be crucial for allowing isochrone fitting and proper synthetic stellar population modeling of the enhanced populations of this cluster. A deeper look into the mean abundances of light elements in each subpopulation is underway, using a machine learning methods on the MUSE spectra (Wang et al., in prep.).

CC acknowledges the contributions to this work via the high performance computing resources at the University of Utah as well as the cluster computing resources of the Max-Planck Institute for Astronomy Heidelberg. ACS, ZW and CC acknowledge support from a Hubble Space Telescope grant GO-16777. M.A.C. acknowledges the support from FONDECYT Postdoctorado project No. 3230727. AFK acknowledges funding from the Austrian Science Fund (FWF) [grant DOI 10.55776/ESP542]. The authors thank Davide Massari and Ricardo Schiavon for helpful discussions.

Facilities: HST(STScI), MUSE@VLT()

Software: Numpy (Harris et al. (2020)), Matplotlib (Hunter 2007), Astropy (Astropy Collaboration et al. 2013, 2018, 2022), Scikit-Learn (Pedregosa et al. 2011)

APPENDIX

In Appendix A we describe in more detail the additional tests which were performed to finalize our algorithm methodologies outlined in Section 3.

A. ALGORITHM VARIATION TESTING

With regards to our sample selections in Section 3.1.1, for several of our quality cuts we experimented with various thresholds, with the goal of balancing sample size with sample quality. Future clustering steps were found to be

strongly affected by outliers in $[\text{Fe}/\text{H}]$ and to rely heavily on this dimension at fainter magnitudes. Therefore, the above criterion of $[\text{Fe}/\text{H}]$ uncertainty < 0.3 was required. Additionally, we tested our clustering algorithm performance when using the DD-Payne iron abundances calculated in Wang et al. (in prep.), using $[\text{M}/\text{H}]$ instead of $[\text{Fe}/\text{H}]$, as well as excluding $[\text{Fe}/\text{H}]$ all together and did not find that any of these provided more pristine and/or distinct subpopulations than using the $[\text{Fe}/\text{H}]$ values derived from the MUSE metallicities.

In Section 3.1.2 where we prepare our data’s clustering dimensions, we also tested several iterations of varied fiducial line choices, including metallicity-dependent fiducial lines, the non-normalization of the colors, and principal component analysis, with no significant improvements made to the tracking of the subpopulations to fainter magnitudes. We additionally tested the performance of our clustering algorithm when using all six available photometric bands, which tended to obscure the subpopulations more and often caused the propagation algorithm to behave in unpredictable ways. Thus, we chose to keep our analysis to the three dimensions listed in the main text, the two delta colors $\Delta_{\text{F336W}, \text{F814W}}$ and $\Delta_{\text{C}_{\text{F275W}, \text{F336W}, \text{F435W}}}$, plus $[\text{Fe}/\text{H}]$.

When we did the N_{subpops} determination in Section 3.2.2, we also note that this number varies (from ~ 12 to ~ 17) based on the scaling of dimensions and the total number of RGB stars considered. However, we find there are considerable disadvantages to fewer or more clusters in later steps of our modeling. Fewer clusters lead to groupings with two or more distinct (by-eye) overdensities on the chromosome map being grouped together, often causing issues with the averaging of their metallicities complicating propagation. With more clusters, we can further distinguish groups that are clearly separate on the RGB. However, these groups are substantially overlapping in color at fainter magnitudes, making it very challenging to separate them with confidence. We performed several checks of clustering performance through the testing of literature subpopulation numbers, such as the 15 found by Bellini et al. (2017). The primary difference here is the separation of the two most metal-rich populations, which for us are grouped together (in Pop. 13). We find the “stealing” by Cluster 13 from the next metal-poor population (12 in our algorithm) is minimized when we do not split Cluster 13 into two.

