Chronologically dating the early assembly of the Milky Way

Josefina Montalbán1,2, J. Ted Mackereth1,2,3, Andrea Miglio5,6, Fiorenzo Vincenzo1,7,8, Cristina Chiappini9, Gael Buldgen10, Benoît Mosser11, Arlette Noels12, Richard Scuflaire12, Mathieu Vrard8,13, Emma Willett1,14, Guy R. Davies1,14, Oliver J. Hall1,14, Martin Bo Nielsen1,14,15, Saniya Khan1,14, Ben M. Rendle1,14, Walter E. van Rossem1,14, Jason W. Ferguson16 and William J. Chaplin1,14

The standard cosmological model predicts that galaxies are built through hierarchical assembly on cosmological timescales1–3. The Milky Way, like other disk galaxies, underwent violent mergers and accretion of small satellite galaxies in its early history. Owing to Gaia Data Release 2 (DR2) and spectroscopic surveys4, the stellar remnants of such mergers have been identified5–7. The chronological dating of such events is crucial to uncover the formation and evolution of the Galaxy at high redshift, but it has so far been challenging due to difficulties in obtaining precise ages for these oldest stars. Here we combine asteroseismology—the study of stellar oscillations—with kinematics and chemical abundances to estimate precise stellar ages (~11%) for a sample of stars observed by the Kepler space mission8. Crucially, this sample includes not only the oldest stars that were formed inside the Galaxy but also stars formed externally and subsequently accreted onto the Milky Way. Leveraging this resolution in age, we provide compelling evidence in favour of models in which the Galaxy had already formed a substantial population of its stars (which now reside mainly in its thick disk) before the infall of the satellite galaxy Gaia-Enceladus/Sausage8,9 around 10 billion years ago.

Recent results based on the European Space Agency (ESA) Gaia mission1 have revealed that the stellar content of the inner halo of the Milky Way (MW) is dominated by debris from some seemingly massive dwarf galaxies, such as the Gaia-Enceladus/Sausage (hereafter GES). The merging event with the GES is now purported to be one of the most important in the Galaxy’s history, shaping how we observe it today9–11. To constrain the effect of such mergers on the MW and other similar galaxies, it is crucial to understand their state both before and following the merger. This requires mapping the temporal sequence of these events with the highest precision possible (~10% to follow the first four billion years after the Big Bang11). Several recent studies have estimated the characteristics and timing of this merging event12,13,14,15, while others (before16,17 and after18,19 Gaia Data Release 2 (DR2)) have aimed to age-date the accreted and in situ stellar populations of the MW halo (see Helmi20 for an extensive review). Although using different kinds of targets and methods, these age-dating techniques are however quite limited in precision and accuracy, as they are based on stellar surface properties and on predictions from stellar evolution models. The latter are known to be affected by, for example, uncertainty in the physics and degeneracy between parameters, which makes it difficult to obtain stellar ages with the required precision and accuracy.

Red giant (RG) stars, being long-lived and intrinsically bright, are excellent candidates to map ages in different regions of the MW12,14,15. However, aging RGs in colour–magnitude space using their surface properties gives yet uncertain results since their colours and luminosities are similar, whatever their mass and age. Fortunately, asteroseismology, which probes the internal structure of stars, provides us with means to reach a precision of 10–20% on age dating individual RG stars12,13,14.

Among the roughly 15,000 oscillating K- and G-type RG stars detected in the field observed by the NASA Kepler space telescope1, a small fraction lie in the low-metallicity regime characteristic of the inner MW halo and high-[α/Fe] disk component ([Fe/H] < −0.5). Of these, some 400 stars have precisely measured element abundances, atmospheric parameters and radial velocities from Data Release 14 of the Apache Point Observatory Galactic Evolution Experiment spectroscopic survey (APOGEE DR14), as well as detailed proper motions from Gaia DR21.

The Kepler data provide oscillation frequency spectra of exquisite quality and resolution, allowing precise estimates to be made of the frequencies of modes of different angular degree (radial ℓ = 0, dipolar ℓ = 1 and quadrupolar ℓ = 2), and hence of fundamental parameters and evolutionary state of the stars. We first use this seismic information to remove from the sample those stars that are in the red clump (RC) phase (that is, low-mass, helium-core-burning stars that probably underwent mass loss earlier in their evolution) and in subsequent phases. Removing these contaminants leaves a
We estimate stellar properties (Supplementary Table 1) using the individual frequencies of radial modes and atmospheric parameters from spectroscopy as observational inputs in AIMS\(^\text{23}\), a Bayesian parameter estimation code, which provides best-fitting stellar properties and full posterior probability distributions by comparing with theoretical stellar models and adiabatic frequencies (Methods and Extended Data Figs. 2 and 3). The precision on age we achieve, of \(11\%\) on average, affords us the ability to unpick the chronological sequence of events some \(\sim 12\) Gyr ago, as we show below.

The robustness of our estimated stellar age distributions has been checked performing different tests (see Methods and Extended Data Figs. 4 and 5 for detailed description). Moreover, as shown in Fig. 1, despite having only fitted modes of degree \(\ell = 0\), the theoretical spectra predicted by the best-fitting model parameters reproduce well also the non-radial modes (\(\ell = 1, 2\)) of the observed spectra, which reinforces the confidence on the quality of the derived stellar parameters.

