The Astrophysical Journal, 977:70 (21pp), 2024 December 10 https://doi.org/10.3847/1538-4357/ad8bac

Strong Chemical Tagging in FIRE: Intra- and Intercluster Chemical Homogeneity in Open Clusters in Milky Way–like Galaxy Simulations

Binod Bhattarai[1], Sarah R. Loebman[1], Melissa K. Ness[2,3], Andrew Wetzel[4], Emily C. Cunningham[3,7], Hanna Parul[5], and Alessa Ibrahim Wiggins[6]
[1] Department of Physics, University of California, Merced, CA 95343, USA; bbhattarai@ucmerced.edu
[2] Research School of Astronomy & Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[3] Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[4] Department of Physics & Astronomy, University of California, Davis, CA 95616, USA
[5] Department of Physics & Astronomy, University of Alabama, Box 870324, Tuscaloosa, AL 35487-0324, USA
[6] Department of Physics and Astronomy, Texas Christian University, Fort Worth, TX 76109, USA
Received 2024 August 1; revised 2024 October 14; accepted 2024 October 25; published 2024 December 3

Abstract

Open-star clusters are the essential building blocks of the Galactic disk; "strong chemical tagging"—the premise that all star clusters can be reconstructed given chemistry information alone—is a driving force behind many current and upcoming large Galactic spectroscopic surveys. In this work, we characterize the abundance patterns for nine elements (C, N, O, Ne, Mg, Si, S, Ca, and Fe) in open clusters (OCs) in three galaxies (m12i, m12f, and m12m) from the Latte suite of FIRE-2 simulations, to investigate the feasibility of strong chemical tagging in these simulations. We select young massive ($\geqslant 10^{4.6} M_{\odot}$) OCs formed in the last ∼100 Myr and calculate the intra- and intercluster abundance scatter for these clusters. We compare these results with analogous calculations drawn from observations of OCs in the Milky Way. We find the intracluster scatter of the observations and simulations to be comparable. While the abundance scatter within each cluster is minimal ($\lesssim 0.020$ dex), the mean abundance patterns of different clusters are not unique. We also calculate the chemical difference in intra- and intercluster star pairs and find it, in general, to be so small that it is difficult to distinguish between stars drawn from the same OC or from different OCs. Despite tracing three distinct nucleosynthetic families (core-collapse supernovae, white dwarf supernovae, and stellar winds), we conclude that these elemental abundances do not provide enough discriminating information to use strong chemical tagging for reliable OC membership.

Unified Astronomy Thesaurus concepts: Open star clusters (1160); Galactic abundances (2002); Magnetohydrodynamical simulations (1966); Galaxy disks (589); Stellar abundances (1577); OB associations (1140)

1. Introduction

Uncovering the history of the Milky Way (MW) stellar disk is a fundamental goal of Galactic astronomy and is tied to the smallest-scale structures where star formation occurs, open clusters (OCs). OCs are groups of stars that are born together and remain gravitationally bound for a period of time; C. J. Lada & E. A. Lada (2003) define an OC as a group of young stars that survive tidal disruption for a period as long as 100 Myr. In numerical simulations, OCs are usually defined as groups of co-forming stars that share a common origin (e.g., C. L. Dobbs et al. 2017; M. Y. Grudić et al. 2023).

OCs are often leveraged to study the dynamical and chemical history of the Galactic disk, as these bright structures are relatively straightforward to observe and characterize (e.g., P. M. Frinchaboy & S. R. Majewski 2008; N. Myers et al. 2022 and references therein). OCs are considered the building blocks of the Galactic disk because a significant fraction of star formation occurs within them and the majority are thought to dissolve quickly (C. J. Lada & E. A. Lada 2003; J. Bland-Hawthorn et al. 2010; J. M. D. Kruijssen et al. 2011; M. R. Krumholz et al. 2019). One of the important objectives of near-field astronomy is to reconstruct the structures that have dispersed due to dynamical processes in the disk, specifically, reassembling structures like OCs using present-day stellar observables.

One stellar observable of particular interest is elemental abundance. This is because stars that belong to a particular OC are thought to form from a common well-mixed molecular cloud with a similar underlying chemistry (F. H. Shu et al. 1987; M. R. Meyer et al. 2000; S. F. Portegies Zwart et al. 2010; Y. Feng & M. R. Krumholz 2014). Based on this assumption, stars that were born within a given OC are believed to have a similar chemical "fingerprint" that would enable OC reconstruction based on chemistry alone (e.g., K. Freeman & J. Bland-Hawthorn 2002; G. M. De Silva et al. 2007, 2009; J. Bland-Hawthorn et al. 2010; J. Bovy 2016; C. Manea et al. 2022). This is the premise behind strong chemical tagging (K. Freeman & J. Bland-Hawthorn 2002), which relies on the accurate measurement of multiple elements in stars to reconstruct cluster membership (e.g., S. R. Majewski et al. 2012; J. Bovy 2016; D. W. Hogg et al. 2016; S. L. Martell et al. 2016; L. Spina et al. 2022b).

The feasibility of strong chemical tagging is a key motivation for several current and upcoming spectroscopic surveys (e.g., GALAH, APOGEE, and WEAVE; G. Dalton et al. 2012; G. M. De Silva et al. 2015; J. A. Kollmeier et al. 2017; S. R. Majewski et al. 2017; S. Buder et al. 2019). For strong chemical tagging to work, the member stars within an OC should be nearly chemically homogeneous, i.e., they

[7] NASA Hubble Fellow.

should have small intracluster dispersion. In addition, each cluster should have a unique chemical signature, so that it can be well distinguished from the stars that belong to other OCs, i.e., there should be significant intercluster dispersion (D. L. Lambert & A. B. S. Reddy 2016).

Whether there is typically a small level of intracluster dispersion (i.e., the level of similarity in the element abundances of member stars) in OCs is still a matter of debate (S. Blanco-Cuaresma et al. 2015; N. Price-Jones & J. Bovy 2019). However, in the limit of small samples, the chemistry of stars that belong to a particular OC has often been shown to share a similar signature when compared with the chemistry of stars that belong to different clusters (e.g., G. M. De Silva et al. 2006; E. Pancino et al. 2010; S. R. Majewski et al. 2012; A. W. Mitschang et al. 2013; D. W. Hogg et al. 2016; S. L. Martell et al. 2016). This indicates that stars that belong to an OC could have a measurable signature that is distinct from other OCs, in the limit that measurement uncertainties in stellar abundances are smaller than the level of the intrinsic element abundance scatter, so that the amplitude of homogeneity in OCs can be uniquely assessed (M. Ness et al. 2015, 2018).

Large surveys have extended studies of OCs to more substantial samples and determined the intrinsic scatter to be <0.03 dex within clusters for up to 15 element abundances (S. Bertran de Lis et al. 2016; J. Bovy 2016; M. Ness et al. 2018; V. J. Poovelil et al. 2020). Interestingly, S. Blanco-Cuaresma et al. (2015) find distinct chemical signatures for stars in different evolutionary stages that belong to the same OC. In fact, theoretical expectations lead to a prediction that intracluster scatter is nonzero when processes like nonthermal equilibrium, atomic diffusion, mixing, the influence of binarity, and planetary engulfment are included in stellar abundance calculations. These predictions have been born out both observationally and in theoretical models (A. Dotter et al. 2017; F. Liu et al. 2019).

Besides debates about intracluster homogeneity, there are also several discussions in the literature about the level of intercluster homogeneity throughout the MW. For example, there is an expectation that the intercluster homogeneity could be a function of galactocentric radius (L. Spina et al. 2022a). In fact, a negative radial gradient in the MW's disk has been found in observations (e.g., C. Boeche et al. 2013; F. Anders et al. 2014; J. Donor et al. 2020; N. Myers et al. 2022; L. Magrini et al. 2023), while radial and azimuthal variations have been detected in several nearby galaxies as well (e.g., L. Sánchez-Menguiano et al. 2016; M. Mollá et al. 2019; K. Kreckel et al. 2020; Z. Li et al. 2023), which could be taken as an indication for the uniqueness of abundances in star clusters found at particular galactocentric locations. However, S. Blanco-Cuaresma et al. (2015) find significant overlap of multiple chemical abundances in their sample of 31 OCs within the MW. Moreover, Y.-S. Ting et al. (2015) show that overdensities in chemical space do not guarantee that such overdensities arise due to a single set of stars from a common birth cloud, thus suggesting an overlap in the intercluster metallicity distribution.

A differing perspective has been found by D. L. Lambert & A. B. S. Reddy (2016), who show significant intercluster variation in the chemical composition of heavy elements (La, Ce, Nd, and Sm) for red giants in a sample of 28 OCs, where all the clusters had nearly solar metallicity. Moreover, C. Manea et al. (2024) find that neutron-capture elements carry a more discerning signature than the lighter elements that have been traditionally considered in the current generation of spectroscopic surveys. However, in the limit of elemental abundances available in APOGEE (S. R. Majewski et al. 2016), M. Ness et al. (2018) find that field stars have similar chemistry as the members of OCs, suggesting an overlap of birth cluster signatures with field stars. Furthermore, in a different study, M. K. Ness et al. (2019) find that, at a fixed [Fe/H], chemistry is a deterministic property of age that does not necessarily change with the birth location.

Perhaps the most critical consideration in assessing the viability of chemical tagging is not simply how chemically homogeneous OCs are themselves, but how homogeneous stars within a cluster are relative to random field stars. The contamination rate using 20 abundances in M. Ness et al. (2018) was found to be around 1% at $[\mathrm{Fe/H}]=0$. That is, while the intracluster dispersion is small, on the order of <0.03 dex, around 1% of random field stars at solar [Fe/H] are as chemically similar as stars within the same cluster. This is very prohibitive for any chemical-tagging pursuit. However, what remains critically unclear is the theoretical expectation for inter- and intracluster homogeneity from simulations. In this work, we test this expectation using galaxy simulations that resolve OCs in a cosmological context.

Previous simulation studies have shown that turbulent mixing plays an important role in homogenizing molecular clouds and reducing stellar abundance scatter (C. F. McKee & J. C. Tan 2002). This can be particularly significant during cloud assembly, potentially explaining the observed chemical homogeneity of stars that originate from the same molecular cloud (Y. Feng & M. R. Krumholz 2014). Extending this work, L. Armillotta et al. (2018) zoom in on the collapse of one giant molecular cloud extracted from a galaxy simulation and study the formation of individual stars down to a spatial resolution of $\approx 10^{-3}$ pc. They find that the star formation process defines a natural size scale of ~ 1 pc for chemically homogeneous star clusters, suggesting stars within ~ 1 pc of each other share similar chemical properties. However, these simulations assume metals are well coupled with gas, overlooking the potential effects of metals in the form of dust grains. Taking a different approach to identify co-natal stars, H. Kamdar et al. (2019) evolve four billion stars over 5 Gyr in a realistic potential with various galactic structures. They find that combining chemical and phase-space information enhances the identification of co-natal populations, suggesting that comoving pairs of stars with velocity separation <2 km s^{-1} and metallicity separation <0.05 dex are likely to be co-natal. Nevertheless, these simulations lack direct *N*-body effects on clusters and individual stars, as well as essential feedback processes.

In this paper, we examine chemical abundance trends in OCs identified in high-resolution cosmological hydrodynamic galaxy simulations generated with the Feedback in Realistic Environments (FIRE-2; P. F. Hopkins et al. 2018) model. These simulations have an adaptive spatial resolution, which enables modeling a large dynamic range (from megaparsecs down to parsecs), and include many relevant star-cluster-scale feedback processes. In this work, we utilize three MW-mass galaxies from the Latte suite (A. R. Wetzel et al. 2016; A. Wetzel et al. 2023) of FIRE-2 simulations, to assess the viability of strong chemical tagging via intercluster and intracluster chemical homogeneity analysis. Importantly, this

study is possible now because FIRE-2 implements physically motivated processes (e.g., subgrid turbulent metal diffusion, stellar feedback, and chemical enrichment from core-collapse supernovae, white dwarf supernovae, and stellar winds) and has high enough resolution in Latte (up to $\sim$1 pc) to distinguish individual star-forming regions.

Previously in FIRE-2, M. A. Bellardini et al. (2021) showed that their MW-mass galaxies have azimuthal scatter in gas at $z = 0$ similar to that observed in nearby galaxies of similar mass. M. A. Bellardini et al. (2022) investigated abundance trends in all young stars across the Latte disks, considering the evolution over time. They found that azimuthal variation dominates at $z \gtrsim 0.8$, whereas radial gradients dominate at late times. Subsequently, A. Carrillo et al. (2023) explored distributions in mono-age populations in one Latte disk, m12i, as a function of [Fe/H], [X/Fe], birth, and present-day locations and found some evidence for inside-out radial growth for stars with ages $<$7 Gyr. A. Carrillo et al. (2023) also examined the age–[X/Fe] relation across the disk and found that the qualitative trends agree with observations (apart from C, O, and Ca) with small intrinsic scatter ($0.01 < \sigma < 0.04$ dex). Additionally, A. Carrillo et al. (2023) found σ to be metallicity-dependent, with $\sigma \approx 0.025$ dex at $[\mathrm{Fe/H}] = -0.25$ dex versus $\sigma \approx 0.015$ dex $[\mathrm{Fe/H}] = 0$ dex. A similar metallicity dependence is seen in the GALAH survey for the elements in common (S. Sharma et al. 2022; A. Carrillo et al. 2023). Moreover, like M. A. Bellardini et al. (2022), A. Carrillo et al. (2023) found that σ is higher in the inner galaxy, where stars are older and formed in less chemically homogeneous environments.

Building on the work of M. A. Bellardini et al. (2022), R. L. Graf et al. (2024) recently showed that the radial metallicity gradient as a function of stellar age in the Latte disks follows a similar trend as in the MW, albeit with shallower gradients in FIRE. Additionally, R. L. Graf et al. (2024) showed that at the present day, the radial metallicity gradient is steepest for the youngest stellar populations, which is again qualitatively consistent with the MW, despite weaker trends in FIRE. Finally, both M. A. Bellardini et al. (2022) and R. L. Graf et al. (2024) considered the azimuthal scatter of [Fe/H] in stars in Latte and found it to be similar to the MW (relative to F. Anders et al. 2014).

While the FIRE studies discussed above focus on all stars in simulated galactic disks, our focus in this study is on chemical trends in OCs specifically. Along these lines, M. Y. Grudić et al. (2023) mapped giant molecular clouds formed self-consistently in a Latte simulation onto a cluster population according to a giant-molecular-cloud-scale cluster formation model calibrated to higher-resolution simulations. This approach enabled them to explore the galaxy's star clusters as a function of mass, metallicity, space, and time. M. Y. Grudić et al. (2023) concluded that massive clusters do not form with metallicities differing from other stars forming within the galaxy.

In this work, we take a different approach: we directly identify young massive OCs in the Latte simulations at the present day to measure their chemical homogeneity in the full galactic context. As of yet, a robust comparison between the chemical homogeneity of observed OCs and OCs formed in cosmological MW-mass galaxy simulations does not exist, and we aim to characterize the precision requirements for metallicity measurements to ensure that star cluster reconstruction could be conducted using chemistry alone. In the long term, we hope to identify clusters, let these disperse into the field, and try to reconstruct them using chemistry alone. However, the reassembly of clusters once mixed into the field, where there is a wide range of chemistry and dynamics at play even for a mono-age population, makes reconstruction a very difficult endeavor at this time. Thus, in this work, we take an intermediary step toward the reconstruction goal and check to see if the chemistry of each cluster is realistic and sufficiently distinct from one another even before dispersing into the field. This essential premise is a necessary condition for strong chemical tagging to be viable. In this study, we leverage OCs identified within the last $\sim$100 Myr of the simulations and select OCs younger than 3 Myr, which allows us to study truly co-forming populations often still embedded in their natal environments. We calculate the intra- and intercluster metallicity dispersion for our OC population and compare it to results from spectroscopic OC observations. Finally, we calculate a chemical difference metric between intracluster and intercluster pairs of stars and discuss the viability of chemical tagging using such chemical differences.