Additionally, we tested our ability to isolate and track 3 primary populations (P1, Im and P2) given by the three streams on the RGB ChD, which proved to be messy below the RGB. Given these issues and others, we chose 14 subpopulations for this analysis, but do not assert that this is strictly the absolute number of subpopulations in ω Cen nor that each is a pristine single stellar population.

During the assembly and testing of our current RGB clustering algorithm in Section 3.2.3, we also tested outcomes using a Gaussian Mixture Model as well as a Bayesian Gaussian Mixture Model as the algorithm which provided initial cluster labels to the RGB. The primary drawback of these methods is that they require a Gaussian shape for each over-density. This would cause some clusters, such as Cluster 11 in the current method, to be extended in all clustering dimensions, causing a large range of $[\text{Fe}/\text{H}]$ values to be assigned. This method also always produces a background term, a low-amplitude high σ term meant to characterize the noise within a sample. This term would need to be removed before each step of propagation, and because of the already strict outlier removal processes, felt unnecessary. These Gaussian mixture models are also typically built on the assumption that each cluster has approximately the same variance, which is certainly not the case, given the extent of e.g. Cluster 11 vs. Cluster 4.

Lastly, in Section 3.3.2, where we begin to propagate our subpopulation labels, we would like to mention that this approach was preceded by many other methods, each of which had its advantages and disadvantages, and none of which performed better than our current algorithm when it came to keeping the color spread and metallicities consistent within each given subpopulation across the CMD. We tried the NearestCentroid algorithm (Pedregosa et al. 2011), assigning each star to its closest centroid, which led to the confusion of populations that cross around the MSTO. This also, in some cases, led to wandering cluster centers when clusters had a large variance. We additionally tested a KNearestNeighbor (Pedregosa et al. 2011) approach where the e.g. 10 nearest-labeled stars to a given unlabeled star voted on which label the target star received. This led to some clusters growing disproportionately large in a cascade of overassignment as we propagated to fainter magnitudes.

We also tested using the cluster assignments as training data for both a Gaussian Mixture Model and a K-Means clustering algorithm, which in theory would simply find the same N_{clusters} with similar centroids and variances in the next faintest bin. However, the initial issue was that the cluster numbers were random and would need to be resorted at every metallicity, which could cause populations with similar metallicities to become mixed with one another. Even after solving this, the re-seeding of the subpopulation centroids, even with strong priors, did not produce consistent results across the CMD.

Column Name	Description
HST ID	Häberle et al. (2024) Catalog Identifier
MUSE ID	Nitschai et al. (2023) Catalog Identifier
delta_275_814	Δ color for F275W - F814W
delta_275_336_435	Δ color for (F275W - F336W) - (F336W - F435W)
Subpopulation Tag	Integer indicating cluster assignment

Table 2. Provided Subpopulation Catalog: Each column of the provided catalog is described. All quantities are unitless.

B. DATA AVAILABILITY

The results of our subpopulation analysis are provided as a machine-readable table **to be added upon acceptance**. The columns provided in this table are outlined in Table 2.