Figure 2 summarizes the chronological, chemical and kinematic properties of the final sample of 95 RGB stars for which we could robustly determine ages. In Fig. 2a, we show its \([\text{Fe/H]}-\text{[Mg/Fe]}\) distribution (coloured by age), together with that of APOGEE DR14 sample (grey points). The grey points clearly show, in addition to two overdensities at higher \([\text{Fe/H]}\) corresponding to the low- and high-\([\alpha/\text{Fe]}\) Galactic Disk populations and metal-rich in situ halo\(^\text{35}\), a scattered population at \([\text{Fe/H}] < -0.7\) and intermediate \([\text{Mg/Fe}] (-0.1\) to 0.2), where the recently characterized GES population lies\(^\text{5,14,26}\). Our final RGB sample (coloured circles) contains members of each of these populations. We also include, for reference, the location of \(\nu\) Indi, a bright subgiant recently dated using seismology, and classified as belonging to the ‘heated’ thick disk\(^\text{27}\).

Recent studies of the Galactic halo and local group dwarfs\(^\text{35}\) suggest that stars in the low-\([\text{Mg/Fe]}\) sequence at \([\text{Fe/H}] \lesssim -0.7\) have probably been accreted to the Galaxy\(^\text{16,29-31}\). The \([\alpha/\text{Fe]}\) ratios in local dwarfs indicate a higher pollution from type Ia supernovae relative to core-collapse supernovae, probably due to inefficient star-formation activity and strong outflows\(^\text{25,26}\). As a consequence, their \([\alpha/\text{Fe]}\) ratios are lower than in situ halo stars, where element abundances are more affected by nucleosynthetic products from core collapse as opposed to type Ia supernovae.

On the basis of the above studies, we classify the asteroseismic RGB sample by making a cut in \([\text{Fe/H]}-\text{[Mg/Fe]}\) space along the line \([\text{Mg/Fe}] = -0.2\) [\text{Fe/H}] + 0.05. Stars below this line are likely to have been formed in dwarf satellites and then accreted, and those above should be born, in majority, in situ. It is conceivable that the in situ and ex situ populations defined in this way will have some contamination from the other group. To mitigate this, we further divide stars below the line into high and low orbital eccentricity groups (calculated as described in Methods). Stars on more radial orbits (eccentricities \(e > 0.7\), open points) are those most likely to have been accreted from the GES progenitor\(^\text{14}\).

Figure 2b, which shows the nickel abundance relative to iron \([\text{Ni/Fe]}\) and the sum of carbon and nitrogen abundance relative to oxygen \([(C+N)/O]\), supports that the applied cuts efficiently isolate different stellar populations\(^\text{25,26}\). The APOGEE DR14 sample below the \([\text{Fe/H]}-\text{[Mg/Fe]}\) line (large grey points) is depleted in both \([\text{Ni/Fe]}\) and \([C+N]/O]\), consistent with local dwarf satellite galaxies, which contain stars with \([\text{Ni/Fe]}\) ratios lower than the MW\(^\text{30}\). The stars of our high-\(e\), low-\([\text{Mg/Fe]}\) sample (hereafter, group A) lie at the lowest values of \([\text{Ni/Fe]}\), and are clearly separated from the other groups, reinforcing our contention that the low-[\text{Mg/Fe}], \(e > 0.7\) group is likely to be formed ex situ. The group made of low-\(e\), low-[\text{Mg/Fe}] stars (hereafter, group B) is probably a mixture of stars of different origin\(^\text{14,35}\): the tail at low eccentricity of GES stars, the low-metallicity end of the thin disk (for instance, the two stars with \([\text{Fe/H}] \gtrsim -0.7\)) or remnants of less massive accretion events, as could be the case for the stars with \([\text{Ni/Fe]}\) and \([C+N]/O]\) patterns similar to those of high-[\text{Mg/Fe}] stars, which could be indicative of a different star-formation history in the galaxy of origin\(^\text{35}\).

The composition of this group is the most sensitive to the details of the classification criterion adopted; however, this does not affect the robustness of our main conclusions about the chronological order of the GES and high-[\text{Mg/Fe}] (hereafter, group C) populations (Methods and Extended Data Fig. 6).

The blue (BS) and red (RS) sequences revealed by Gaia DR2\(^\text{35}\) in the colour–magnitude diagram (CMD) of the kinematically defined halo have been associated to a population of extragalactic origin, and to the in situ halo and/or heated thick disk, respectively\(^\text{37}\). As shown in Fig. 3, most of the stars of our group C, which we classified as an in situ high-[\text{Fe/H}] disk/halo population, naturally occupy the RS, while likely GES stars lie in the BS. The high precision ages afforded by asteroseismology allow us to order chronologically the formation of the accreted population with respect to the high-[\text{Fe/H}] in situ one.

Figure 4 shows our main finding: the distribution in age and orbit eccentricity (coloured by \([\text{Fe/H]}\)) of stars in our sample. The top panel shows the marginalized posterior distributions in age for our three groups of stars: group A, [Mg/Fe] below the cut and \(e > 0.7\).
**Fig. 2** | Chronological, chemical and kinematic properties of the seismic RGB sample. 