This paper is structured as follows. In Section 2, we discuss our data and methodology. Specifically, in Section 2.1 we introduce the simulations; in Section 2.2, we discuss how we identify OCs in these simulations; and in Section 2.3, we introduce the chemical difference metric we use to characterize the similarity of elemental abundances. In Section 3, we present our results, and finally, in Section 4, we present our discussion in the context of the observational literature and our conclusions.

2. Data and Methodology

2.1. Simulations

We use the Latte suite of FIRE-2 cosmological MW-mass galaxy simulations (A. R. Wetzel et al. 2016; A. Wetzel et al. 2023) to study the properties of OCs. As mentioned earlier, FIRE stands for Feedback in Realistic Environments; the code is able to simulate the interplay between the Interstellar Medium (ISM) and the stellar feedback processes in a cosmological environment (P. F. Hopkins 2015; P. F. Hopkins et al. 2018). As detailed in P. F. Hopkins (2015), these simulations use the Lagrangian meshless finite-mass hydrodynamics method implemented in Gizmo. Gizmo solves the fluid equations using a moving particle distribution that is automatically adaptive to resolution.

FIRE-2 uses a physically motivated metallicity-dependent radiative-transfer heating and cooling approach for gas that includes free–free, photoionization, and recombination processes, Compton scattering, photoelectric heating, and collisional dust effects, as well as molecular, metal-line, and fine-structure processes. The implementation of FIRE we use in this work additionally includes magnetohydrodynamics, where the equations of ideal magnetohydrodynamics are solved explicitly (P. F. Hopkins et al. 2020).

We use versions of these simulations that we resimulated over the final 110 Myr, specifically to store finer snapshot time spacing (1 Myr) than in the original simulations ($\approx$22 Myr). This enables us to track OCs in detail in phase space across time. While these resimulations are not publicly available, snapshots of the original simulations are publicly available (A. Wetzel et al. 2023).

FIRE-2 tracks 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe) across a temperature range of 10 to 10^{10} K. These simulations incorporate a spatially uniform, redshift-dependent UV background, as described by C.-A. Faucher-Giguère et al. (2009). As typically done by observers, metallicity and elemental abundances are scaled to solar values; in these simulations, we use values for the Sun from M. Asplund et al. (2009).

Star formation proceeds according to the following recipe in these simulations: each star particle is generated from a self-gravitating gas cell that is Jeans-unstable, cold ($T < 10^4$ K), dense ($n > 1000\,\mathrm{cm}^{-3}$), and molecular (following M. R. Krumholz & N. Y. Gnedin 2011); when star formation criteria are met, there is a one-to-one conversion of a gas cell into a star particle. That is, a star particle acquires its mass and elemental makeup from the gas cell it originates from and depicts a simple stellar population characterized by a Kroupa initial mass function (P. Kroupa 2001). Each star particle is individually evolved assuming a standard stellar population model (STARBURST99 v7.0; C. Leitherer et al. 1999) and produces time-resolved stellar feedback from both core-collapse and white dwarf supernovae. The rates of core-collapse supernovae are based on STARBURST99 (C. Leitherer et al. 1999) and the rates of white dwarf supernovae are from F. Mannucci et al. (2006). FIRE-2 also incorporates stellar feedback due to continuous mass loss, radiation pressure, photoionization, and photoelectric heating. The nucleosynthetic yields we use are based on K. Iwamoto et al. (1999) for white dwarf supernovae and K. Nomoto et al. (2006) for core-collapse supernovae. Stellar wind yields are sourced primarily from O, B, and AGB stars; we use a model for this assembled by R. P. C. Wiersma et al. (2009) that incorporates expectations from L. B. van den Hoek & M. A. T. Groenewegen (1997), P. Marigo (2001), and R. G. Izzard et al. (2004). Further details of this can be found in Appendix A of A. Carrillo et al. (2023) and in P. F. Hopkins et al. (2018).

FIRE-2 incorporates subgrid diffusion and the mixing of elemental abundances in gas, which occurs through unresolved turbulent eddies (K.-Y. Su et al. 2017; I. Escala et al. 2018; P. F. Hopkins et al. 2018). Our model assumes that subgrid mixing is primarily influenced by the largest unresolved eddies, which effectively smooth out abundance variations among gas elements. I. Escala et al. (2018) demonstrate that incorporating this subgrid model is essential for accurately replicating the observed distributions of stellar metallicities in galaxies. M. A. Bellardini et al. (2021) show that while the details of this diffusion model do not significantly affect the radial or vertical gradients in FIRE-2 MW-mass galaxies, they do affect the azimuthal scatter. M. A. Bellardini et al. (2021) also show that these galaxies have azimuthal scatter in gas at $z = 0$ similar to that observed in nearby galaxies of similar mass.

The Latte suite contains several MW-like galaxies; in this work, we analyze three Latte simulations (m12f, m12i, and m12m). The galaxies we analyze represent some of the most massive MW-like isolate disks in the Latte suite and span a range of assembly histories. Key differences among these simulations are that m12i has a massive disk with a late-forming MW-mass halo, m12f has an LMC-like satellite merger with the MW-like disk 3 Gyr prior to the present day, and m12m has a strong bar at high redshift and an earlier-forming halo (see R. E. Sanderson et al. 2020). A full description of the criteria used to generate these galaxies and their properties at the present day can be found at A. Wetzel et al. (2023) and references therein. A brief description of the relevant details for this work follows. Each galaxy simulation assumes a flat ΛCDM cosmology with parameters consistent with the Planck Collaboration et al. (2020): $h = 0.68$–0.71, $\Omega_\Lambda = 0.69$–0.734, $\Omega_m = 0.266$–0.31, $\Omega_b = 0.0455$–0.048, $\sigma_8 = 0.801$–0.82, and $n_s = 0.961$–0.97.[8]

In this work, "MW-like" refers to the total mass of each galaxy at the present day. The Latte galaxies were selected to have halo masses roughly the same as the MW's: $M_{200} = 1$–$2 \times 10^{12} M_\odot$, where M_{200} refers to the total mass within the radius containing 200 times the mean matter density of the Universe. These simulations contain three types of material: dark matter, stars, and gas. The resolution of each dark matter particle is $3.5 \times 10^4 M_\odot$ and the initial mass resolution of baryonic material is $7070\,M_\odot$. As noted earlier, each star particle should be thought of as a single stellar population. Because of this, as a star particle ages, it loses a portion of its stellar mass. Thus, at $z = 0$, a typical star particle has a mass of $\sim 5000\,M_\odot$. In the Latte galaxies, star and dark matter particles have a fixed gravitational force softening length of 4 and 40 pc (Plummer equivalent), respectively. Gas cells have adaptive gas smoothing: in regions where there is a high density of gas cells, the gravitational force softening length can resolve down to a minimum of 1 pc.

Given their spatial and mass resolution, the Latte simulations can resolve massive star clusters with multiple star particles that represent portions of each OC. And, as mentioned earlier, with a temporal resolution for snapshots of 1 Myr for these resimulations, they can track in detail how star particles in OCs evolve in phase space over time. See Section 2.2 for more details on OC selection and characteristics.

2.2. Star Cluster Identification

We identify OCs in our simulations using the friends-of-friends algorithm (M. J. Geller & J. P. Huchra 1983). It works by identifying members within a fixed distance known as the linking length; in this work, we adopt a linking length of 4 pc, which corresponds to our softening length. To ensure that the OCs are identified before the onset of core-collapse supernovae, we select OCs from star particles that are less than 3 Myr old over the last 110 Myr of these simulations. We require a minimum of five star particles for a cluster to be identified, and thus with our fiducial mass resolution ($7070\,M_\odot$), we resolve clusters down to a cluster mass of $10^{4.6}\,M_\odot$.

With these selection criteria, the OCs we identify have a characteristically small velocity dispersion ($\sigma_{1D} \sim 4\,\mathrm{km\,s}^{-1}$), a small size ($R_{\rm half} \sim 3$ pc), and a small spread in age ($\sigma_{\rm age} \sim 0.25$ Myr). We do not require that the natal giant molecular cloud be bound for our OCs to form inside them; despite this, roughly half of the clusters that we identify are born bound ($\log_{10}(2KE/PE) \geqslant 0$) and the global population follows a Gaussian distribution of boundedness peaked at 0. While the size and boundedness of the population that we identify may be better described as OCs and associations, for the purposes of this study, which aims to characterize the spread in metallicity inside and between small clustered star-forming environments

[8] Where h is the dimensionless equivalent of the Hubble Parameter H_0, Ω_Λ is the cosmological constant or vacuum density at the present time, Ω_m is the total matter (dark plus baryonic) density today, Ω_b is the baryonic matter density, σ_8 is the rms fluctuation of the density perturbations at 8 h^{-1} Mpc scale, and n_s is the scalar spectral index.

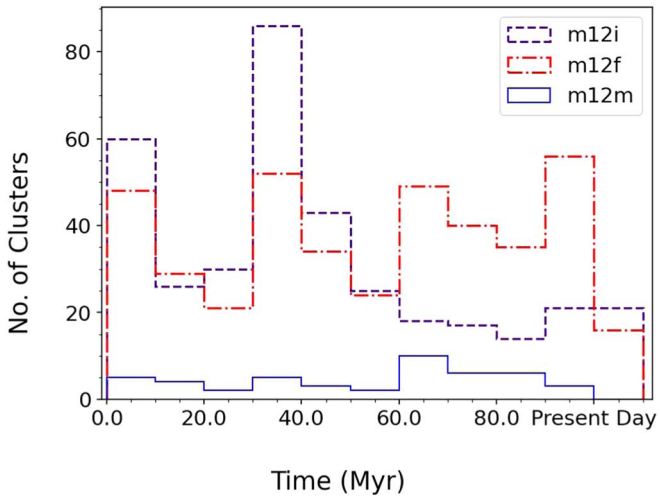

Figure 1. Number of young (<3 Myr) massive OCs identified in the last ∼100 Myr of the simulations. OC selection has been run at a 3 Myr interval to avoid double counting. m12i is indicated by the dashed lines colored in indigo, m12f is indicated by the dashed–dotted lines colored in red, and m12m is indicated by the solid blue lines. A varying rate of star cluster formation is observed in all three simulations, with m12m having the lowest number of clusters produced throughout the 100 Myr period. This reflects a globally suppressed SFR for m12m at the present day. See Section 2.2 for further discussion.

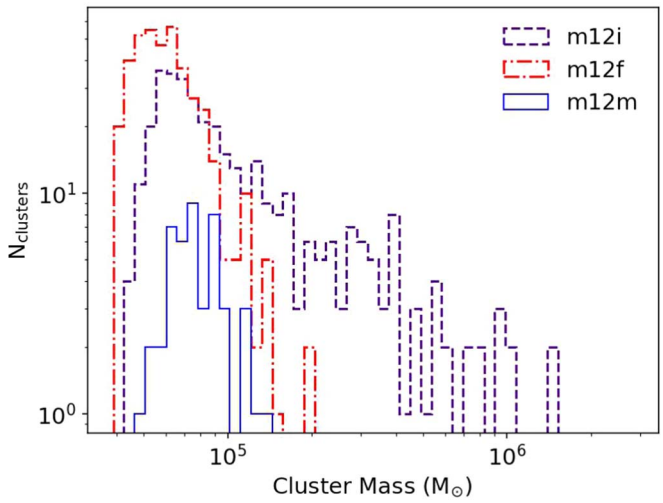

Figure 2. Mass distribution of the OCs from all three simulations. m12i is indicated by the dashed lines colored in indigo, m12f is indicated by the dashed–dotted lines colored in red, and m12m is indicated by the solid blue lines. The slope of the cluster-mass function is similar for all three simulations, with m12i having a shallower slope with a few high-mass outliers.

across a full galactic disk, we find it sufficient not to distinguish between OCs and associations. We also note that the age selection criteria that we adopt do not significantly impact any of these properties (D. Masoumi 2024). Because of this, we feel that the OC populations that we identify sufficiently resemble OCs within a full simulated cosmological MW-like galaxy environment to proceed with the elemental abundance study we present here; a more detailed study that focuses on the impact of the selection criteria we use on the resultant properties of the OCs we identify can be found in D. Masoumi (2024). Note we identify star clusters similarly in all three MW-like analog simulations mentioned in Section 2.1.

In this work, we identify a total of 361, 404, and 46 unique OCs in m12i, m12f, and m12m, respectively, using the criteria discussed above. We analyze the OCs from each galaxy simulation as a single data set. In Figure 1, we show a histogram of the number of clusters identified throughout the final 110 Myr of the simulations. These three simulations (m12i, m12f, and m12m) represent a varying rate of star cluster formation over the 110 Myr epoch; m12m has a relatively consistent albeit low number of clusters formed over this period of time, while m12i forms roughly a quarter of all the clusters over just a 3 Myr span, and finally m12f has a fluctuating rate of cluster formation that oscillates on a ∼30 Myr timescale. We note that the low rate of m12m's cluster formation reflects the overall depressed star formation rate (SFR) of the entire galaxy at the present day.

The cluster-mass distribution of these simulations is shown in Figure 2. We note that m12f and m12m have similar slopes for the cluster-mass function. However, m12i has a shallower slope and contains a few high-mass outliers that form during a period of vigorous star formation in the plane of the disk. It is worth noting that the SFRs at the present day of these galaxies are $\sim 10 M_{\odot}\,\mathrm{yr}^{-1}$, while the MW's current SFR has been estimated to be $2.0 \pm 0.7\ 1 M_{\odot}\,\mathrm{yr}^{-1}$ (D. Elia et al. 2022), and thus these simulations produce more high-mass clusters more frequently than the MW does. These OCs are best thought of as Young Massive Clusters (YMCs; see Table 1 from S. F. Portegies Zwart et al. 2010) and will be discussed further in D. Masoumi et al. (2024, in preparation). While the lowest-mass OC that we resolve is $3.89 \times 10^4 M_{\odot}$, which is very massive when compared with observations (e.g., T. Cantat-Gaudin 2022; T. Cantat-Gaudin & L. Casamiquela 2024; E. L. Hunt & S. Reffert 2024), there are multiple YMCs in the MW that fall in our resolvable mass limit (e.g., RSGC1, RSGC2, RSGC3, and Wd1; as described in S. F. Portegies Zwart et al. 2010, Table 2). We anticipate that the subsequent versions of these simulations will generate OCs to a lower mass limit. However, for the time being, the OCs in this paper are the highest-resolution sample generated in a cosmological context in an MW-like setting to date and offer the first opportunity to explore previously inaccessible connections between OC properties and the Galactic environment.

In Figure 3, we show a top-down view of m12i, m12f, and m12m near the present day (13.693 Gyr). Here, we show gas in the background, which has been binned in 200 pc x 200 pc squares and colored by the mean value of [Fe/H] in each volume. Plotted on top of the average [Fe/H] distribution for gas are contours of the density distribution of stars that are less than 50 Myr old. These contours highlight where there is active star formation corresponding to the spiral structure. The circles overlaid on top indicate the mean [Fe/H] of the OCs and associations (linking length $\leqslant$18 pc) that were identified at this moment in time; while we have included larger associations here for visualization purposes, in this paper we only include OCs found using a linking length of 4 pc for analysis. We note that Figure 3 shows that the average [Fe/H] of OCs and associations generally reflects the gas metallicity distribution across the disk.