REFERENCES

- Anderson, J. 2002, in *Astronomical Society of the Pacific Conference Series*, Vol. 265, *Omega Centauri, A Unique Window into Astrophysics*, ed. F. van Leeuwen, J. D. Hughes, & G. Piotto, 87
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33, doi: [10.1051/0004-6361/201322068](#)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123, doi: [10.3847/1538-3881/aabc4f](#)
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167, doi: [10.3847/1538-4357/ac7c74](#)
- Bekki, K., & Freeman, K. C. 2003, *MNRAS*, 346, L11, doi: [10.1046/j.1365-2966.2003.07275.x](#)
- Bellini, A., Milone, A. P., Anderson, J., et al. 2017, *ApJ*, 844, 164, doi: [10.3847/1538-4357/aa7b7e](#)
- Callingham, T. M., Cautun, M., Deason, A. J., et al. 2022, *MNRAS*, 513, 4107, doi: [10.1093/mnras/stac1145](#)
- Clontz, C., Seth, A. C., Dotter, A., et al. 2024, *ApJ*, 977, 14, doi: [10.3847/1538-4357/ad8621](#)
- Clontz, C., Seth, A. C., Wang, Z., et al. 2025, *ApJ*, 984, 162, doi: [10.3847/1538-4357/adc1c5](#)
- Dondoglio, E., Milone, A. P., Marino, A. F., et al. 2025, *arXiv e-prints*, arXiv:2509.16719, doi: [10.48550/arXiv.2509.16719](#)
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, *ApJS*, 178, 89, doi: [10.1086/589654](#)
- Dupree, A. K., & Avrett, E. H. 2013, *ApJL*, 773, L28, doi: [10.1088/2041-8205/773/2/L28](#)
- Forbes, D. A. 2020, *MNRAS*, 493, 847, doi: [10.1093/mnras/staa245](#)
- Häberle, M., Neumayer, N., Bellini, A., et al. 2024, *arXiv e-prints*, arXiv:2404.03722, doi: [10.48550/arXiv.2404.03722](#)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](#)
- Hema, B. P., Pandey, G., Kurucz, R. L., & Allende Prieto, C. 2020, *ApJ*, 897, 32, doi: [10.3847/1538-4357/ab93bd](#)
- Hilker, M., Kayser, A., Richtler, T., & Willemsen, P. 2004, *A&A*, 422, L9, doi: [10.1051/0004-6361:20040188](#)
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90, doi: [10.1109/MCSE.2007.55](#)
- Husser, T.-O., Latour, M., Brinchmann, J., et al. 2020, *A&A*, 635, A114, doi: [10.1051/0004-6361/201936508](#)
- Johnson, C. I., Dupree, A. K., Mateo, M., et al. 2020, *AJ*, 159, 254, doi: [10.3847/1538-3881/ab8819](#)
- Johnson, C. I., & Pilachowski, C. A. 2010, *ApJ*, 722, 1373, doi: [10.1088/0004-637X/722/2/1373](#)
- Joo, S.-J., & Lee, Y.-W. 2013, *ApJ*, 762, 36, doi: [10.1088/0004-637X/762/1/36](#)
- Karakas, A. I., Marino, A. F., & Nataf, D. M. 2014, *ApJ*, 784, 32, doi: [10.1088/0004-637X/784/1/32](#)
- Kroupa, P. 2001, *MNRAS*, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](#)
- Latour, M., Calamida, A., Husser, T. O., et al. 2021, *A&A*, 653, L8, doi: [10.1051/0004-6361/202141791](#)
- Leaman, R. 2012, *AJ*, 144, 183, doi: [10.1088/0004-6256/144/6/183](#)
- Limberg, G., Souza, S. O., Pérez-Villegas, A., et al. 2022, *ApJ*, 935, 109, doi: [10.3847/1538-4357/ac8159](#)
- Majewski, S. R., Nidever, D. L., Smith, V. V., et al. 2012, *ApJL*, 747, L37, doi: [10.1088/2041-8205/747/2/L37](#)
- Marino, A. F., Milone, A. P., Piotto, G., et al. 2011, *ApJ*, 731, 64, doi: [10.1088/0004-637X/731/1/64](#)
- . 2012, *ApJ*, 746, 14, doi: [10.1088/0004-637X/746/1/14](#)
- Marino, A. F., Milone, A. P., Renzini, A., et al. 2019, *MNRAS*, 487, 3815, doi: [10.1093/mnras/stz1415](#)
- Mason, A. C., Schiavon, R. P., Kamann, S., et al. 2025, *arXiv e-prints*, arXiv:2504.06341, doi: [10.48550/arXiv.2504.06341](#)

- Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012, *ApJ*, 744, 58, doi: [10.1088/0004-637X/744/1/5810.1086/141918](https://doi.org/10.1088/0004-637X/744/1/5810.1086/141918)
- Milone, A. P., Piotto, G., Renzini, A., et al. 2017, *MNRAS*, 464, 3636, doi: [10.1093/mnras/stw2531](https://doi.org/10.1093/mnras/stw2531)
- Milone, A. P., Marino, A. F., Renzini, A., et al. 2018, *MNRAS*, 481, 5098, doi: [10.1093/mnras/sty2573](https://doi.org/10.1093/mnras/sty2573)
- . 2020, *MNRAS*, 497, 3846, doi: [10.1093/mnras/staa2119](https://doi.org/10.1093/mnras/staa2119)
- Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, *MNRAS*, 488, 1235, doi: [10.1093/mnras/stz1770](https://doi.org/10.1093/mnras/stz1770)
- Nitschai, M. S., Neumayer, N., Clontz, C., et al. 2023, *ApJ*, 958, 8, doi: [10.3847/1538-4357/acf5db](https://doi.org/10.3847/1538-4357/acf5db)
- Norris, J. E. 2004, *ApJL*, 612, L25, doi: [10.1086/423986](https://doi.org/10.1086/423986)
- Norris, J. E., & Da Costa, G. S. 1995, *ApJ*, 447, 680, doi: [10.1086/175909](https://doi.org/10.1086/175909)
- Norris, J. E., Freeman, K. C., & Mighell, K. J. 1996, *ApJ*, 462, 241, doi: [10.1086/177145](https://doi.org/10.1086/177145)
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, 12, 2825
- Pfeffer, J., Lardo, C., Bastian, N., Saracino, S., & Kamann, S. 2021, *MNRAS*, 500, 2514, doi: [10.1093/mnras/staa3407](https://doi.org/10.1093/mnras/staa3407)
- Piotto, G., Villanova, S., Bedin, L. R., et al. 2005, *ApJ*, 621, 777, doi: [10.1086/427796](https://doi.org/10.1086/427796)
- Piotto, G., Milone, A. P., Bedin, L. R., et al. 2015, *AJ*, 149, 91, doi: [10.1088/0004-6256/149/3/91](https://doi.org/10.1088/0004-6256/149/3/91)
- Reddy, A. B. S. 2020, *MNRAS*, 495, 383, doi: [10.1093/mnras/staa1081](https://doi.org/10.1093/mnras/staa1081)
- Salaris, M., Chieffi, A., & Straniero, O. 1993, *ApJ*, 414, 580, doi: [10.1086/173105](https://doi.org/10.1086/173105)
- Tailo, M., Di Criscienzo, M., D’Antona, F., Caloi, V., & Ventura, P. 2016, *MNRAS*, 457, 4525, doi: [10.1093/mnras/stw319](https://doi.org/10.1093/mnras/stw319)
- Ting, Y.-S., Rix, H.-W., Conroy, C., Ho, A. Y. Q., & Lin, J. 2017, *ApJL*, 849, L9, doi: [10.3847/2041-8213/aa921c](https://doi.org/10.3847/2041-8213/aa921c)
- Villanova, S., Geisler, D., Gratton, R. G., & Cassisi, S. 2014, *ApJ*, 791, 107, doi: [10.1088/0004-637X/791/2/107](https://doi.org/10.1088/0004-637X/791/2/107)
- Villanova, S., Piotto, G., King, I. R., et al. 2007, *ApJ*, 663, 296, doi: [10.1086/517905](https://doi.org/10.1086/517905)
- Wang, Z., Hayden, M. R., Sharma, S., et al. 2022, *MNRAS*, 514, 1034, doi: [10.1093/mnras/stac1425](https://doi.org/10.1093/mnras/stac1425)
- Wang, Z., Seth, A. C., Latour, M., et al. 2025, *arXiv e-prints*, arXiv:2510.00330, doi: [10.48550/arXiv.2510.00330](https://doi.org/10.48550/arXiv.2510.00330)
- Xiang, M., Ting, Y.-S., Rix, H.-W., et al. 2019, *ApJS*, 245, 34, doi: [10.3847/1538-4365/ab5364](https://doi.org/10.3847/1538-4365/ab5364)