**a.** The [Mg/Fe]-[Fe/H] plane for our sample (points coloured by age), compared with the rest of APOGEE DR14 at [Fe/H] < −0.5 (small grey points). The dashed line ([Mg/Fe] = −0.2[Fe/H] + 0.05) demonstrates the simple division we make between likely in situ (above) and ex situ (below) stars. Likely ex situ stars with $e > 0.7$ are shown as open points. The diamond represents $v_{\text{ind}}$ values\(^{22}\)\(^{22}\). 

**b.** The ([C + N]/O)-[Ni/Fe] distribution of the three groups defined in a. The underlying grey points represent the entire APOGEE DR14 sample with [Fe/H] < −0.5, with those which lie below the [Fe/H]-[Mg/Fe] division shown as larger points. The low-[Mg/Fe], $e > 0.7$ stars have atypical element abundances relative to the other groups, exhibiting very low [Ni/Fe] and a small depletion in carbon and nitrogen relative to oxygen.

---

As in other recent papers\(^{19,30}\), we find in our sample a fraction of ‘apparently young’ stars, despite chemical markers indicative of old ages (elevated $\alpha$-element abundances and a high C/N ratio). As asteroseismology assigns a high mass to these targets, they have previously been identified as ‘overmassive’ $\alpha$-rich stars (likely product of mass transfer\(^{30}\)).

We fit a hierarchical model to the stellar ages in each group, assessing the mean age and the intrinsic age spread of each population. We assume that the true age of each star in each group is contaminated by a wider normal distribution by some fraction $\epsilon$, which are probably ex situ in origin have a similar (but slightly younger) mean age compared with the majority of the stars in the in situ high-[Mg/Fe] population. This suggests that these ex situ high-$e$ stars were probably formed at roughly the same epoch as, or even after, the high-[Mg/Fe] population. The contamination by the overmassive (and therefore young in appearance) stars is of the order 10%, with a consistent age and spread among each population.

---

**Table 1** | Properties of the inferred age distribution of the three populations

| Group | $\mu$ (Gyr) | $\tau$ (Gyr) | $\mu_{\epsilon}$ (Gyr) | $\tau_{\epsilon}$ (Gyr) | $\epsilon$ |
|-------|-------------|-------------|-----------------------|-----------------------|-----------|
| A: low [Mg/Fe], $e > 0.7$ | $9.7 \pm 0.6$ | $0.8^{+0.9}_{-0.4}$ | $4.5 \pm 2.0$ | $2.9^{+5.7}_{-2.0}$ | $0.15^{+0.12}_{-0.08}$ |
| B: low [Mg/Fe], $e < 0.7$ | $8.2 \pm 0.8$ | $0.8^{+1.0}_{-0.5}$ | $4.9 \pm 2.0$ | $2.8^{+6.1}_{-1.8}$ | $0.06^{+0.07}_{-0.03}$ |
| C: high [Mg/Fe] | $10.4 \pm 0.3$ | $0.5^{+0.2}_{-0.3}$ | $4.2 \pm 0.8$ | $2.1^{+4.2}_{-1.4}$ | $0.16^{+0.05}_{-0.04}$ |

Median and $1\sigma$ error interval of: mean age $\mu$, intrinsic age spread $\tau$ of the main population and contaminant ($\mu_{\epsilon}, \tau_{\epsilon}$) population (that of ‘overmassive’ stars), and the contaminant fraction $\epsilon$ for the three populations of stars defined in the sample of metal-poor Kepler giants. The high- and low-$\alpha$, low-[Mg/Fe] stars have significantly different age distributions. The high-$e$, low-[Mg/Fe] stars, which are probably ex situ in origin have a similar (but slightly younger) mean age compared with the majority of the stars in the in situ high-[Mg/Fe] population. This suggests that these ex situ high-$e$ stars were probably formed at roughly the same epoch as, or even after, the high-[Mg/Fe] population. The contamination by the overmassive (and therefore young in appearance) stars is of the order 10%, with a consistent age and spread among each population.
and removing them from the sample does not change the age distribution. This difference is visible in the posterior age distribution for these stars shown in Fig. 4 (yellow histogram). (2) Population A stars (which we associate with the GES progenitor) have a mean age and spread consistent with those of population C. This suggests that these stars, which are likely to have been born ex situ, were formed contemporaneously to, if not slightly after, the high-[α/Fe] population C that was formed in the MW starting ~10 to 11.5 Gyr ago, as shown by this study.

The precise ages inferred here for individual objects in the BS and RS of Fig. 3 provide crucial constraints to MW formation models and to the more general debate on the dominant drivers of star formation. The models and to the more general debate on the dominant drivers of star formation and RS of Fig. 3 provide crucial constraints to MW formation models and to the more general debate on the dominant drivers of star formation and associated accretion. Asteroseismology provides, in addition to atmospheric parameters, information about the internal stellar structure, angular momentum content and the characteristic rotation speed of the stars.

**Methods**

**APOGEE, Gaia and Kepler data.** We select targets with SDSS-IV/APOGEE spectra and NASA Kepler light curves by cross-matching the APOGEE DR14 catalogue with the Kepler Input Catalog (KIC). We then cross-match again with the Gaia DR2 catalogue, which provides parallax, position and proper motion measurements for the relevant stars. APOGEE provides, in addition to atmospheric parameters (effective temperature and detailed abundances for ~230,000 stars), highly precise (~1%) radial velocities for all targets.