We note that the top-down perspective on the galactic disks shown in Figure 3 highlights the overall radial metallicity gradient in all three galaxies; while each galaxy has a different normalization and slope of the radial metallicity gradient, overall it is clear that a negative radial gradient is present in the global disk for gas at the present day (see M. A. Bellardini et al. 2021 for further discussion). While here we have shown the

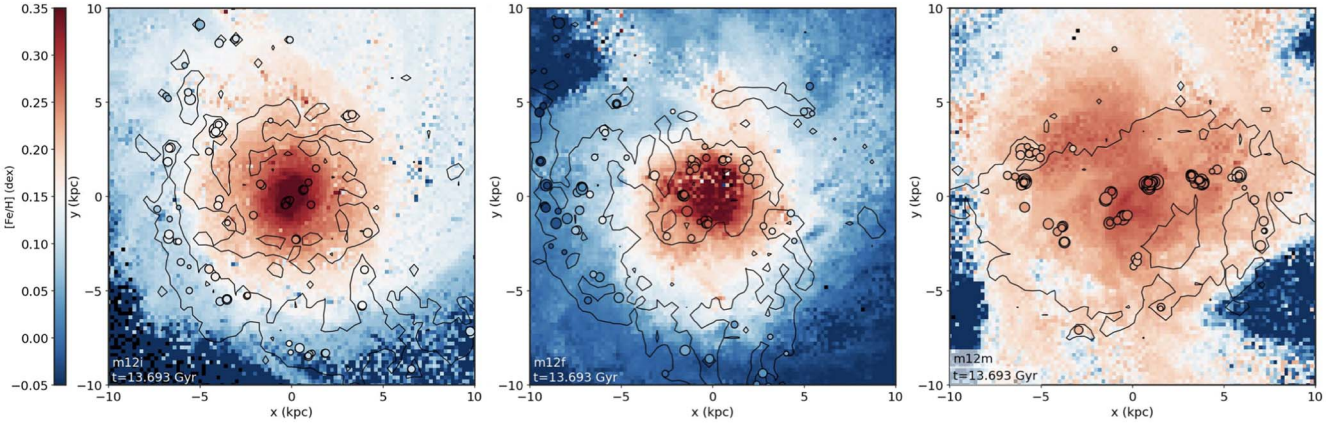

Figure 3. Snapshots from three Latte simulations of m12i, m12f, and m12m. Background: mean [Fe/H] for gas in the disk of each simulation. The corresponding values are shown by the color bar on the left. Contours: density of stars < 50 Myr old that highlight the spiral arms. Circles: mean [Fe/H] of OCs and associations identified at 13.693 Gyr. Note that the mean [Fe/H] of the clustered star formation reflects the global metallicity distribution of the gas in the disk.

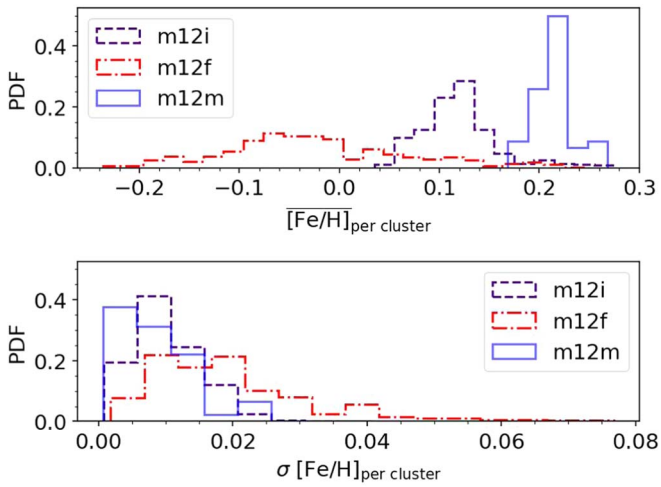

Figure 4. Distribution of mean [Fe/H] (upper panel) and dispersion of [Fe/H] in the cluster sample of all three simulations (lower panel). m12i is indicated by the dashed lines colored in indigo, m12f is indicated by the dashed–dotted lines colored in red, and m12m is indicated by the solid blue lines. m12f has a larger range of mean and scatter in [Fe/H].

distribution of [Fe/H], we will be using a total of nine elemental abundances (Fe, C, N, O, Mg, Si, S, Ca, and Ne) throughout the remainder of this work. We note that these elements are highly correlated with each other in our simulations. We show this correlation in Appendix A in Figures 10 and 12. Such a strong correlation is also seen in the MW (E. J. Griffith et al. 2024); in fact, M. K. Ness et al. (2022) show that, using a linear regression fit to Fe and Mg alone, one can predict eight supernova elements within 5% accuracy. Despite such strong correlations, we elect to use all elements at our disposal in this work.

In Figure 4, we show normalized histograms of the mean [Fe/H] (upper panel) and standard deviation of [Fe/H] (lower panel) for OCs in all three simulations. The upper panel shows the presence of scatter in the mean [Fe/H]. While m12f has a large range of OC mean [Fe/H], m12i and m12m have a narrower range of values. Each of these populations peaks at a different value of [Fe/H]; however, there is some overlap between these samples. The standard deviation histogram for each galaxy shown in the bottom panel of Figure 4 illustrates that there are very few OCs that have a large standard deviation of [Fe/H]; moreover, the peak of each distribution is at a small value. Given these essential properties—there is scatter in the average [Fe/H] in OCs in the simulations and there is a small spread in [Fe/H] in each OC—in this work, we are able to test the feasibility of uniquely identifying stars from the same OC using their elemental abundances alone.

2.3. Chemical Difference Metric

The goal of this investigation is to understand the similarity of stellar abundances for stars within a given OC; to simplify this, we define a simple chemical difference metric using a pseudo-Cartesian measurement of distance in abundance space. A similar approach is used in M. Ness et al. (2018), which was adapted from A. W. Mitschang et al. (2013). The metric is defined as

$$d^2_{nn'} = \sum_{i=1}^{I} [x_{ni} - x_{n'i}]^2,$$

where n and n' are two stars being compared, with star n containing the abundance x_{ni} and star n' containing the abundance $x_{n'i}$ for all elements 1 to I. Simply put, this metric evaluates the squared chemical difference between any two stars, summed over all available elements, and provides a single number that accesses the chemical similarity for each pair of stars. In this work, we initially evaluate the chemical difference metric summed over eight element abundances ([Mg/Fe], [Ca/Fe], [S/Fe], [Si/Fe], [Ne/Fe], [O/Fe], [N/Fe], and [C/Fe]), without any errors being considered.

After this initial assessment, we further consider the chemical difference metric with APOGEE errors added to our simulated OCs (drawn from the Chemical Abundances and Mapping Survey; N. Myers et al. 2022). The difference metric with added errors is defined as

$$d^2_{nn'} = \sum_{i=1}^{I} \frac{[x_{ni} - x_{n'i}]^2}{\sigma^2_{ni} + \sigma^2_{n'i}},$$

where σ_{ni} refers to the error measurement for x_{ni} and $\sigma_{n'i}$ represents the error measurement for $x_{n'i}$. Note that we omit Ne (neon) from the above list of elements in this analysis, as it is a relatively rare element that is not measured in the APOGEE OCs.

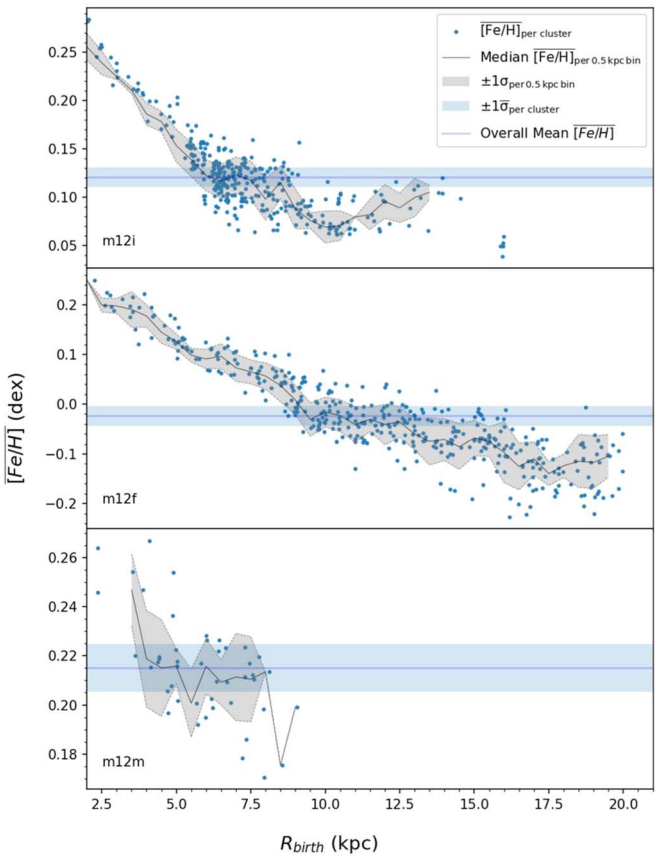

Figure 5. Scatter plot of the mean [Fe/H] vs. birth radius of OCs from three simulations: m12i (top panel), m12f (middle panel), and m12m (bottom panel). The blue dots indicate the mean [Fe/H] of each cluster and the solid black curve represents the median trend for data in 0.5 kpc radial bins. The gray shaded area represents the $\pm 1\sigma$ range within 0.5 kpc bins. The blue horizontal line represents the average of all OCs in the sample, and the blue horizontal shaded region represents the typical $\pm 1\sigma$ scatter of [Fe/H] within any given cluster. It can be seen that the metallicity distribution varies differently in the inner disk than the outer, and many clusters have a similar mean [Fe/H], despite being found at varying locations across the disk.

3. Results

3.1. The Relationship Between [Fe/H] and Birth Radius for OCs

In this section, we consider the radial metallicity gradient: the average [Fe/H] as a function of galactocentric radius, which is often thought to be a unique indicator of where stars are born (e.g., I. Minchev et al. 2018). Here, we measure the radial metallicity gradient using OCs, as has often been done in the literature (K. A. Janes 1979; P. Sestito et al. 2008; L. Spina et al. 2022a; L. Magrini et al. 2023). Our goal is to assess how truly unique the mean [Fe/H] is for the OCs in our simulations. We note that the dispersion in each cluster does not change substantially across the disk.

In Figure 5, we show the mean [Fe/H] versus the birth radius ($R_{\rm birth}$) for all OCs identified over the final 110 Myr in m12i (top panel), m12f (middle panel), and m12m (bottom panel). Note that the range of values indicated on the y-axes are slightly different for each galaxy, which is very important when comparing the slopes between each system. Each blue dot indicates the mean [Fe/H] value of an individual OC. We calculate a running median using a bin size of 0.5 kpc (shown in a black solid line) and indicate the $\pm 1\sigma$ confidence region around the median values, shaded in gray. In addition to this, we indicate in each figure the global average [Fe/H] (solid blue horizontal line) of all our OCs and also the average dispersion of [Fe/H] (shaded blue region) inside a typical OC. We later subsample OCs drawn from this blue-shaded region to perform a chemical-difference-metric test (see Section 3.3).

We find a very different trend in the metallicity gradient in the three galaxies we assess. For m12i (top panel), there is a steep gradient within $R_{\rm birth} < 6$ kpc (with a slope of ~ -0.02 dex kpc^{-1}) and a small scatter about the median trend line (with scatter < 0.01 dex). This combination of a steep slope and small scatter means that, within this region, [Fe/H] strongly constrains the birth location. However, at $R_{\rm birth} > 6$ kpc, the slope of the radial metallicity gradient substantially flattens out and the scatter significantly increases (> 0.02 dex). Moreover, outside of 10 kpc, there is an inversion in the slope of the metallicity gradient. Taken together, there is no longer a one-to-one mapping of [Fe/H] and birth radius outside 6 kpc and [Fe/H] should not be taken as a reliable indicator of the birth location.

However, when we consider m12f (middle panel), the predictive power of [Fe/H] as a unique indicator of birth location extends to a significantly larger region (out to 10 kpc). Within this radial range, there is a small scatter around the median trend line ($<$0.02 dex). Moreover, like m12i, within the inner region of m12f, the slope has a constant negative value of -0.04 dex kpc^{-1}. It is noteworthy that, for m12f, outside of 10 kpc, the scatter increases significantly, doubling to $\sim$0.04 dex in the the outer regions of the disk. And, like m12i, at the largest Galactocentric birth radii, there is a subtle upturn in the median trend line. Thus, again, at large Galactocentric radii, [Fe/H] no longer has strong predictive power as an indicator of birth location.

Finally, m12m (lower panel), has a significantly smaller number of clusters than the other two systems, and these clusters are all found at Galactocentric radii < 9 kpc. While the slope of the radial metallicity gradient for m12m is negative, it is substantially shallower (< -0.008 dex kpc^{-1}) than m12i and m12f. Moreover, while the typical scatter about the median trend line is similar to the inner region of m12i ($\sim$0.01 dex), in combination with the relatively shallow gradient, there is limited discerning power for [Fe/H] to indicate birth location in this system.

Overall, all three systems appear to have a negative gradient at small radii that generally flattens at larger radii, consistent with the analysis of these galaxies in M. A. Bellardini et al. (2022). Such a gradient and flattening in the average metallicity of OCs as a function of radius has often been modeled in the MW using a two-component fit (e.g., J. Donor et al. 2020; N. Myers et al. 2022; L. Spina et al. 2022a). The presence of such a transition in the gradient is an indicator that the variation in abundances is different for different regions of our Galactic disk.

Note that in Section 3.3, we will discuss the shaded blue region in Figure 5 and further consider the additional power of individual abundances to draw connections to the birth location. This can be done by selecting OCs that fall within $\pm 1\sigma$ of a fixed [Fe/H] value. As we mentioned earlier, σ is set by the typical dispersion within an OC, while the fixed [Fe/H] value is set by the average [Fe/H] of all the OCs discussed here.

In Figure 6, we show the metallicity distribution function for a subset of OCs found within the fixed [Fe/H] range discussed above. These histograms were generated using a Gaussian kernel-smoothing technique. Each curve is color-coded by where the

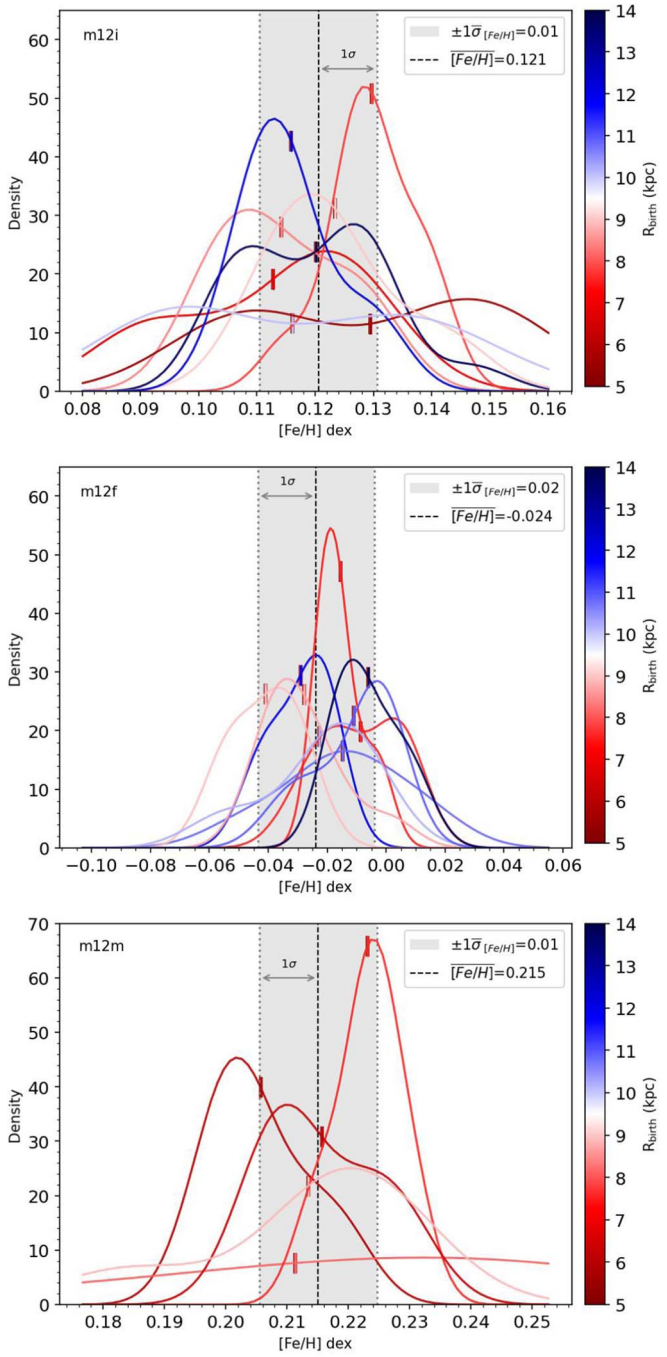

Figure 6. The [Fe/H] distribution for a subset of OCs from the subsample highlighted in blue from Figure 5 for each simulation. Each smoothed histogram represents the metallicity distribution for a single OC and was generated using Gaussian kernel smoothing. The shaded region indicates the $\pm1\sigma$ from the overall mean of the distribution of all clusters from that simulation. Despite their locations in the disk, it is possible to find OCs that have similar abundance distributions at different R.