**Distances.** The distance estimates using parallaxes from Gaia DR2 for our sample have a mean relative error of 15% (median 11%), and for 22% of the sample that value is larger than 20%. We also take distance estimates from the astroNN catalogue, which are based on neural network models of the APOGEE spectra, trained on the Gaia DR2 parallaxes. astroNN distances, which have relative uncertainties of roughly 10%, and provide a more robust measure of the stellar distances than the parallax information for these more distant stars, have been used in the determination of the orbital parameters.

**Orbital parameters.** Orbital parameters are estimated for the sample in question using the Stäckel approximation-based fast orbit estimation method implemented in galpy. We take 100 samples of the covariance matrices for each star, formed from the observed right ascension, declination, proper motion in right ascension and declination, distance and radial velocity, and their uncertainties and correlation coefficients (in this case, the distance and radial velocity are measured independently, so their uncertainties are uncorrelated). We then estimate the orbital parameters for each of these samples assuming the simple MWPotential2014 potential, which is adequate in this case, since the majority of these stars have halo-like kinematics and are not likely to be affected by non-axisymmetries in the disk and bulge. We assume the position of the Sun to be \( R_{\odot} = 8.125 \) kpc (Galactocentric distance) and \( z_{\odot} = 0.02 \) kpc (height above the Galactic mid-plane), and its velocity to be \( v_{\odot} = [ U, V, W ] = [ -11.1, 245.6, 7.25 ] \) km s\(^{-1}\) (in the left-handed cartesian Galactic coordinate system), based on the Sagittarius A* proper motion and the solar motion. We estimate pericentre and apocentre radii, orbital eccentricity and the maximum vertical excursion, their uncertainties and correlation coefficients for each star. These orbital parameters will later allow us to verify the accreted nature (or not) of stars in our sample.

**Seismic data.** We retrieve Kepler light curves from MAST (https://archive.stsci.edu/kepler/publiclightcurves.html) and measure individual radial-mode frequencies following the approach in ref.\(^{14}\). These results were cross-matching with the radial frequency modes using the automatic pipeline PBJam.
We measured frequency of at least 3 individual radial modes in 278 targets over 400. From that sample, we remove the stars in the AGB (~50%). Their current masses (those inferred from seismology) are probably the result of some mass loss in previous evolutionary phases, and hence their age estimates would depend on the poorly known mass-loss prescription itself. This classification is based on the value of the gravity-mode period spacing crank, when available, and from visual inspection of the power spectra. This classification has been also cross-checked with results in other studies. Among the non-core-helium-burning giants, we restrict the sample to stars with $v_{\text{max}}$ larger than 15 Hz. This mitigates contamination from early asymptotic giant branch (AGB) stars and removes stars with relatively low $v_{\text{max}}$, a domain where seismic constraints are less numerous (the number of radial modes decreases with $v_{\text{max}}$) and robust.

**Final sample.** After the above refereed cuts, our final sample contains 105 stars, likely in the RGB, with at least four radial modes detected (8, 19, 47 and 31 with 4, 5, 6 and 7 modes, respectively). Their frequencies have a mean uncertainty of 0.0858% (median 0.0595%). The typical uncertainty of the effective temperature ($T_{\text{eff}}$), a key parameter for metalliclicity determination, is 1.7%. The end of the sample is [Fe/H] = -0.66, with 25% of the targets having an iron content lower than -0.85. The typical error quoted in APOGEE DR14 is ~0.008. That is substantially smaller than the typical accuracy of APOGEE DR14 chemical abundances, as assessed using different and independent spectroscopic analyses of APOGEE stars, and 10 to 20 times smaller the step used in the grid of models. Hence, by increasing $v_{\text{max}}$ for the typical error in the domain of interest, we have $v_{\text{max}}$ = 0.05 dex.

Concerning the x-elements, 50% (40%, 8% and 2%) of the targets have $\alpha$/Fe = 0.2 (0.3, 0.1 and 0.4, respectively).

**Bayesian stellar parameter inference.** The stellar parameters of each star in our sample have been estimated using a Markov chain Monte Carlo (MCMC) ensemble sampler, and selects stellar models that best fit the observation data by interpolating (evolutionary tracks and frequencies) in a pre-computed grid. As demonstrated by several studies using individual frequencies as observational constraints contributes to significantly reducing the uncertainties affecting estimated global stellar parameters with respect to the precision and accuracy resulting from scaling relations. The drawback of using individual frequencies is that theoretical values should be corrected by the surface effects. In this study, we use the frequencies of individual radial modes and their uncertainties as observational seismic constraints, and correct the theoretical frequencies using a two-terms prescription. The surface-effect corrections involve in that case two free parameters ($a_0$ and $a_1$, equation (4) in ref. 1) to be derived by the fitting procedure for each target. Other parameters to be determined are the stellar mass, the initial mass fraction of metals and the stellar age. We define specific priors for $v_{\text{max}}$, $a_0$ and $a_1$ if an initial calculation has led to unexpected surface-effect corrections. For those cases, we re-run AIMS using uniform priors for these parameters, the domains of which are estimated from the other suitable fits. As ‘classical’ constraints, we adopt the spectroscopic values of effective temperature and surface metal content from APOGEE DR14, and the average seismic index $v_{\text{max}}$ from the sample is $v_{\text{max}} = \frac{v_{\text{rms}}}{\eta_{\text{rms}}}$ and $\eta_{\text{rms}} = \frac{\sigma_{\text{rms}}}{\sigma_{\text{rms}}}$, where $\sigma_{\text{rms}}$ is the root mean square of the residuals of the fit to the mean value of the oscillation frequency for these time series, AIMS does not converge if luminosity is used as an observational constraint (regardless of the adopted Gaia zero-offset parallax).

The properties of the 95 stars of our final sample are collected in the Supplementary Table 1. Its last columns contain the values of stellar mass, radius and age of the models that best match observations, as well as their at interval values. These internal statistical errors based on the sampling of the posterior probability distributions obtained with the grid of models used. Extended Data Figs. 2 and 3 present the posterior distributions for six relevant stellar parameters for the targets KIC 4134367 and KIC 12111110, respectively. In both cases, these distributions were obtained using as constraints in AIMS: six radial modes, surface mass fraction of metals, effective temperature and $v_{\text{rms}}$. KIC 4134367 is one of the targets shown in Fig. 1, while KIC 12111110 is the object at age ~10 Gyr, and eccentricity 0.99 (Fig. 1), which shows a large and very asymmetric uncertainty. In this case, the posterior distributions are clearly bimodal. Although the best match with observation (dot-dashed vertical line) is achieved for the older group of models, a large number of AIMS models still have a high probability. An uncertainty in luminosity smaller than 10% should be needed to critically discriminate between the two solutions.

**Robustness tests against systematic uncertainties.** It is widely accepted that the effective temperature strongly depends on the age through the mean molecular mass parameter ($\mu_{\text{eff}}$) and on the adopted atmospheric boundary conditions. A systematic difference between $T_{\text{eff}}$ of the models and observations could indicate that those parameters are not the adequate ones to represent observational data, creating a tension, leading to systematic larger or smaller stellar masses, and hence affecting the estimated ages. To check for the robustness of our results for $T_{\text{eff}}$, we have $\Delta T_{\text{eff}}$ = 85 K (AIMS results—with the grid of models above described—are typically 85 K hotter than observed $T_{\text{eff}}$). Although the likelihood is generally higher for the temperature scale shifted by +85 K, the stellar parameters retrieved do not change. In fact, the fitting is dominated by the individual frequencies, with a lower impact of $T_{\text{eff}}$ directly or through $\mu_{\text{eff}}$.

In the fitting process, we do not interpolate in the parameter [Fe/H], but we select the grid with the closest value to that estimated from spectroscopy. To estimate the effect of the [Fe/H] step on the derived stellar parameters, we have run AIMS using a grid of models computed with [Fe/H] values shifted by ±0.1 dex, and compared their ages. The differences between ages obtained using the nominal [Fe/H] or that shifted, divided by their uncertainties, have a standard deviation of 0.3.

AIMS allows different prescriptions for the surface-effects correction. We check the effect of using the one-term correction. The differences between one- and two-term prescriptions depends on the number of modes observed. For four modes, the results with the two approaches are in good agreement. However, as the number of modes increases, the one-term prescription appears clearly inefficient at fitting observed oscillation frequencies, and systematically provides much younger ages.

We expect that a large part of the AGB contamination will have been removed from the sample by filtering out $v_{\text{max}}$ values smaller than 15 Hz and selecting targets with at least 4 radial modes. Nevertheless, we test the effect on the age determination of a mass of mis-classification BHB/AGB. We select synthetics AGB models, and spectra, and derive their stellar parameters using AIMS, as above described, that is, with a grid of models that stop at $R = 25 R_\odot$ in the RGB phase. As a consequence of the mass-classification, AIMS either does not converge or provides very high (and unrealistic—20–25 Gyr) stellar ages.

Computing $\nu - 2$ oscillation modes is very time consuming and we have not used all of them in the fitting process. However, their frequencies for our sample have been determined at the same time as the radial ones, allowing us to estimate the mean value of the small frequency separation ($\delta \nu_{\text{02}} = \nu_{02} - \nu_{01}$). This seismic index, the values of which depend also on the evolutionary state, is a good proxy of the stellar mass** and can help us to identify genuine massive stars. The trend of $\nu_{02}$ with $\delta \nu_{\text{02}}$ is generally consistent with the masses assigning that our sample is formed by RGB. For a subset of 22 targets, we also have measured the dipole-mode period spacing. These values and the $\nu_{02}$ ones are consistent with classifying these targets as RGB stars, including among them two of the massive/ young stars.

**Cross-checking using Gaia parallaxes.** We check the consistency of derived stellar parameters with luminosity values from Gaia DR2. These values are obtained using Gaia DR2 parallaxes, 2MASS K, apparent magnitudes and bolometric corrections** appropriate to the atmosphere parameters (seismic surface gravity, and spectroscopic $T_{\text{eff}}$ and photopsheric chemical composition). The most important contribution to the uncertainty in the absolute uncertainty comes from parallax, which suffers from a zero-point offset of the order of few tens of microarcseconds. The effect of different extinction estimates turns out to be only 0.008 dex for the magnitude $K_s$.