OCs formed in the disk. The density shown on the *y*-axis is calculated such that the area under the curve is 1, using a bin width of 0.001. The black dashed line along the center replicates the horizontal line in Figure 5, which represents the global mean [Fe/H] for all clusters in a given simulation, and the gray-shaded region inside the black dotted lines indicate the 1σ dispersion from the typical standard deviation inside each OC.

The clusters shown in Figure 6 represent a subset of all clusters shown in the blue-shaded region of Figure 5 and probe a large range of $R_{\rm birth}$ across the disk. The short vertical line on each curve indicates the mean value of [Fe/H] for that OC. Here, we can see that while these clusters have similar mean [Fe/H] (all fall within the gray-shaded region), their individual metallicity distribution function can vary in shape. While these clusters form at a very large range in radii, their metallicity distribution function does not map uniquely onto their formation location. Notably, the mean [Fe/H] of a young OC does not always map strictly onto the formation location, nor does the dispersion of [Fe/H] inside a given OC. This reinforces the idea that the mean metallicity may not be enough to uniquely determine the formation locations of OCs.

3.2. Inter- and Intracluster Metallicity Dispersion

The intracluster dispersion provides a measure of the chemical homogeneity of member stars within a given cluster for a given element. If such a measurement is small, it indicates a similarity in the elemental abundance of member stars inside a cluster. Here, we define the intracluster dispersion as the weighted mean of the dispersion inside each cluster across all OCs for each simulation. Here, the weight is proportional to the number of star particles in each cluster.

Similarly, intercluster dispersion can provide a measure of the scatter in the mean elemental abundance between different clusters. A large value of intercluster dispersion indicates a significant difference in the mean elemental abundance (e.g., [Fe/H]) between OCs. One might expect the intercluster dispersion to be higher than the intracluster dispersion, especially when comparing clusters across a large range of radii. In this limit, one might expect strong chemical tagging to be feasible. With a larger dispersion between clusters and a smaller dispersion inside a cluster, it would be easier to identify stars that belong to the same. However, as we saw in Section 3.1, regions of the disk can have a weak radial metallicity gradient. Because of this, it is valuable to quantify the comparative significance of the intra- versus intercluster dispersion to test the feasibility of strong chemical tagging. We will further discuss the comparative significance of these dispersions utilizing a chemical difference metric in the next section.

In Figure 7, we show the intra- and intercluster dispersion for all OCs in each simulation for each of the nine elements that we trace (Fe, C, N, O, Mg, Si, S, Ca, and Ne). M12i, m12f, and m12m are shown in the top, middle, and bottom panels, respectively; within each panel, the lower curve in blue indicates the weighted average intracluster dispersion per element, hereafter represented by the symbol $\bar{\sigma}_{[X/H]}$. The vertical indigo lines corresponding to each element indicate the $\pm1\sigma$ scatter of the intracluster dispersion in all clusters. Since this is a dispersion of the dispersion, it was calculated by first converting these standard deviations into variances. For n standard deviations, σ_1, σ_2,…,σ_n, the variance of the sample standard deviations can be calculated as

$$\mathrm{Var}(\sigma_i) = \sigma_i^2.$$

We then calculate the average of these variances as

$$\mathrm{Mean(Var)} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}(\sigma_i).$$

Finally, the standard deviation of the standard deviation is

$$\text{Standard Deviation}(\sigma_\sigma) = \sqrt{\mathrm{Mean(Var)}}.$$

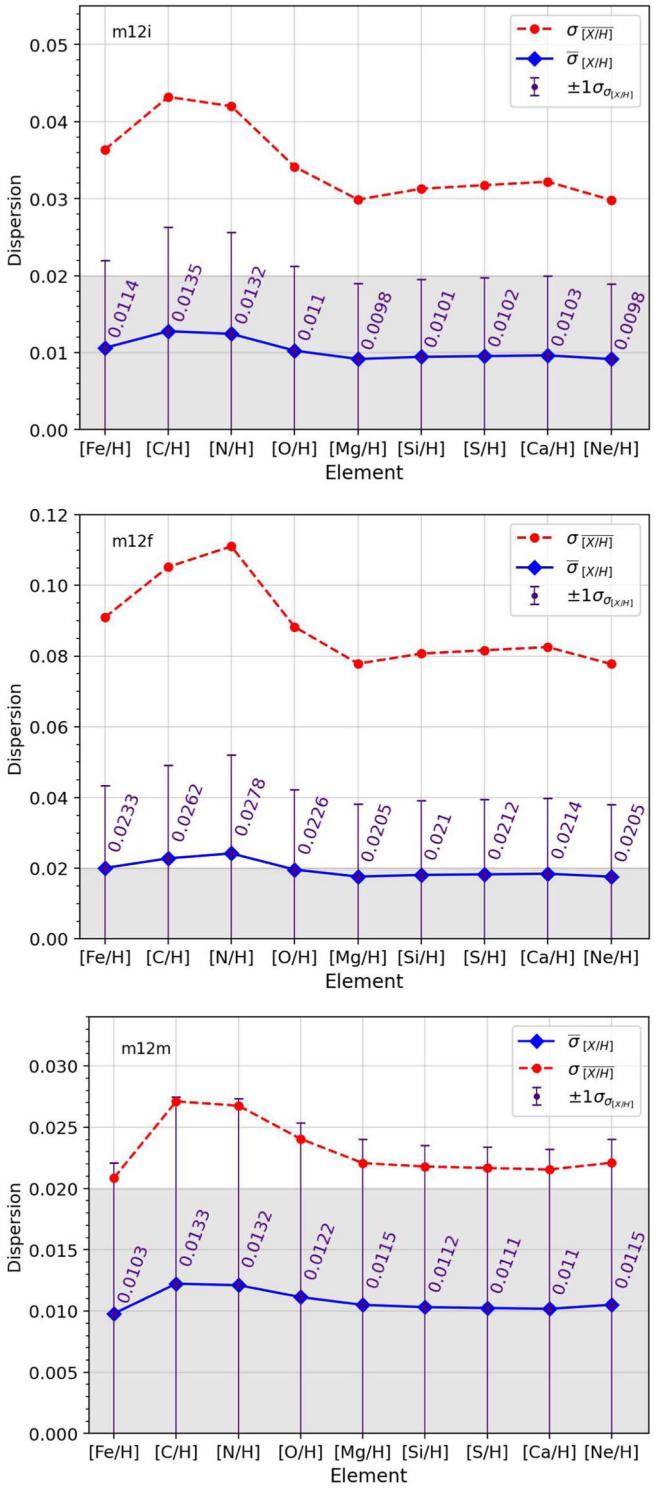

Figure 7. For each of the simulations in each panel: blue—the intracluster dispersion indicated as the mean dispersion per element; and indigo—the dispersion around the average dispersion. The indigo splines indicate the $\pm 1\sigma$ around it. The gray shaded region highlights the typical measurements of such dispersion from observations, with the gray horizontal line indicating the upper limit. Dashed red—intercluster dispersion. m12f has larger dispersion in both inter- and intracluster abundances, most probably due to an LMC-like merger resulting in a more mixed disk compared to other simulations that do not incorporate such a merger.

The typical measurements of the intrinsic dispersion from observations for these elements fall under the shaded gray region (drawn from M. Ness et al. 2018). The upper red curve indicates the measurement of intercluster dispersion (hereafter represented by the symbol $\sigma_{[\bar{X}/H]}$). This has been calculated as the weighted standard deviation of the mean elemental abundance, $[\bar{X}/H]$, in each cluster.

As indicated by these measurements, the average intracluster dispersion inside each FIRE OC is quite small and generally less than 0.020 dex; these measurements are comparable to the intrinsic dispersion calculated by several observational studies, which find that OCs have intrinsic dispersion <0.03 dex (e.g., S. Bertran de Lis et al. 2016; J. Bovy 2016; M. Ness et al. 2018; V. J. Poovelil et al. 2020). This is also consistent with previous studies of m12i, which have considered the global dispersion for stars in the disk at the present day. As mentioned in Section 1, A. Carrillo et al. (2023) investigated the intrinsic dispersion at fixed metallicity for all disk stars in m12i and found it to be ≈0.025 dex at [Fe/H] = –0.25 dex and ≈0.015 dex at [Fe/H] = 0 dex. The dispersion of the intracluster dispersion shown by the vertical indigo splines spanning $\pm 1\sigma$ of the intracluster dispersion tells us the width of the distribution of the intrinsic scatters. These measurements are also very small, which is an indication that our simulated clusters have some element scatter that is nonzero, which is similar for all clusters.

Similarly, as indicated by the upper red curves, the intercluster dispersion measurements for all three simulations m12i, m12f, and m12m are larger than the intracluster dispersion, as expected. We measure mean intercluster dispersions of ≈0.03 dex for m12i, ≈0.09 dex for m12f, and ≈0.02 dex for m12m. We further test if we can use individual element abundances to discriminate between clusters. To do this, we examine pairs of stars within and between clusters to measure their chemical similarity. A requirement of reconstructing the formation conditions of stars using only their individual abundances is that the chemical distance between stars within a cluster (intracluster pairs) is smaller than that of stars between clusters (intercluster pairs). A minimum requirement for a pair of stars to be from the same cluster is that they have the same [Fe/H]. We therefore now work only in a narrow [Fe/H] bin near the overall mean [Fe/H] of the OCs in each simulation.

3.3. Intra- and Intercluster Chemical Difference Metric

In this section, we discuss the chemical difference metric to see if we can distinguish between two stars that are formed in one OC and two stars randomly drawn from different OCs. We employ the chemical difference metric described in Section 2.3, which calculates the squared chemical difference between two stars for each element and is summed over all elemental abundances. Such a metric allows us to consider how chemically similar two stars are via one number, which is small (approaching zero) when two stars are similar.

We take the same sample of OCs highlighted in blue from Figure 5 that is within $\pm 1\sigma$ of the galaxy-wide mean [Fe/H]. We use this sample to calculate the chemical difference between member stars, to explore if we can discern the difference within and between clusters. We first draw all possible pairs of stars from within an OC and calculate the chemical difference for each pair. This procedure is iteratively repeated for each OC in our sample, and the results are presented in the blue histogram in Figure 8. To maintain a comparative framework, we apply the same metric for an equivalent number of random pairs, but this time, each star is

drawn from a different randomly selected OC. This distribution is shown in the red histogram in Figure 8. We form 17,747 intra- and intercluster pairs in m12i, 1,685 pairs in m12f, and 374 in m12m.

In Figure 8, the blue histogram consists of the difference metric for the pairs of stars drawn from the same OC. The majority of the pair's chemical differences fall within the smallest bins of the histogram for all three simulations. If two stars are measured with such a small chemical difference, one would naively assume they come from the same OC. However, these smallest-chemical-difference-metric bins are also significantly populated by stars from intercluster pairs, as indicated by the red histograms. In m12i, ≈86/73% of the total intra-/intercluster pairs have a difference metric of <0.0002. Similarly, in m12f, ≈66/43% of the total intra-/intercluster pairs have a difference metric of <0.0002. Finally, in m12m, ≈84/65% of the total intra-/intercluster pairs have a difference metric of <0.0002. Because of this, if we observe two stars with a chemical difference of <0.0002, it would be extremely difficult to know if these two stars came from the same OC or were drawn from two different OCs.[9] That is, for the majority of pairs, it would be difficult to distinguish between stars that fall within the same OCs or different OCs using the vector of abundances for these elements alone. Given that the typical uncertainties in the abundance measurements are ≈0.03 dex (M. Ness et al. 2018), it is unlikely that strong chemical tagging is feasible. To further illustrate the impact of the uncertainties on the observed chemical differences, we calculate the chemical difference metrics for intra- and intercluster pairs after adding the APOGEE errors, which we discuss in Section 3.4.

To better quantify the degree of inherent similarity between the two distributions shown in Figure 8, we perform a two-sample Kolmogorov–Smirnov (K-S) test. Essentially, such a K-S test quantifies the likelihood that the observed data sets were drawn from the same underlying distribution and generates the probability that the observed differences between the two in-hand data sets are due to chance selection. There are two metrics reported from a K-S test: the K-S statistic and the *p*-value; both metrics need to be considered together to fully understand the significance of the differences between the observed data sets. The K-S statistic measures the maximum discrepancy between the cumulative distributions of two data sets. The *p*-value indicates the probability of observing such a discrepancy by random chance alone if the null hypothesis is true; in a two-sample K-S test, the null hypothesis is that the two observed data sets were drawn from the same underlying distribution. Values of the K-S statistic and the *p*-value range between 0 and 1. A high value of the K-S statistic (>0.6) and a small *p*-value (<0.05[10]) suggest that the observed samples are very different and are very likely to be drawn from two different underlying distributions, whereas a low value of the K-S statistic (<0.4) and a small *p*-value (<0.05) suggest that while the observed samples are similar, it is still very likely that they are drawn from two different underlying distributions.

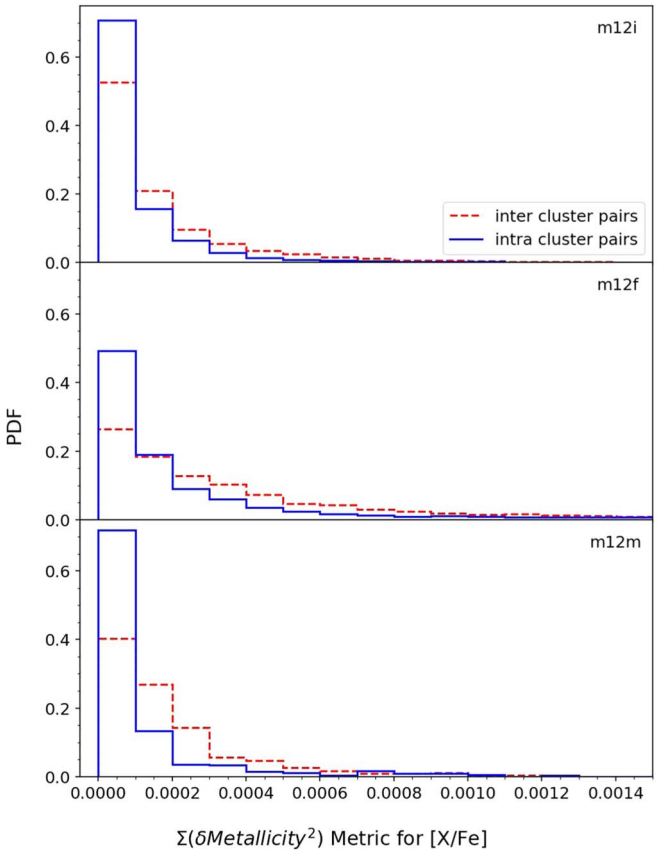

Figure 8. Histograms of the chemical difference metric for pairs of stars within OCs indicated by blue histograms (intracluster pairs) and between OCs indicated by the dashed red histograms (intercluster pairs) for each simulation m12i (top), m12f (middle), and m12m (bottom) in each panel. These clusters were selected to fall within $\pm 1\sigma$ from the overall mean distribution of [Fe/H] in each simulation. The bin with the smallest chemical difference metric obtained for both inter- and intracluster pairs is abundantly populated, indicating that such a difference is pretty small to uniquely identify the chemistry of individual clusters.