We perform two new runs of stellar parameter estimation using luminosity (with offset of 30 mas and 50 mas) instead of $v_{\text{max}}$ as observational constraint in
AIMEES. The results are generally in good agreement with previous ones. The fits are dominated by highly precise frequencies and the still large errors affecting luminosity do not allow us to discriminate in case of multimodal posterior distributions not even in the reliability of a ≈ value.

The stellar radii derived from the Stefan–Boltzmann law with the spectroscopic effective temperatures and Gaia luminosity values have been compared with those inferred using AIME. The residuals divided by the relevant uncertainties have a standard deviation close to one, with no apparent trend with, for example, Δνmax. The median offset is, on the other hand, sensitive to the assumed zero-point parallel offset (better agreement with 50 μas zero-point parallel offset) and to the effective temperature scale (consistent with the results above).

Cross-checking using global seismic parameters for a larger set of ex situ and in situ stars. As an additional test for robustness, we use the code PARAM\(^2\) to infer more detailed structure, radii and ages for the larger set of stars (Extended Data Fig. 1). We consider all stars (RC and RGB) with average seismic parameters (Δν, ν\text{max}) determined by the COR pipeline\(^\text{5}\). While the results from the detailed analysis are more precise and more accurate (the median age uncertainty given by PARAM is 25% instead 11%), we use the age distribution of the wider sample to check whether our conclusion that the stars in these groups are likely to belong to in situ and ex situ stellar populations, and therefore such a modelling provides a means of statistically comparing the age distributions while taking the age uncertainties properly into account. We assume that age measurements of stars in a given population are drawn from a normal distribution with a mean age μ and intrinsic age spread σ, with some measurement error σm (derived from the posterior probability given by AIME). We include an outlier term in our model, assuming that in each population there is an overdensity at younger age due to our measurement of groups which are ‘overmassive’ (probably due to binary interactions) and therefore appear young. We assume that these outliers are also distributed normally with a mean age μ, a spread τ, and contributing some fraction ε.

Although we have no good way of determining whether these stars are bona fide young stars or indeed, overall massive stars, we are confident that this would not fundamentally change the results of the HBM analysis we perform. In particular, since the model always assumes contamination at younger ages, the only effect of including some bona fide young stars (which, as we have determined, are probably separate from the GC stars) would be to artificially increase the value of our epsilon (contamination fraction) parameter, and would not significantly alter the mean age of the target population.

We sample the posterior probability distribution given the data in each group in abundance space using pmc3. We make use of the the No-U-Turn Sampler (NUTS), a variant of Hamiltonian Monte Carlo, which uses the gradients of the likelihood function to facilitate rapid convergence and sampling of the posteriors over many parameters. For each population, we take 1,000 samples of the posterior over 4 independent chains after allowing 1,000 burn-in steps, for a total of 4,000 samples.

Effect of selection criteria. The classification in in situ and ex situ populations of our low-metal sample ([Fe/H] < −0.5) is based on a particular cut in the [α/Fe] plane and in eccentricity. Unfortunately, there is no consensus in the literature on which cut should be adopted nor on the value of [Fe/H] defining the low-metallicity end of the Galactic thin disk. We have analysed the effect of considering different division lines in the [α/Fe] plane and of shifting the threshold of eccentricity from 0.7 to 0.6 (Extended Data Fig. 6). Besides the selection made in the main text, we have used two other division lines: [Mg/Fe] = −0.5 [Fe/H] − 0.3 (used in a previous work\(^\text{3}\)) and [Mg/Fe] = −0.2 [Fe/H] (used in a previous work\(^\text{3}\)) which have a lower value than the one used in the main text but with a zero point of 0 instead of 0.05. The conclusions of the paper do not change: in the in situ high-age population (group C) is slightly older than that of GES (group A). Different selection criteria modify mainly the composition of the group B, which is not associated with a particular stellar population. It contains what is not in A or C, and in particular may contain some contamination from the low-metallicity end of the thin disk. The two stars at the high-metallicity end ([Fe/H] > 0.7) of our group B are probably part of the thin disk as indicated by their orbital and chemical properties. They can end in a different group depending on the cut used; however, that does not fundamentally change the age distributions of the other two groups. We believe that the contamination from the thin disk does not affect other stars of group B, as these two are the only ones having a maximum vertical excursion lower than 1 kpc.

Data availability
All raw observational data are publicly available: Kepler light curves at https://archive.stsci.edu/kepler/publiclightcurves.html; Gaia DR2 at https://gea.esac.esa.int/archive and APOGEE DR14 may be accessed via https://www.sdss.org/dr14/. APOGEE DR14 raw data have been used in the top panels of Fig. 2. Processed data, such as individual frequencies, orbital and stellar parameter, are available in Supplementary Table 1 or on request.

Evolutionary tracks are publicly available at https://doi.org/10.5281/zenodo.4032320 and theoretical stellar models and oscillation frequencies are available on request.