We obtain a K-S statistic of 0.19 for m12i, which indicates a small maximum difference between the Probability Density Function (PDFs) of the calculated chemical difference metric for the inter- and intracluster pairs. However, the *p*-value is quite small (∼0.00 for m12i), thus indicating that it is unlikely that these two samples are drawn from the same global distribution. In other words, while there are some similar aspects to both data sets (the maximum discrepancy in the global PDF is small), they are not in fact drawn from the same parent distribution. This can be seen visually by inspecting the red and blue lines in the top panel of Figure 8; while both distributions are strongly peaked at the lowest values of the chemical difference metric, at larger values, the shape of the intercluster distribution (shown in red) is different from the shape of the intracluster distribution (shown in blue). The K-S statistic is 0.25 for m12f; this is slightly higher than m12i's K-S statistic, but still a critically low value, indicating that there is similarity in the two observed m12f data sets. However, again, the *p*-value is very low (∼0.00 for m12f), which indicates that

[9] We choose <0.0002 as a selection criterion here, as it is close to the mean chemical difference in all three simulations (see Table 3 for relevant statistics for the chemical difference metric for each simulation). In Appendix B, we explore the impact of changing the statistic we use and of the number of elements we leverage to discern between intra- and intercluster pairs. Table 1 shows that regardless of the statistic we use for subdividing our sample for the smallest-chemical-difference-metric values, we always see significant contamination with intercluster pairs. We note that using fewer elements to calculate the chemical difference metric marginally decreases the ability to discern between the intra- and intercluster pairs at small values of the chemical difference metric.

[10] What is considered a "small" *p*-value is based on the chosen level of significance. In this work, we choose the level of significance to be 5%, thus we can trust the K-S statistic if the *p*-value is < 0.05.

it is highly unlikely that both data sets are drawn from the same underlying distribution. Finally, m12m has the largest K-S statistic (∼0.37) of all three simulations; however, it still falls below the 0.4 threshold, thus we can say that the maximum differences in m12m's distributions are quite slight. Despite this, yet again, the *p*-value is vanishingly small (∼0.00 for m12m), indicating that it is extremely unlikely that these samples were drawn from the same parent distribution.

Overall, we conclude while the intra- and intercluster chemical-difference-metric data sets were drawn from two different global distributions, there is little power with this set of elements in using small values of the chemical difference metric to select stars that definitively came from one OC. That is, the contamination rate of intercluster pairs at the smallest chemical difference metric is far too significant to trust OC reconstruction.

3.4. Intra- and Intercluster Chemical Difference Metric with APOGEE Errors in Simulated OCs

In order to test how the incorporation of errors in abundance measurements affects our chemical-difference-metric calculations, in this section, we add errors drawn from an observational cluster catalog to our simulated OCs. The errors we use come from the Open Cluster Chemical Abundances and Mapping Survey (OCCAM VI; N. Myers et al. 2022), which was derived from APOGEE DR17 (Abdurro'uf et al. 2022). To add errors to the simulation, we first generate a normal distribution for each element for each star; the mean of the distribution is set to the true value from the simulation. We set the standard deviation of the distribution by calculating the mean error for each element from the OCCAM cluster catalog. We then draw from this distribution a single value to represent the chemical abundance with error included, which is assigned to each star. This process is repeated for all elements and all stars. In this analysis, we use the seven elements ([C/Fe], [N/Fe], [O/Fe], [Mg/Fe], [Si/Fe], [S/Fe], and [Ca/Fe]) in common with the OCCAM sample. We use the same sample of clusters discussed in Section 3.3 for our analysis here.

In Figure 9, each panel shows the histograms for the chemical difference metric from the intra- and intercluster star pairs with APOGEE errors included; we show each distribution with a blue solid and red dashed line, respectively. Again, the top, middle, and bottom panels correspond to m12i, m12f, and m12m, respectively. With errors included, the shape and the extent of the distribution shown in Figure 9 changes substantially. The distribution now peaks closer to the degree of freedom that is equal to the number of elements used to calculate the chemical difference metric. That is, each element contributes equally to the error-included chemical difference metric. Unlike the histograms of the metrics with no error in abundance measurements, which peaked at the smallest chemical difference, the histograms now peak at ≈7. All three panels from different simulations show a similar trend with the addition of errors, and most of the bins until the metric value of 15 for each simulation are almost equally populated by both intra- and intercluster pairs. Considering a limit of 7, which is closer to the mean of the chemical difference metrics for both intra- and intercluster pairs, ≈55/56% of the total intra-/intercluster pairs have the difference metric of <7 in m12i. Similarly, for m12f, ≈55/54% of the total intra-/intercluster pairs have a difference metric of less than 7. For m12m, the percentages of pairs with the difference metric of <7 are

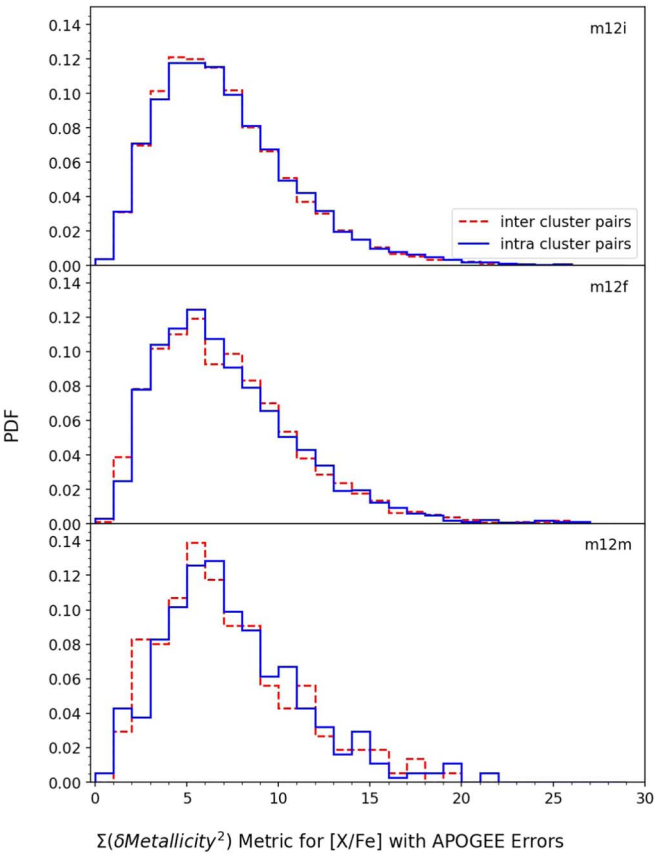

Figure 9. Histograms of the chemical difference metric for pairs of stars with mock APOGEE errors in OCs (blue) and between OCs (dashed red) for each simulation. These clusters were selected to fall within the mean $\pm 1\sigma$ of the typical intracluster dispersion from the overall mean distribution of [Fe/H] in each simulation. Most of the bins are abundantly populated with the chemical difference metrics obtained from both inter- and intracluster pairs, indicating that such a difference is pretty small to uniquely identify the chemistry of individual clusters.

≈52/56% for intra-/intercluster pairs, respectively. Also, the maximum PDF obtained for either the intra- or intercluster pairs is slightly below ≈0.13 at the mean difference metric, thus indicating there is no prospect of identifying if the star pairs are coming from the same cluster or different clusters. This analysis clearly shows that, with errors, there is not enough chemical difference between the stars to identify them uniquely.

4. Discussion and Conclusions

In this paper, we have characterized the degree of scatter in abundance distributions in young OCs identified in three galaxies from the Latte suite of FIRE-2 simulations (m12i, m12f, and m12m). Using a range of elements (C, N, O, Ne, Mg, Si, S, Ca, and Fe) that are generated from three distinct nucleosynthetic families (core-collapse supernovae, white dwarf supernovae, and stellar winds), we have shown that the abundance dispersion in FIRE OCs is comparable to the observations from M. Ness et al. (2018).

In this section, we compare our results with other observational studies that have measured the level of chemical homogeneity present in individual OCs using spectroscopic analysis. Dispersions measured from observations are influenced by two factors: intrinsic scatter—the true difference in

elemental abundance between stars within an OC—and observational noise, which originates from measurement uncertainties. Most observational studies consider an OC to be chemically homogeneous if the metallicity dispersion inside a cluster is less than or equal to the typical measurement uncertainties. While measurement uncertainties vary between studies (e.g., 0.03 dex versus 0.008–0.036 dex, from F. Liu et al. 2016 and M. Ness et al. 2018, respectively), typically the measurement uncertainty is $\leqslant$ 0.03 dex (Y.-S. Ting et al. 2012; J. Bovy 2016).

Simulations do not suffer from observational uncertainties, and the typical intrinsic dispersion inside a simulated Latte OC is < 0.02 dex. This is broadly consistent with the intrinsic dispersion measured for these elements from observational studies. In Latte, such homogeneity is a consequence of two dominant factors: (1) at the present day, the gas within individual GMCs is very well mixed (see Figure 7 of M. A. Bellardini et al. 2021); and (2) the star formation within an individual OC is typically short-lived ($\sigma_{\rm age} < 1$ Myr; D. Masoumi, et al. 2024, in preparation), which limits internal enrichment.

Note that some level of inhomogeneity is expected in observations driven by, among other things, atomic diffusion, mixing processes, and planet engulfment (S. Blanco-Cuaresma et al. 2015; C. Bertelli Motta et al. 2018; L. Spina et al. 2021). Such processes are expected to have an impact on the metallicity of evolved populations and are not resolved in Latte OCs, as star particles have a fixed abundance throughout their lifetime. In observational studies, dispersion measurements that make use of both main-sequence and giant stars within an OC are expected to have an increased metallicity dispersion; as a result, some studies consider dwarf and giant stars separately (S. Blanco-Cuaresma et al. 2015; L. Casamiquela et al. 2020). Moreover, different types of elements (e.g., light versus heavy elements) are expected to have a different amount of scatter (G. M. De Silva et al. 2009; C. Manea et al. 2024). While the current generation of Latte simulations do not track heavy elements, such an improvement is anticipated in upcoming generations of FIRE.

With this somewhat complicated observational context in mind, we now present results from the literature. Note that, as far as possible, we try to include both the stellar populations and the elements studied in each work. In one of the earliest works on chemical homogeneity, G. M. De Silva et al. (2006) calculated chemical scatter below their measurement uncertainties for dwarf stars in the Hyades OC. In a follow-up study, G. M. De Silva et al. (2007) analyzed abundances of the light elements in 12 red giant stars in the OC Collinder 261 and estimated the intrinsic scatter to be $<$0.05 dex across these elements.

In line with G. M. De Silva et al. (2006, 2007), many other groups have found that OCs are chemically homogeneous from spectroscopic analysis (e.g., E. Pancino et al. 2010; Y.-S. Ting et al. 2012; A. Bragaglia et al. 2014; K. Cunha et al. 2015). J. Bovy (2016) used APOGEE spectra for 49 giants in M67, NGC 6819, and NGC 2420 and showed that the scatter is $<$0.03 dex for 15 elemental abundances (C, N, O, Na, Mg, Al, Si, S, K, Ca, Ti, V, Mn, Fe, and Ni). Among these elements, C and Fe had strong limits ($<$ 0.02 dex at 95% confidence); however, for Na, S, K, Ti, and V, the scatter was $<$0.05 dex. V. J. Poovelil et al. (2020) also calculated the abundance scatter in APOGEE red giants in 10 OCs for eight elements—Mg, Al, Si, Ca, Fe, Si, Mn, and Ni—and found that the dispersions for these clusters were within the limits indicated by J. Bovy (2016), of $<$0.03 dex for most of the elements, with only Mg, Al, and Si having larger scatter, which could be consistent with dredge-up effects that preferentially impact light elements. Recently, A. Sinha et al. (2024) studied 26 OCs across 20 elements and showed that we can place 3σ upper limits on OC homogeneity within 0.02 dex or less for most of the elements. Our typical intrinsic scatter measurements of $<$0.02 are in line with these results.

However, some studies have found measurable chemical inhomogeneity larger than the measurement uncertainties in OCs (e.g., S. Blanco-Cuaresma et al. 2015; F. Liu et al. 2016; C. Bertelli Motta et al. 2018; L. Casamiquela et al. 2020). For example, F. Liu et al. (2016) studied the abundances of 19 elements (C, O, Na, Mg, Al, Si, S, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, and Ba) in 16 solar-type stars in the Hyades cluster, with abundance errors ranging from a low of 0.008 dex for Si to a high of 0.036 dex for S. For most elements, the abundance dispersion they measured was significantly larger than the measurement errors, by a factor of $\approx$1.5–2. They concluded that Hyades is chemically inhomogeneous at the $\geqslant$0.02 dex level, which contrasts with the conclusions for Hyades drawn from G. M. De Silva et al. (2006) and G. M. De Silva et al. (2007). F. Liu et al. (2016) suggest the pollution of metal-poor gas in portions of the protocluster cloud combined with supernova ejection in other regions with incomplete mixing as the root cause of such dispersion.

Overall, we find that different observational studies have differing criteria for determining if an OC is chemically homogeneous. This conclusion can be impacted by the stellar sample used (containing evolved, main-sequence, or a combination of the two populations), the elements considered (light and/or heavy elements), and the measurement errors (which are instrument-resolution- and line-fitting-methodology-dependent). Dwarf stars and heavy elements show the greatest discriminating power for the measurement of homogeneity. Most studies agree that the typical level of intracluster scatter inside MW OCs is $\leqslant$ 0.03 dex for most elements; OCs with scatter less than this per element relative to H are considered chemically homogeneous. In our simulations, the intrinsic OC intracluster scatter in [X/H] for C, N, O, Ne, Mg, Si, S, Ca, and Fe is $\leqslant$ 0.02 dex; these Latte OCs would be considered chemically homogeneous by most observational studies.

While there is abundant research supporting small intracluster dispersion in OCs, the assumption of large intercluster dispersion, which is crucial for strong chemical tagging to work, has only recently been explored (e.g., S. Blanco-Cuaresma et al. 2015; D. L. Lambert & A. B. S. Reddy 2016). One of the earliest studies to consider intercluster dispersion was G. M. De Silva et al. (2009). Using 10 elements from 24 OCs, they showed varying levels of dispersion for each element. In particular, they found an intercluster scatter of $\geqslant$0.20 dex in heavy elements like Mn and Ba, which they concluded is an indication of abundance variation between clusters.

In the time since, there have been several significant studies that investigate chemical abundances in large spectroscopic samples of OCs surveyed in a homogeneous manner. These include the Bologna Open Clusters Chemical Evolution project (A. Bragaglia & M. Tosi 2006), the WIYN Open Cluster Study (R. D. Mathieu 2000; H. R. Jacobson et al. 2011), Gaia ESO

OC papers such as L. Magrini et al. (2014), and APOGEE studies such as the OCCAM survey (N. Myers et al. 2022). While these studies provide tremendous insights into radial trends in the MW, unfortunately, as of yet, they do not explicitly consider intercluster trends.

Subsequent to G. M. De Silva et al. (2009), A. W. Mitschang et al. (2013) used 30 distinct OCs from the literature and measured both the inter- and intracluster dispersion. They found a typical intercluster dispersion of ∼0.120 dex and considerably smaller intracluster scatter of ∼0.045 dex (i.e., the intercluster dispersion is ∼three times greater than the intracluster dispersion). Thereafter, the first study (to our knowledge) that considered intercluster scatter is S. Blanco-Cuaresma et al. (2015). In this work, they compiled stellar spectra from 31 old and intermediate-age OCs to determine 17 abundance species and showed that stars at different evolutionary stages have distinct chemical patterns. Upon separating stars into dwarfs and giants, they found only a few OCs show distinct chemical signatures, with the majority of them possessing a high degree of overlap. They also analyzed five OCs (IC4651, M67, NGC 2447, NGC 2632, and NGC 3680) with stars in various evolutionary stages and found that the differences in abundances for [Fe/H] and [X/Fe] were less than 0.05 dex for almost every element considered, except for Na, which can be enhanced in giants. Such enhancement likely occurs as a result of mixing from the first dredge-up in giant stars.

On the other hand, D. L. Lambert & A. B. S. Reddy (2016) studied intercluster abundance differences for 23 elements in 70 red giants from 28 OCs. Their analysis was comparable in scope and accuracy to that reported by S. Blanco-Cuaresma et al. (2015), except that their selection included several additional heavy elements—La, Ce, Nd, Sm, and Eu—that are produced by neutron-capture processes. Using a set of 20 clusters, they calculated the scatter in light elements, which was 0.16 dex for Na and Al, 0.12 dex for Mg and Si, and 0.11 dex for Ca (see Table 1 from D. L. Lambert & A. B. S. Reddy 2016). They observed intercluster scatter up to 0.5 dex for heavy elements like La and Ce, with uncertainty <0.07–0.08 dex. D. L. Lambert & A. B. S. Reddy (2016) point out that while the intercluster scatter for heavy elements is large, many clusters are individually chemically homogeneous in heavy elements with unique chemical patterns. They suggest that combining measurements of one or more light elements like Mg, Al, Si, Cu, and Zn with heavy elements like La, Ce, Nd, Sm, and Eu in future chemical-tagging studies might produce insightful results (D. L. Lambert & A. B. S. Reddy 2016).

Overall, at present, there is limited discussion of intercluster scatter in the literature and differing perspectives are offered in the two most recent works. S. Blanco-Cuaresma et al. (2015) have found a high degree of overlap (small intercluster scatter) in the chemical signatures of light elements across OCs; however, they do not directly quantify the intercluster scatter in their analysis. On the other hand, D. L. Lambert & A. B. S. Reddy (2016) have found a significant intercluster scatter in heavy neutron-capture elements (up to 0.5 dex) and a smaller scatter in lighter elements (on average ≈0.13 dex). As previously mentioned, at this time, we only track light elements in our simulations. In this limit, the OC intercluster scatter we measure for [X/H] for C, N, O, Ne, Mg, Si, S, Ca, and Fe is two to five times larger than the intracluster dispersion in individual OCs. Specifically, when averaged across all nine elements, the inter- versus intracluster dispersion is for m12i: 0.03 versus 0.01; for m12f: 0.09 versus 0.02; and for m12m: 0.02 versus 0.01. We find that when the intercluster scatter is considered by element, only C and N have a slightly higher scatter than the average (albeit it at a very small level; see Figure 7). This is consistent with the result presented by A. W. Mitschang et al. (2013), which showed that the intercluster dispersion is ∼three times greater than the intracluster dispersion. D. L. Lambert & A. B. S. Reddy (2016) have shown intercluster dispersion to be >10 times larger than the intracluster dispersion for heavy elements like La, Ce, Nd, and Sm (e.g., 0.4 dex of the intracluster scatter versus 0.03 dex of the intercluster scatter for La). However, if we assume that the individual OCs from D. L. Lambert & A. B. S. Reddy (2016) are homogeneous at the ≈0.03 dex level, then the ratio of the inter- to intracluster scatter for light elements is ∼4, which is in line with our results.

As one might expect, the intercluster scatter that we report for OCs is less than what has been seen for all young stars in the Latte disks. For example, M. A. Bellardini et al. (2022) selected young stars (< 500 Myr old) at the present day from 11 of the Latte galaxies at a range of radii and showed that at small radii, the scatter across the entire annulus is 0.03–0.04 dex. Moreover, at the present day, the typical variation across the fiducial solar cylinder ($R = 8$ kpc) in [Fe/H] is ≈0.05 dex (see the top panel of Figure 9 of M. A. Bellardini et al. 2022; R. L. Graf et al. 2024). However, at this annulus, the intercluster scatter for OCs in [Fe/H] is only 0.02 dex.

In addition to intra- and intercluster scatter, we have calculated a chemical difference metric similar to the one first introduced by A. W. Mitschang et al. (2013). Such a metric helps to determine if one can distinguish pairs of stars formed in the same OC from pairs formed across different respective OCs. We calculated the percentage of pairs that have small chemical-difference-metric values (less than the mean chemical difference for the entire sample) in the hope that this population would be predominantly populated by intracluster pairs. However, for m12f, 43% of the intercluster and 66% of the intracluster pairs had similar indistinguishably small chemical difference values. Moreover, in the other two galaxies, the rate of chemical difference overlap was even greater: 65% versus 84% for m12m and 73% versus 86% for m12i in inter- versus intracluster pairs. We find it noteworthy that a significant fraction of the smallest-chemical-difference values belong to intercluster pairs, which one would typically assume to be populated by only intracluster pairs.

A. W. Mitschang et al. (2013) used the chemical difference metric to compare 30 distinct OCs from across the literature. Their abundance measurements ranged from five to 23 elements per cluster across 291 stars. They found a distinct peak in the distribution of the chemical difference metric for both the intra- and the intercluster distributions. In their analysis, the two distributions had a critical crossing point at a chemical metric value of ∼0.07—at this value, both the intra- and intercluster star pairs showed identical chemical difference distributions and thus the metric had no constraining power for strong chemical tagging. In our simulated OCs, such a crossing point occurs at a very small chemical difference value (∼0.0001); this crossing point is near the median of both skewed Gaussian distributions and is very well sampled by both intra- and intercluster pairs. Notably, the chemical difference metric has no constraining power to determine the

origin at this location. In our extended analysis of the chemical difference metric, we add APOGEE errors to the abundance measurements of our OCs. We find that the error-added chemical metric distributions for intra- and intercluster pairs overlap significantly at almost all measured values, further reinforcing that there is little prospect of discerning intra- and intercluster pairing using the chemical difference metric from light elements alone.

In a slightly different scenario, leveraging 600 intracluster pairs and 1,018,581 field pairs with similar log(g) and $T_{\rm eff}$ values, M. Ness et al. (2018) utilized 20 element abundances to calculate the chemical difference metric. The chemical difference distributions for intracluster and field star pairs were visibly distinct but not totally disjoint, with 0.3% of field stars having similar chemical differences as the median difference in intracluster pairs. When stars of a fixed solar metallicity ([Fe/H] $= 0 \pm 0.02$) were considered, the percentage of field star pairs with overlapping intracluster chemical difference values increased to $\sim$1.0%. M. Ness et al. (2018) refer to these field stars as solar "doppelgangers." They conclude that such a significantly populated doppelganger population implies that strong chemical tagging in a strict sense would not work with this data set. While we are intrigued by this result, we leave a similar calculation utilizing our OC sample and Latte field stars to future work.

In conclusion, the OCs in FIRE show realistic intracluster homogeneity (typical OC elemental dispersion $\lesssim$0.02 dex) and comparable intercluster dispersion to observations (two to five times larger than intracluster dispersion). However, when using a chemical difference metric to distinguish between intra- and intercluster pairs, we do not find significantly large chemical differences to distinguish the OC origin for the majority of stars. This is likely due to limited and/or redundant information encoded in the light elements we track (C, N, O, Ne, Mg, Si, S, Ca, and Fe); with the addition of heavy elements, the chemical difference metric may produce more reliable results for chemical tagging in the future. However, as informed by our simulated OC sample (drawn from three realistic MW-mass cosmological galaxy simulations with varied merger histories and environments), in the limits of light elements only, we are dubious that strong chemical tagging can reconstruct individual OCs reliably to a high level of confidence.

Acknowledgments

B.B. and S.L. would like to thank Peter Frinchaboy and Natalie Myers for help with integrating APOGEE OCCAM errors into the simulation data. B.B. acknowledges support from the Center for Computational Astrophysics (CCA), which enabled the remote participation in the 2021 predoctoral program.

B.B. acknowledges support from NSF grant AST-2109234. S.L. acknowledges support from NSF grant AST-2109234 and HST grant AR-16624 from STScI. A.W. received support from: NSF, via CAREER award AST-2045928 and grant AST-2107772; NASA ATP grant 80NSSC20K0513; and HST grant GO-16273, from STScI. E.C.C. acknowledges support for this work provided by NASA through the NASA Hubble Fellowship Program grant HST-HF2-51502 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555.

We generated simulations using: XSEDE, supported by NSF grant ACI-1548562; Blue Waters, supported by the NSF; Frontera allocations AST21010 and AST20016, supported by the NSF and TACC; and Pleiades, via the NASA HEC program through the NAS Division at Ames Research Center.

FIRE-2 simulations are publicly available (A. Wetzel et al. 2023) at http://flathub.flatironinstitute.org/fire. Additional FIRE simulation data are available at https://fire.northwestern.edu/data. A public version of the GIZMO code is available at http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html.

Software: Python (F. Perez & B. E. Granger 2007), Matplotlib (J. D. Hunter 2007), Numpy (C. R. Harris et al. 2020), Scipy (P. Virtanen et al. 2020), `halo_analysis` (A. Wetzel & S. Garrison-Kimmel 2020a), `gizmo_analysis` (A. Wetzel & S. Garrison-Kimmel 2020b).

Appendix A
Correlation Between Different Elements

In this appendix, we consider the scatter and correlation between elemental abundances for all stars in all OCs found in m12i and presented in this paper. A discussion of how yields are implemented in FIRE-2 can be found in Appendix A of A. Carrillo et al. (2023) and in P. F. Hopkins et al. (2018).

In Figure 10, we consider [X/H] versus [Fe/H] for X = C, N, O, Mg, Si, and Ca. We find that each of these elements is highly correlated with [Fe/H], as indicated by the Pearson correlation coefficient (>0.99 in all cases). This can be seen by comparing the trend in the data in each panel to the diagonal red dotted one-to-one line. Not only is there limited scatter in elemental abundance at any value of [Fe/H] (as shown in Figure 11, which presents [X/Fe] versus [Fe/H], to highlight the limited scatter at any value of [Fe/H]), but the data in Figure 10 also follow a simple linearly increasing trend in all panels, with a slope that is nearly identical to the one-to-one line. While there are some slight departures from a single linear distribution (e.g., at [Fe/H]$\sim$0.15, a transition to a shallower slope for [Mg/H] and a transition to a steeper slope for [C/H] and [N/H]), to first order, all the elements we consider in this paper carry largely the same chemical signature as [Fe/H]. That is, modulo a normative offset, given a value of [Fe/H] for any star in any of our OCs in m12i, [X/H] can be predicted with high confidence for C, N, O, Mg, Si, and Ca.

To further reiterate this point, in Figure 12 we explore the relationships of individual elements with one another by plotting [X/H] versus [Y/H] for X,Y = C, N, O, Si, Mg, and Ca. Here, we color-code each star by its [Fe/H] value to reflect the trend shown in Figure 10. Again, we see that these elements are highly correlated with one another when considering a Pearson correlation coefficient or inspecting the trends relative to the one-to-one line. Of all the elements, [Mg/H] is the least correlated with [C/H], with a Pearson correlation coefficient of 0.989. Indeed, there is a slight scatter in [Mg/H] for small values of [C/H] and a modest transition in slope at high values of [C/H]. To this end, in our subsequent chemical-difference-metric analysis presented in Appendix B, we utilize Mg and C as maximally independent elements. However, it bears repeating that all of the elements presented here carry very similar discriminating power. Note that while this analysis indicates that chemical information is redundant in our simulations, something similar has been found in the MW. For example, E. J. Griffith et al. (2024) and M. K. Ness et al. (2022) have shown that if Fe and Mg are measured alone, then eight other supernova elements can be predicted to within 5% of their true values using a simple linear regression (see Figure 1 of M. K. Ness et al. 2022 for an analogous figure to Figure 12, but for the MW).

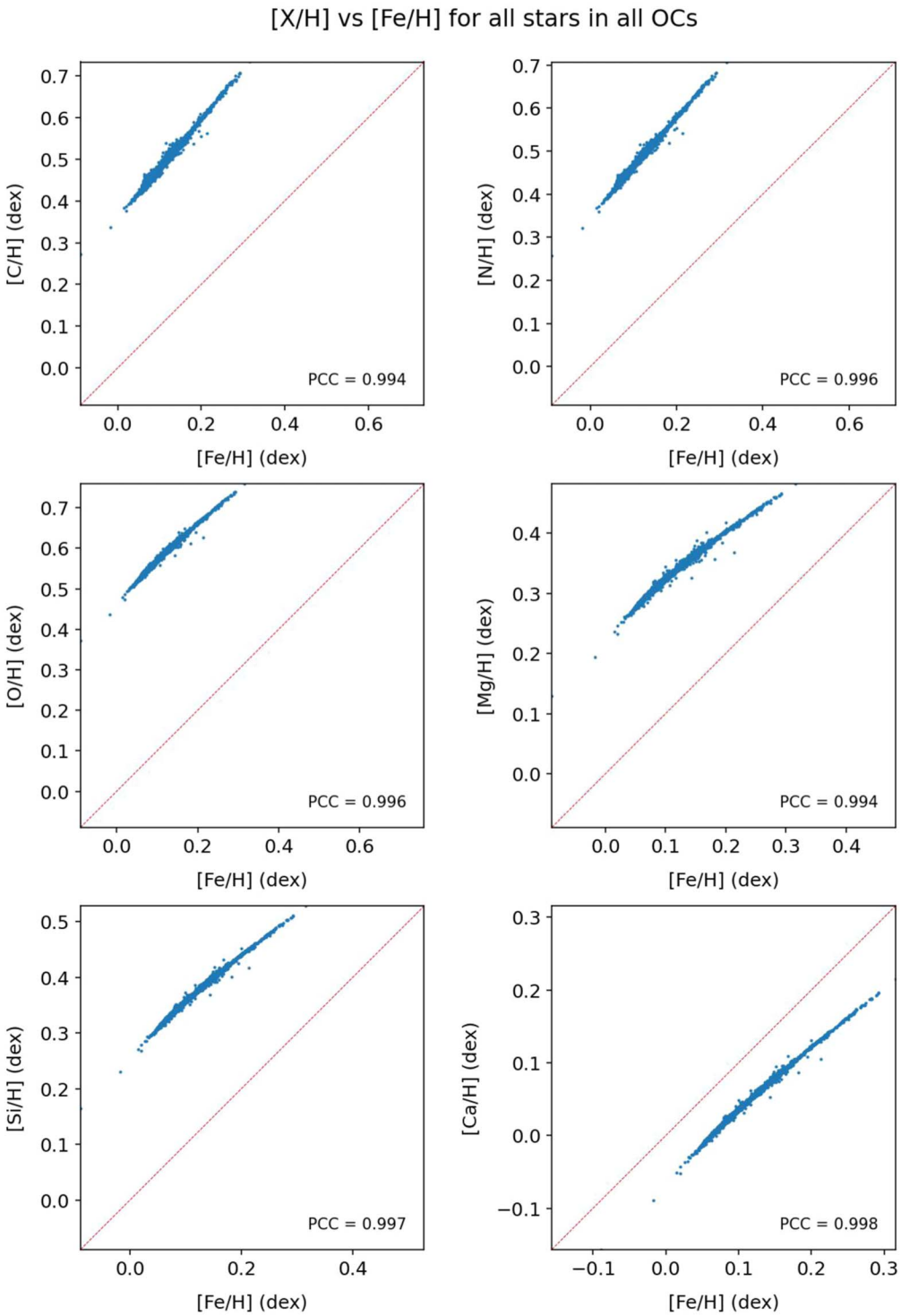

Figure 10. From left to right, top to bottom: [X/H] vs. [Fe/H] for C, N, O, Mg, Si, and Ca for stars from OCs presented in this paper in m12i. Elements are highly correlated with [Fe/H], as indicated by the Pearson correlation coefficient, which is labeled on each plot (>0.99 for all elements considered). The dotted diagonal red line is the one-to-one line. Note that the slope of the [X/H] vs. [Fe/H] is similar to the one-to-one line and there is limited scatter in [X/H] at any value of [Fe/H]. This indicates that all the elements considered carry similar discriminating power as [Fe/H].