References
1. Peeples, P. J. E. Principles of Physical Cosmology (Princeton Univ. Press, 1993).
2. Kauffmann, G., White, S. D. M. & Guiderdoni, B. The formation and evolution of galaxies within merging dark matter haloes. Mon. Not. R. Astron. Soc. 264, 201–218 (1993).
3. Gaia Collaboration Gaia Data Release 2. Summary of the contents and survey properties. Astron. Astrophys. 616, A1 (2018).
4. Majewski, S. R. et al. The Apache Point Observatory Galactic Evolution Experiment (APOGEE). Astron. J. 154, 9 (2017).
5. Helmi, A. et al. The merger that led to the formation of the Milky Way’s inner stellar halo and thick disk. Nature 563, 85–88 (2018).
6. Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E. & Deason, A. J. Co-formation of the disc and the stellar halo. Mon. Not. R. Astron. Soc. 478, 611–619 (2018).
7. Myeong, G. C., Vasilev, E. Iorio, G., Evans, N. W. & Belokurov, V. Evidence for two early accretion events that built the Milky Way stellar halo. Mon. Not. R. Astron. Soc. 488, 1235–1247 (2019).
8. Borucki, W. J. Kepler mission: development and overview. Rep. Prog. Phys. 79, 036901 (2016).
9. Belokurov, V. et al. The biggest splash. Mon. Not. R. Astron. Soc. 494, 3880–3898 (2020).
10. Di Matteo, P. et al. The Milky Way has no in-situ halo other than the heated thick disc. Composition of the stellar halo and age-dating the last significant merger with Gaia DR2 and APOGEE. Astron. Astrophys. 632, A4 (2019).
11. Vincenzo, F. et al. The fall of a giant. Chemical evolution of Enceladus, alias the Gaia Sausage. Mon. Not. R. Astron. Soc. 487, L47–L52 (2019).
12. Miglio, A. et al. PLATO as it is: a legacy mission for galactic archaeology. Astron. Nachr. 338, 664–667 (2017).
13. Haywood, M. et al. In disguise or out of reach: first clues about in situ and accreted stars in the stellar halo of the Milky Way from Gaia DR2. Astrophys. J. 863, 113 (2018).
14. Mackerezh, J., Kerber, B. L. C., L. et al. Origin of accreted stellar halo populations in the Milky Way using APOGEE, Gaia, and the EAGLE simulations. Mon. Not. R. Astron. Soc. 482, 3426–3442 (2019).
15. Grand, R. J. J. et al. The dual origin of the Galactic thick disc and halo from the gas-rich Gaia Enceladus–Sausage merger. Mon. Not. R. Astron. Soc. 497, 3808–3820 (2020).
16. Schuster, W. J., Moreno, E., Nissen, P. E. & Pichardo, B. Two distinct halo populations in the solar neighborhood. III. Evidence from stellar ages and orbital parameters. Astron. Astrophys. 538, A21 (2012).
17. Hawkins, K., Jofrè, P., Gilmore, G. & Maserer, T. On the relative ages of the α-rich and α-poor stellar populations in the Galactic halo. Mon. Not. R. Astron. Soc. 445, 2575–2588 (2014).
18. Das, P., Hawkins, K. & Jofrè, P. Ages and kinematics of chemically selected, accreted Milky Way halo stars. Mon. Not. R. Astron. Soc. 493, 5195–5207 (2020).
19. Gallart, C. et al. Uncovering the birth of the Milky Way through accurate stellar ages with Gaia. Nat. Astron. 3, 932–939 (2019).
20. Helmi, A. Streams, substructures, and the early history of the Milky Way. Annu. Rev. Astron. Astrophys. 58, 205–256 (2020).
Acknowledgements

J.M., J.T.M., A.M., F.V. and E.W. acknowledge support from the ERC Consolidator Grant funding scheme (project ASTEROCHRONOMETRY, G.A. no. 772293). F.V. acknowledges the support of a Fellowship from the Center for Cosmology and AstroParticle Physics at The Ohio State University. M.V. is supported by FEDER - Fundo Europeu de Desenvolvimento Regional through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by grants PTDC/FIS-AST/30389/2017 and POCI-01-0145-FEDER-030389. C.C. acknowledges partial support from DFG Grant CH1188/2-1 and from the CHETEC COST Action (CA16117), supported by COST (European Cooperation in Science and Technology). G.B. acknowledges fundings from the SNF AMBIZZONE grant no. 185805 (Seismic inversions and modelling of transport processes in stars) and from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (project STAREX). G.R.D. acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (CartographY GA 804752). M.B.N. acknowledges support of the UK Space Agency. O.J.H. acknowledges the support of the UK Science and Technology Facilities Council (STFC). This article made use of AIMS, a software for fitting stellar pulsation data, developed in the context of the SPACEINN network, funded by the European Commission’s Seventh Framework Programme. This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the Gaia Multilateral Agreement. The computations described in this paper were performed using the University of Birmingham’s BlueBEAR HPC service, which provides a high-performance computing service to the university’s research community. See http://www.birmingham.ac.uk/bear for more details. We thank S. McGee for reading and commenting on the manuscript.