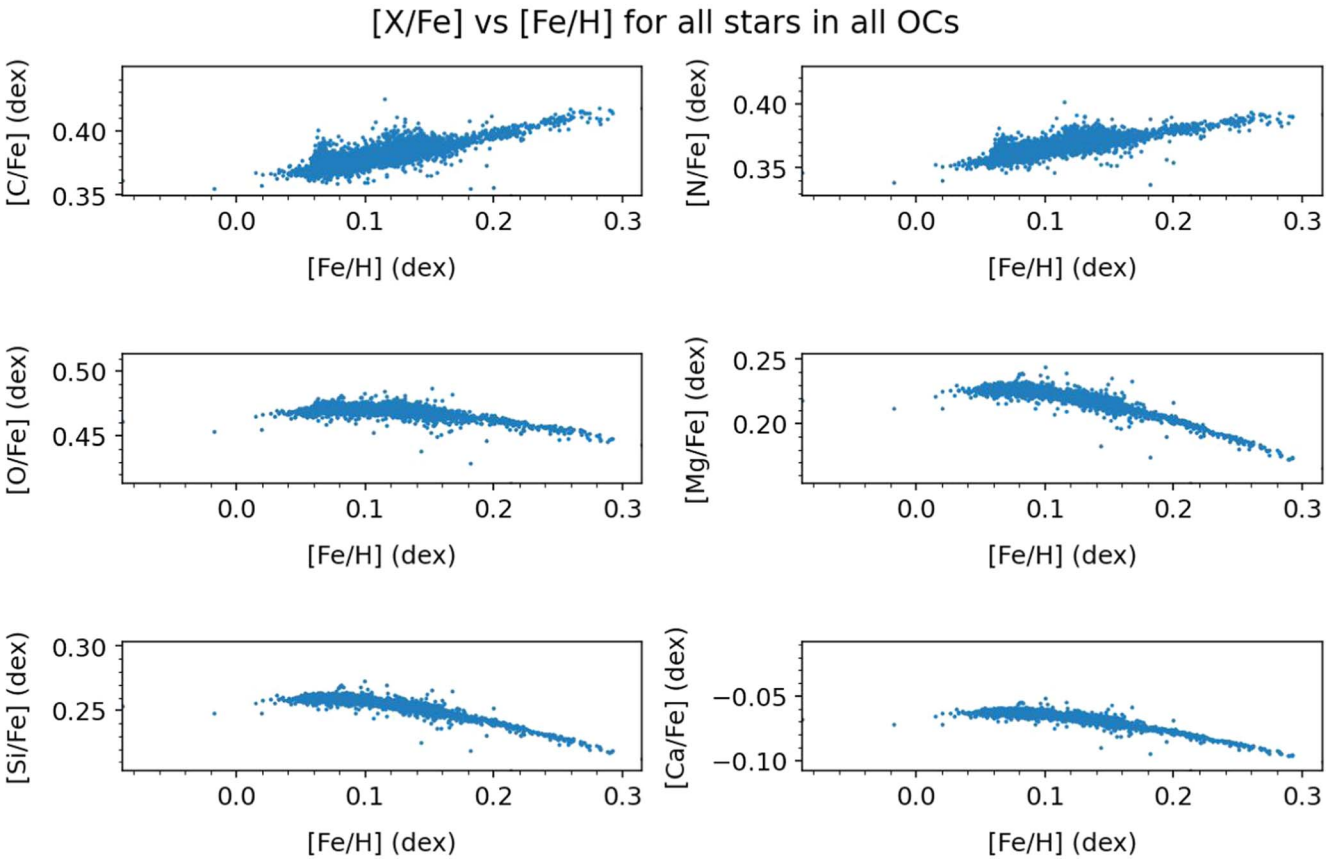

Figure 11. From left to right, top to bottom: [X/Fe] vs. [Fe/H] for C, N, O, Mg, Si, and Ca for stars from OCs presented in this paper in m12i. Elements have a very tight scatter for a given value of [Fe/H], as indicated by the small range in [X/Fe] shown on the *y*-axes. The scatter is more substantial at lower values of [Fe/H]; however, the scatter decreases substantially at [Fe/H]$\geqslant$0.19 dex.

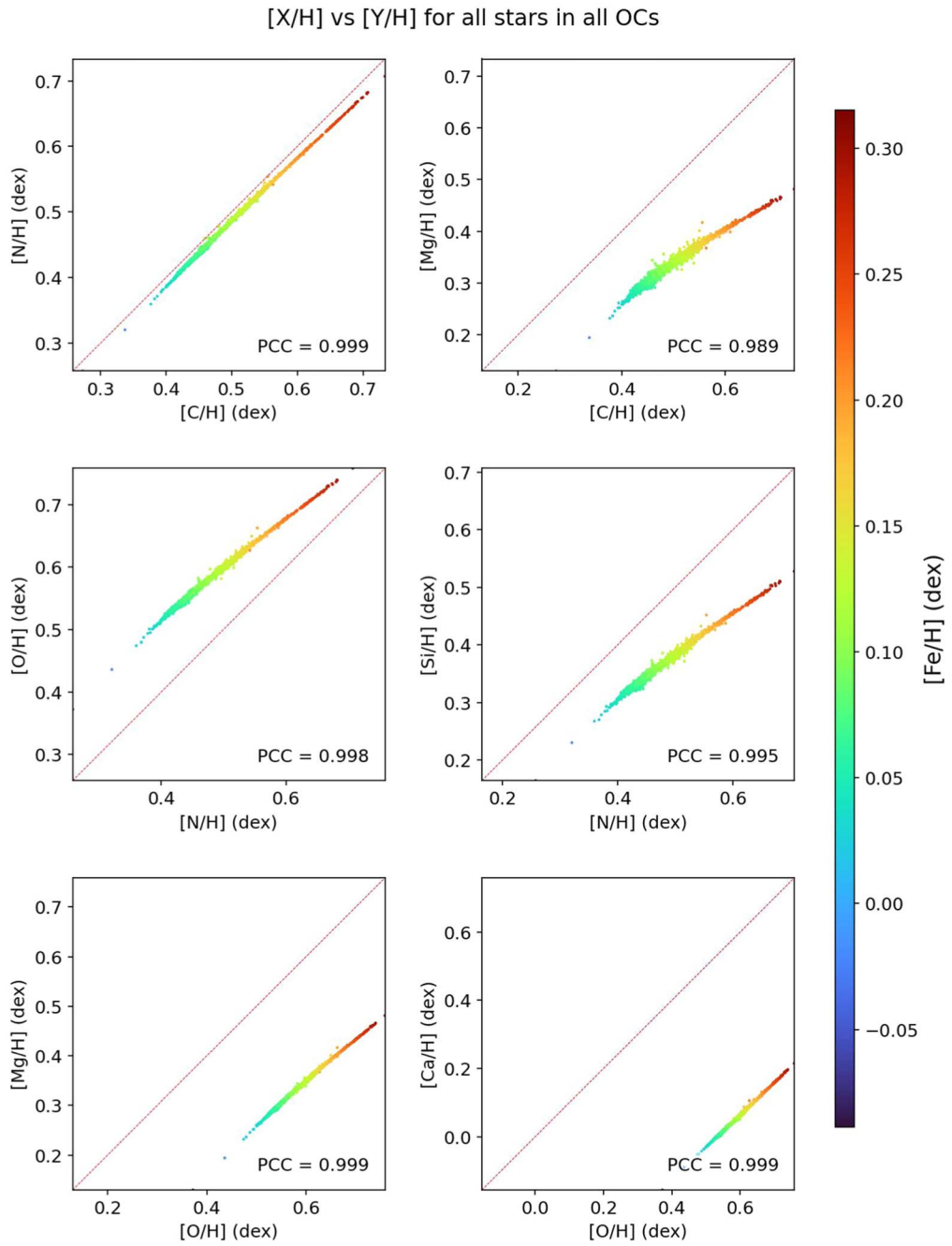

Figure 12. Correlation between [X/H] vs. [Y/H] for C, N, O, Mg, and Ca, color-coded by [Fe/H] (shown in the color bar). Elements are highly correlated with one another, as indicated by the Pearson correlation coefficient ($\geqslant$0.989 for all elements considered). The dotted diagonal red line is the one-to-one line. Note that the slope of the [X/H] vs. [Y/H] is similar to the one-to-one line and there is limited scatter in [X/H] at any value of [Y/H]. This indicates that all the elements considered carry similar discriminating power.

Appendix B
Chemical Difference Metric Using Only Three Elements

Here, we consider how the chemical difference metric is impacted by the number of elements used or the types of elements chosen to compute the metric. We have shown our analysis using all elements (C, N, O, Mg, Si, S, Ca, and Ne) in Section 3.3. However, as shown by Figures 10, 11, and 12, the elements are highly correlated with each other in our simulations. Thus, here we discuss the distribution of chemical differences in two other cases: calculated using only the elements Mg, C, and N, as well as calculated using only the elements Mg, C, and O. We have selected these elements as the maximally uncorrelated elements presented in Appendix A. We compare these results alongside the case where we use all elements to compute the metric.

We show histograms of the chemical difference metric in Figure 13 for pairs of stars within OCs, indicated by blue histograms (intracluster pairs), and between OCs, indicated by the dashed red histograms (intercluster pairs), for each simulation: m12i (top panel), m12f (middle panel), and m12m (bottom panel). These chemical difference distributions are computed similarly to those presented in Figure 8, which was generated by selecting OCs that fall within $\pm 1\sigma$ of the overall mean distribution of [Fe/H] in each simulation, then

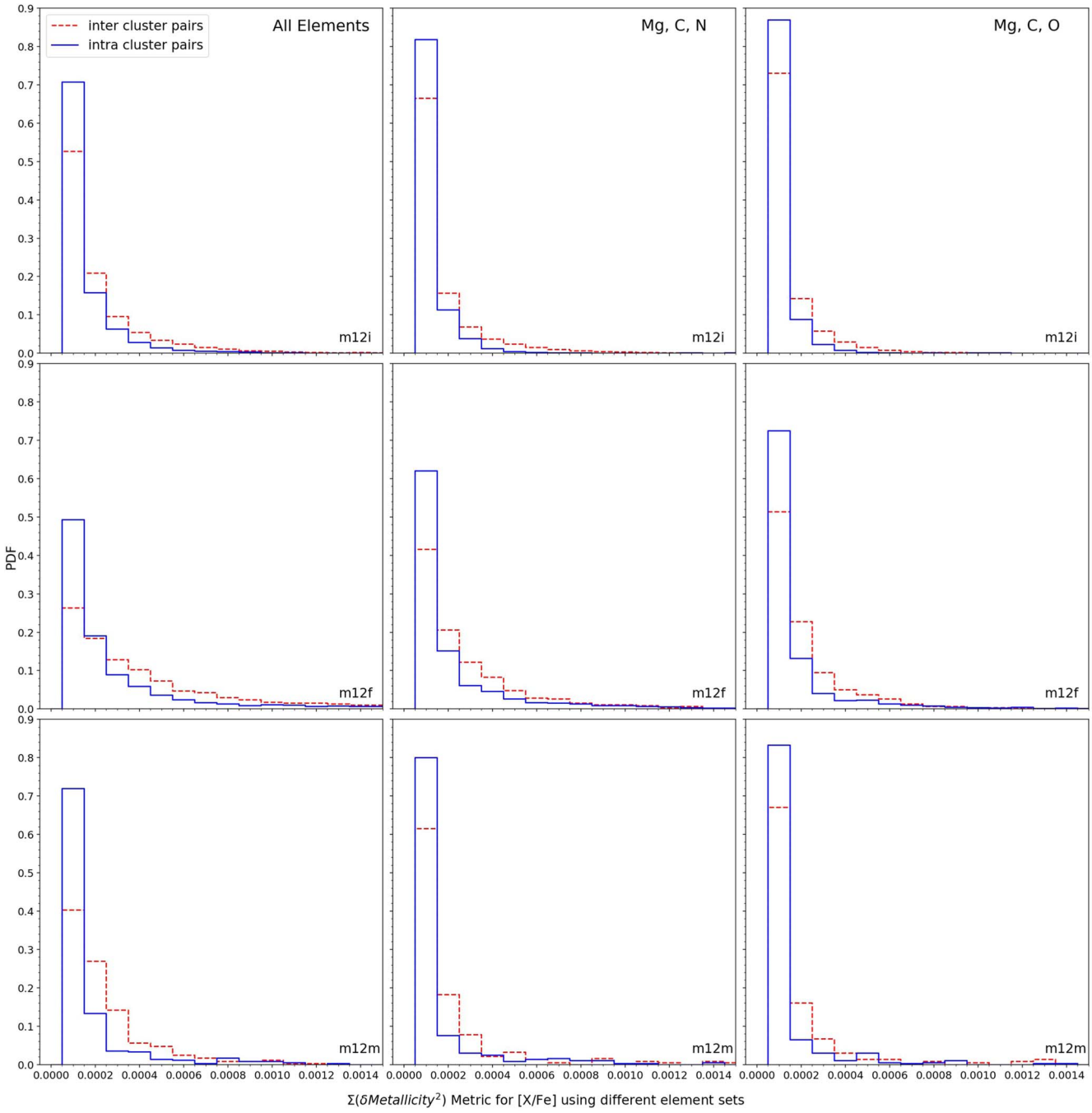

Figure 13. Histograms of the chemical difference metric for pairs of stars within OCs, indicated by blue histograms (intracluster pairs), and between OCs, indicated by the dashed red histograms (intercluster pairs), for each simulation: m12i (top), m12f (middle), and m12m (bottom). The chemical difference distributions shown in each column are computed using the method presented Section 2.3, but leverage different elements to compute the chemical difference metric. The left panel is a reproduction of Figure 8, which was generated using all relevant elements in our simulations (C, N, O, Mg, Si, S, Ca, and Ne). The middle column shows the chemical difference distribution calculated using three elements (Mg, C, and N). The right column shows the chemical difference distribution calculated using a different subset of three elements (Mg, C, and O). We note that decreasing the number of elements increases both the number of intracluster pairs and intercluster pairs that populate the smallest-chemical-difference-metric bin. There is slightly more contamination from intercluster pairs when fewer elements are used to calculate the chemical difference metric. While broadly all of the distributions presented here look very similar, the maximum differences between the inter- and intracluster global distributions are found when all elements are used to calculate the metric.

applying the formula presented in Section 2.3 to relevant pairs drawn for the selected OCs. The first column on the left shows the distribution of the chemical difference metric when all elements are used. The middle column and right column are for the cases where only three elements are used: Mg, C, and N and Mg, C, and O. We note that decreasing the number of elements increases both the number of intracluster pairs and intercluster pairs that populate the smallest-chemical-difference-metric bin. There is slightly more contamination from intercluster pairs when fewer elements are used to calculate the chemical difference metric. While broadly all of the distributions presented here look very similar, the maximum differences between the inter- and intracluster global distributions are found when all elements are used to calculate the metric.