Author contributions

J.M. led the project, with help from J.T.M., A.M., F.V., C.C. and W.J.C. G.R.D. designed the pipeline for the light-curve analysis. J.M., J.T.M., A.M., F.V., G.R.D., O.I.H., M.B.N., S.K. and W.E.v.R. worked on extracting mode parameters from the Kepler light curves. J.M., A.N. and R.S. performed the stellar modelling and theoretical oscillation frequency computations. J.M., G.B. and B.M.R. worked on the stellar parameter determination from individual frequencies using Bayesian inference code AIMS. A.M. estimated stellar parameters from global observational constraints using the code PARAM. B.M. and M.V. provided global seismic parameters. J.T.M. and F.V. performed the kinematics and chemical composition analysis from Gaia DR2 and APOGEE DR14 datasets. E.W. derived absolute stellar luminosity from Gaia DR2. J.W.F. provided radiative opacity data at low temperature for the alpha-enhanced chemical mixture used in the stellar evolution code. All authors have contributed to the interpretation of the data and the results, and discussion and giving comments on the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-021-01347-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-021-01347-7.

Correspondence and requests for materials should be addressed to J.M.

Peer review information Nature Astronomy thanks David Soderblom and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021
Extended Data Fig. 1 | Data samples. a, Diagram [$\alpha$/Fe] versus [Fe/H] for all the Kepler-APOGEE-DR14 sample (grey dots). Orange symbols are the targets in our sub-sample: red giant stars with [Fe/H] < -0.5, and blue ones are the first ascending red giant branch targets selected for characterization in this paper. b, $T_{\text{eff}}$ versus $\nu_{\text{max}}$ diagram (equivalent to Kiel diagram) of our target sample (color-coded by metallicity), overlying the complete Kepler-APOGEE-DR14 one (grey empty and full symbols). The dashed lines corresponds to two [$\alpha$/Fe]=0.2 evolutionary tracks: blue $M=0.9$ $M_\odot$, [Fe/H]=-1.0; orange, same mass but [Fe/H]=-0.5.
Extended Data Fig. 2 | Posterior probability distributions for KIC 4143467 stellar properties as inferred by AIMS. a-f: age, mass, radius, mean density, luminosity and frequency at maximum power, respectively. The oscillation spectra of this target is shown in first panel of Fig. 1. The vertical dash-dotted lines indicate the value of the corresponding parameter in the best-fitting model from the MCMC sampling.
Extended Data Fig. 3 | Posterior probability distributions for KIC 12111110 stellar properties as inferred by AIMS. a-f: age, mass, radius, mean density, luminosity and frequency at maximum power, respectively. The vertical dash-dotted lines indicate the value of the corresponding parameter in the best-fitting model from the MCMC sampling.
Extended Data Fig. 4 | Age distribution using PARAM for the APOGEE-Kepler sample with stellar radius limited to 14 $R_\odot$. a–b, [$\alpha$/Fe] vs. [Fe/H] distribution of the sample coloured by age (a) and eccentricity (b). The symbol size scales with $\nu_{\text{max}}$. c, Age distributions of accreted and in-situ stars, so classified from their [$\alpha$/Fe] and eccentricity values; d, Kiel diagram of the sample coloured by metallicity. Notice that the ‘very old’ (yellow dots $T_{\text{eff}} > 5400$ K) suggest that we have underestimated the mass loss for those stars.
Extended Data Fig. 5 | Age distribution using PARAM for the APOGEE-Kepler sample with stellar radius limited to 8 R⊙. a–b, [α/Fe] vs. [Fe/H] distribution of the sample coloured by age (a) and eccentricity (b). The symbol size scales with νmax. c, Age distributions of accreted and in-situ stars, so classified from their [α/Fe] and eccentricity values; d, Kiel diagram of the sample coloured by metallicity. Notice that the ‘very old’ (yellow dots Teff > 5400 K) suggest that we have underestimated the mass loss for those stars.
Extended Data Fig. 6 | Age and eccentricity distributions for different selection criteria for in-situ and accreted populations. Age against eccentricity (e) for the stars in the sample coloured by [Fe/H]). Circles represent age values of the best fitting models, and horizontal lines their uncertainties (16%-84% C.I. from full posterior distributions). Uncertainties on e are smaller than the symbol size. The diamond represents ν Indi\textsuperscript{25} (not included in the distributions). The histogram above reflects the combined posterior distributions for the stars in each selection. a,c, division line [Mg/Fe] = -0.5 [Fe/H]-0.3 (ref. 14). b,d, division line [Mg/Fe] = -0.2 [Fe/H] (ref. 26). Top and bottom panels correspond to eccentricity threshold 0.7 and 0.6 respectively.
Extended Data Fig. 7 | Probabilistic graphical model of that used to fit the mean age and intrinsic age spread of the *in*- and *ex-situ* populations defined on the basis of element abundances and orbital parameters. We assume the measured ages are drawn from an underlying true age $\theta$ distribution that is Gaussian with a mean $\mu$ with a standard deviation $\tau$. We assume that the true age distribution is contaminated by stars whose mass is higher than expected (and therefore appear younger), likely due to some poorly understood process such as binary interactions. We model these contaminants as also being drawn from another normal distribution with a mean $\mu_c$ and spread $\tau_c$ which has a fractional contribution $\varepsilon$ to the total age distribution (hence the main population contributes $1-\varepsilon$).