Table 1
Chemical-difference-metric Analysis Using All Elements vs. a Subset of Elements (Mg, C, and N and Mg, C, and O)

Simulation	Selection of Chemical Difference Metric	Intracluster All Elements (%)	Intracluster Mg, C, and N (%)	Intracluster Mg, C, and O (%)	Intercluster All Elements (%)	Intercluster Mg, C, and N (%)	Intercluster Mg, C, and O (%)
m12i	< Intra Mean	76.23	74.94	75.94	59.61	58.82	59.07
	< Intra Median	50.00	50.00	50.00	31.60	33.23	32.16
	< Intra+Inter Mean	81.68	81.64	82.21	66.63	66.39	66.85
	< Intra+Inter Median	59.54	58.55	59.24	40.46	41.45	40.76
	< 0.00010 (Fixed)	70.47	81.66	86.79	52.36	66.41	72.92
	< 0.00020 (Fixed)	86.18	92.95	95.57	73.20	82.13	87.21
m12f	< Intra Mean	83.20	79.47	84.33	70.39	65.76	71.57
	< Intra Median	49.97	49.97	49.97	27.77	28.66	28.25
	< Intra+Inter Mean	86.94	83.98	88.07	79.11	74.24	80.71
	< Intra+Inter Median	62.08	60.42	61.36	37.92	39.58	38.64
	< 0.00010 (Fixed)	47.54	61.66	72.28	25.34	40.77	50.92
	< 0.00020 (Fixed)	65.93	76.74	85.40	43.03	61.01	73.47
m12m	< Intra Mean	83.16	81.55	82.62	62.03	67.65	66.04
	< Intra Median	50.00	50.00	50.00	17.65	19.79	16.58
	< Intra+Inter Mean	85.83	83.16	84.76	72.19	72.19	72.46
	< Intra+Inter Median	67.11	63.64	65.24	32.89	36.36	34.76
	< 0.00010 (Fixed)	70.59	79.14	82.62	38.77	61.50	67.11
	< 0.00020 (Fixed)	83.69	86.63	89.04	64.71	79.68	83.16

Table 2
Purity Percentage of Intracluster Pairs

Simulation	Statistic	All Elements Purity (%)	Mg, C, and N Purity (%)	Mg, C, and O Purity (%)
m12i	< Intra Mean	56.13	56.05	56.23
	< Intra Median	61.29	60.09	60.91
	< Intra+Inter Mean	55.09	55.12	55.16
	< Intra+Inter Median	59.54	58.55	59.24
	< 0.00010 (Fixed)	57.36	55.14	54.36
	< 0.00020 (Fixed)	54.08	53.12	53.21
m12f	< Intra Mean	54.14	54.56	54.94
	< Intra Median	64.69	64.52	64.15
	< Intra+Inter Mean	54.13	53.34	53.57
	< Intra+Inter Median	62.08	60.42	61.36
	< 0.00010 (Fixed)	47.54	61.66	72.28
	< 0.00020 (Fixed)	65.93	76.74	85.40
m12m	< Intra Mean	59.64	60.04	60.31
	< Intra Median	55.23	54.91	54.84
	< Intra+Inter Mean	56.03	55.47	55.23
	< Intra+Inter Median	67.11	63.64	65.24
	< 0.00010 (Fixed)	57.34	55.65	55.12
	< 0.00020 (Fixed)	54.39	53.76	53.92

In Table 1, we show the percentage of pairs that have a chemical difference metric smaller than the mean, median, or a fixed value when different elements are used to calculate the chemical difference metric. While there is a slight increase or decrease in the percentage of pairs falling below a chosen limit, the chemical difference overlap—that is, the relative amount of intercluster contamination—does not change significantly; there is always significant contamination at small values of the chemical difference metric. In Table 2, we present the purity (the percentage of intracluster pairs relative to the total pairs) of each sample for each metric. Each population is significantly impure—the contamination rate is typically 40%–50%. We conclude that no matter what metric we use or subset of elements we consider, we always have significant contamination of intercluster pairs.

We note that the mean and median of the chemical difference distribution decreases very slightly when few elements are considered, and it decreases significantly when the mean and median of the metrics are computed when combining the intra- and intercluster distributions into one global distribution; we present this for completeness in Table 3.

Table 3
Chemical-difference-metric Statistics

Simulation	Statistic	All Elements	Mg, C, and N	Mg, C, and O
m12i	Intra Mean	0.00007	0.00007	0.00006
	Intra Median	0.00003	0.00003	0.00002
	Intra +Inter Mean	0.00016	0.00010	0.00008
	Intra+Inter Median	0.00007	0.00004	0.00003
m12f	Intra Mean	0.00024	0.00024	0.00019
	Intra Median	0.00006	0.00006	0.00004
	Intra +Inter Mean	0.00065	0.00032	0.00026
	Intra+Inter Median	0.00017	0.00009	0.00006
m12m	Intra Mean	0.00012	0.00012	0.00010
	Intra Median	0.00002	0.00002	0.00002
	Intra +Inter Mean	0.00024	0.00014	0.00012
	Intra+Inter Median	0.00009	0.00005	0.00004

ORCID iDs

Binod Bhattarai https://orcid.org/0000-0002-7707-1996
Sarah R. Loebman https://orcid.org/0000-0003-3217-5967
Melissa K. Ness https://orcid.org/0000-0001-5082-6693
Andrew Wetzel https://orcid.org/0000-0003-0603-8942
Emily C. Cunningham https://orcid.org/0000-0002-6993-0826
Hanna Parul https://orcid.org/0009-0007-3431-4269
Alessa Ibrahim Wiggins https://orcid.org/0009-0008-0081-764X

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Anders, F., Chiappini, C., Santiago, B. X., et al. 2014, A&A, 564, A115
Armillotta, L., Krumholz, M. R., & Fujimoto, Y. 2018, MNRAS, 481, 5000
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Bellardini, M. A., Wetzel, A., Loebman, S. R., & Bailin, J. 2022, MNRAS, 514, 4270
Bellardini, M. A., Wetzel, A., Loebman, S. R., et al. 2021, MNRAS, 505, 4586
Bertelli Motta, C., Pasquali, A., Richer, J., et al. 2018, MNRAS, 478, 425
Bertran de Lis, S., Allende Prieto, C., Majewski, S. R., et al. 2016, A&A, 590, A74
Blanco-Cuaresma, S., Soubiran, C., Heiter, U., et al. 2015, A&A, 577, A47
Bland-Hawthorn, J., Krumholz, M. R., & Freeman, K. 2010, ApJ, 713, 166
Boeche, C., Siebert, A., Piffl, T., et al. 2013, A&A, 559, A59
Bovy, J. 2016, ApJ, 817, 49
Bragaglia, A., Sneden, C., Carretta, E., et al. 2014, ApJ, 796, 68
Bragaglia, A., & Tosi, M. 2006, AJ, 131, 1544
Buder, S., Lind, K., Ness, M. K., et al. 2019, A&A, 624, A19
Cantat-Gaudin, T. 2022, Univ, 8, 111
Cantat-Gaudin, T., & Casamiquela, L. 2024, NewAR, 99, 101696
Carrillo, A., Ness, M. K., Hawkins, K., et al. 2023, ApJ, 942, 35
Casamiquela, L., Tarricq, Y., Soubiran, C., et al. 2020, A&A, 635, A8
Cunha, K., Smith, V. V., Johnson, J. A., et al. 2015, ApJL, 798, L41
Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, Proc. SPIE, 8446, 84460P
De Silva, G. M., Freeman, K. C., Asplund, M., et al. 2007, AJ, 133, 1161
De Silva, G. M., Freeman, K. C., & Bland-Hawthorn, J. 2009, in IAU Symp. 254, The Galaxy Disk in Cosmological Context, ed. J. Andersen, B. m Nordströara, & J. Bland-Hawthorn (Cambridge: Cambridge Univ. Press), 133
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
De Silva, G. M., Sneden, C., Paulson, D. B., et al. 2006, AJ, 131, 455
Dobbs, C. L., Adamo, A., Few, C. G., et al. 2017, MNRAS, 464, 3580
Donor, J., Frinchaboy, P. M., Cunha, K., et al. 2020, AJ, 159, 199
Dotter, A., Conroy, C., Cargile, P., & Asplund, M. 2017, ApJ, 840, 99
Elia, D., Molinari, S., Schisano, E., et al. 2022, ApJ, 941, 162
Escala, I., Wetzel, A., Kirby, E. N., et al. 2018, MNRAS, 474, 2194
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Feng, Y., & Krumholz, M. R. 2014, Natur, 513, 523
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Frinchaboy, P. M., & Majewski, S. R. 2008, AJ, 136, 118
Geller, M. J., & Huchra, J. P. 1983, ApJS, 52, 61
Graf, R. L., Wetzel, A., Bellardini, M. A., & Bailin, J. 2024, arXiv:2402.15614
Griffith, E. J., Hogg, D. W., Dalcanton, J. J., et al. 2024, AJ, 167, 98
Grudić, M. Y., Hafen, Z., Rodriguez, C. L., et al. 2023, MNRAS, 519, 1366
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hogg, D. W., Casey, A. R., Ness, M., et al. 2016, ApJ, 833, 262
Hopkins, P. F. 2015, MNRAS, 450, 53
Hopkins, P. F., Chan, T. K., Garrison-Kimmel, S., et al. 2020, MNRAS, 492, 3465
Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800
Hunt, E. L., & Reffert, S. 2024, A&A, 686, A42
Hunter, J. D. 2007, CSE, 9, 90
Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, ApJS, 125, 439
Izzard, R. G., Tout, C. A., Karakas, A. I., & Pols, O. R. 2004, MNRAS, 350, 407
Jacobson, H. R., Pilachowski, C. A., & Friel, E. D. 2011, AJ, 142, 59
Janes, K. A. 1979, ApJS, 39, 135
Kamdar, H., Conroy, C., Ting, Y.-S., et al. 2019, ApJ, 884, 173
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
Kreckel, K., Ho, I.-T., Blanc, G. A., et al. 2020, MNRAS, 499, 193
Kroupa, P. 2001, MNRAS, 322, 231
Kruijssen, J. M. D., Pelupessy, F. I., Lamers, H. J. G. L. M., Portegies Zwart, S. F., & Icke, V. 2011, MNRAS, 414, 1339
Krumholz, M. R., & Gnedin, N. Y. 2011, ApJ, 729, 36
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lambert, D. L., & Reddy, A. B. S. 2016, ApJ, 831, 202
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Li, Z., Wisnioski, E., Mendel, J. T., et al. 2023, MNRAS, 518, 286
Liu, F., Asplund, M., Yong, D., et al. 2019, A&A, 627, A117
Liu, F., Yong, D., Asplund, M., Ramírez, I., & Meléndez, J. 2016, MNRAS, 457, 3934
Magrini, L., Randich, S., Romano, D., et al. 2014, A&A, 563, A44
Magrini, L., Viscasillas Vázquez, C., Spina, L., et al. 2023, A&A, 669, A119
Majewski, S. R. & APOGEE Team 2016, AN, 337, 863
Majewski, S. R., Nidever, D. L., Smith, V. V., et al. 2012, ApJL, 747, L37
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Manea, C., Hawkins, K., & Maas, Z. G. 2022, MNRAS, 511, 2829
Manea, C., Hawkins, K., Ness, M. K., et al. 2024, ApJ, 972, 69
Mannucci, F., Della Valle, M., & Panagia, N. 2006, MNRAS, 370, 773
Marigo, P. 2001, A&A, 370, 194
Martell, S. L., Shetrone, M. D., Lucatello, S., et al. 2016, ApJ, 825, 146
Masoumi, D. 2024, PhD thesis, Univ. of California, Merced
Mathieu, R. D. 2000, in ASP Conf. Ser. 198, Stellar Clusters and Associations: Convection, Rotation, and Dynamos, ed. R. Pallavicini, G. Micela, & S. Sciortino (San Francisco, CA: ASP), 517
McKee, C. F., & Tan, J. C. 2002, Natur, 416, 59
Meyer, M. R., Adams, F. C., Hillenbrand, L. A., Carpenter, J. M., & Larson, R. B. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson, AZ: Univ. Arizona Press), 121
Minchev, I., Anders, F., Recio-Blanco, A., et al. 2018, MNRAS, 481, 1645
Mitschang, A. W., De Silva, G., Sharma, S., & Zucker, D. B. 2013, MNRAS, 428, 2321
Mollá, M., Wekesa, S., Cavichia, O., et al. 2019, MNRAS, 490, 665
Myers, N., Donor, J., Spoo, T., et al. 2022, AJ, 164, 85
Ness, M., Hogg, D. W., Rix, H. W., Ho, A. Y. Q., & Zasowski, G. 2015, ApJ, 808, 16
Ness, M., Rix, H. W., Hogg, D. W., et al. 2018, ApJ, 853, 198
Ness, M. K., Johnston, K. V., Blancato, K., et al. 2019, ApJ, 883, 177
Ness, M. K., Wheeler, A. J., McKinnon, K., et al. 2022, ApJ, 926, 144
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, NuPhA, 777, 424
Pancino, E., Carrera, R., Rossetti, E., & Gallart, C. 2010, A&A, 511, A56
Perez, F., & Granger, B. E. 2007, CSE, 9, 21
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Poovelil, V. J., Zasowski, G., Hasselquist, S., et al. 2020, ApJ, 903, 55
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Price-Jones, N., & Bovy, J. 2019, MNRAS, 487, 871
Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2016, A&A, 587, A70
Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2020, ApJS, 246, 6
Sestito, P., Bragaglia, A., Randich, S., et al. 2008, A&A, 488, 943
Sharma, S., Hayden, M. R., Bland-Hawthorn, J., et al. 2022, MNRAS, 510, 734
Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23
Sinha, A., Zasowski, G., Frinchaboy, P., et al. 2024, ApJ, 975, 89
Spina, L., Magrini, L., & Cunha, K. 2022a, Univ, 8, 87
Spina, L., Magrini, L., Sacco, G. G., et al. 2022b, A&A, 668, A16
Spina, L., Sharma, P., Meléndez, J., et al. 2021, NatAs, 5, 1163
Su, K.-Y., Hopkins, P. F., Hayward, C. C., et al. 2017, MNRAS, 471, 144
Ting, Y.-S., Conroy, C., & Goodman, A. 2015, ApJ, 807, 104
Ting, Y.-S., De Silva, G. M., Freeman, K. C., & Parker, S.-J. 2012, MNRAS, 427, 882
van den Hoek, L. B., & Groenewegen, M. A. T. 1997, A&AS, 123, 305
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wetzel, A., & Garrison-Kimmel, S., 2020a HaloAnalysis: Read and Analyze Halo Catalogs and Merger Trees, Astrophysics Source Code Library, ascl:2002.014
Wetzel, A., & Garrison-Kimmel, S., 2020b GizmoAnalysis: Read and Analyze Gizmo Simulations, Astrophysics Source Code Library, ascl:2002.015
Wetzel, A., Hayward, C. C., Sanderson, R. E., et al. 2023, ApJS, 265, 44
Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, ApJL, 827, L23
Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009, MNRAS, 399, 574