Galactic Archeology
Prospects and issues
Sofia Feltzing
Lund Observatory
Galactic Archeology

• Builds on the understanding of
  – stars as time-capsules
    ★ Solar-like stars retain, in their atmospheres, the same composition of elements as the gas cloud from which they formed.
  – each star formation event has a unique signature
    ★ The chemistry in each star formation event is influenced by a unique set of chemical enrichment events.
  – star clusters disperse in the Galactic potential

Freeman & Bland-Hawthorn 2002 ARA&A 40 487
• So far we have likely
  – truly identified 1 (one!) such event
    ★ The moving group HR1614 (de Silva et al.)
  – we have disproved several, e.g.,
    ★ Herculus moving group (Bensby et al.)
    ★ KFR08 stream (Liu, Ruchti, Feltzing)
  – have we just been unlucky or is it really hard to do this? (see also Mitschang et al. 2014, Ting et al. 2012)
Gaia changes all

Unparalleled dataset with motions and positions for 10^9 stars across the Milky Way, 10^4 times more stars with full phase-space information; 10^6 volume increase; 100x more accurate.

GAIA’S REACH

The Gaia spacecraft will use parallax and ultra-precise position measurements to obtain the distances and ‘proper’ (sideways) motions of stars throughout much of the Milky Way, seen here edge-on. Data from Gaia will shed light on the Galaxy’s history, structure and dynamics.

Previous missions could measure stellar distances with an accuracy of 10% only up to 100 parsecs.

Gaia’s limit for measuring distances with an accuracy of 10% will be 10,000 parsecs.

Gaia will measure proper motions accurate to 1 kilometre per second for stars up to 20,000 parsecs away.

From A. Helmi @ ESO in 2020
Ground-based follow-up to Gaia is essential as not all stars will have spectra and hence will not have, e.g., [Fe/H] derived

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astrometry</td>
<td>7 ( \mu )arcsec at ( V = 10 )</td>
</tr>
<tr>
<td></td>
<td>12 – 25 ( \mu )arcsec at ( V = 15 )</td>
</tr>
<tr>
<td></td>
<td>100 – 300 ( \mu )arcsec at ( V = 20 )</td>
</tr>
<tr>
<td>Photometry</td>
<td>low resolution prism spectra to ( V = 20 )</td>
</tr>
<tr>
<td>Radial velocities</td>
<td>1 – 15 km s(^{-1}) to ( V \lesssim 17 )</td>
</tr>
</tbody>
</table>

Ca \( \Pi \) infra-red triplet (847 – 874 nm)

RVs down to \( \sim 15.2 \) and abundances to 11

Pre-launch estimates
Ground-based follow-up

Adapted from Gilmore et al. (2012)
Ground-based follow-up
Ground-based follow-up

Surface gravity

Elemental abundances
Ground-based follow-up

Adapted from Gilmore et al. (2012)

Requirements on spectrographs come from how well we need to know elemental abundances and stellar parameters.
Size of features

Example of the precision/accuracy you wish to have.

The typical size of features seen in abundance trends are of ~0.2 dex, or less.

Plot based on data from Klaus Furhmann’s studies (priv. comm.)
Nieva and Przybilla: BA-type supergiants + B-stars & HII-regions (Esteban+ 2005)

- only recently have more and extensive samples with information beyond the the solar neighbourhood become available (SEGUE, APOGEE)
- however, they only cover certain parameter spaces for stars (e.g. no OB stars)
Sizes to measure

Chemical tagging

Chemical tracing

HR1614
Hyades
Collinder 261

σ ≤ 0.05 dex

σ ≤ 0.1 - 0.2 dex
Size of survey

Precision vs. # of stars.

Low probability of false detection (p = 0.01)
Examples

Precision vs. # of stars.

Low probability of false detection (p = 0.01)
Assuming we know what $\sigma$ we want
Resolution vs. Signal-to-Noise ratio in spectra for abundance determination

Ability to retrieve a given equivalent width (EW) at any $\lambda$

$\sigma(\text{EW})/\text{EW}$ prop. $(R)^{-1/2}$

SNR=185 per Å
SNR=154 per Å

EW=25mA
S/N=154 per Å, ~50 per pix.
S/N=185 per Å, ~60 per pix.
S/N=230 per Å, ~75 per pix.
S/N=308 per Å, ~100 per pix.

0.1 dex

Luca Sbordone for 4MOST consortium, used for PDR
See also Gustafsson (1992)
Galactic Archeology

• Builds on the understanding of
  – stars as time-capsules
  – each star formation event has a unique signature
  – star clusters disperse in the Galactic potential

• Pure dynamical studies can also lead us further
  – Most knowledge is from the solar neighbourhood
Disk dynamics

Expectation: Bar and spiral arms in the model ➝ makes for more complex velocity-plane

Antoja thesis

Kinematic structures prevail at ∼1 kpc (RAVE, Antoja et al. arXiv:1205.0546)
Galactic Archeology II

• “Quantifying the strength of radial migration in the Milky Way is one of the most pertinent action items for the next generation of Milky Way surveys”
  – … but, how do you that?
    ★ What stars should we select?
    ★ What properties of the stars should we measure?
  – …
  – Do we have an answer to such questions?
# Existing instruments

<table>
<thead>
<tr>
<th>NIR</th>
<th>UV</th>
<th>UV</th>
</tr>
</thead>
</table>
| **APOGEE (I+II)**  
Multifibre spectrograph part of SDSS  
R ~ 22 500 (1.51 - 1.70 μm)  
300 fibres  
https://www.sdss3.org/future/apogee2.php  
Survey: APOGEE – 10⁵ stars | **HERMES**  
Multifibre spectrograph on AAT  
R ~ 28 000 a  
390+ fibres  
Survey: GALAH – 10⁶ stars | **FLAMES**  
Multifibre spectrograph on VLT  
R ~ 20 000  
> 100 fibres  
Survey: Gaia-ESO – 10⁵ stars |
### Upcoming survey instruments

<table>
<thead>
<tr>
<th>NIR</th>
<th>UV</th>
<th>UV</th>
</tr>
</thead>
</table>
| **MOONS**  
NIR multifibre spectrograph being built for VLT  
R ~ 5000 (0.64-1.8 μm)  
R ~ 9000, 20 000, 20 000 (0.7-0.9, 1.17-1.26, 1.52-1.63 μm)  
1024 fibres  
Being built by consortium lead by ATC, UK  
PI: Michielie Cirasuolo  
http://www.roe.ac.uk/~ciras/MOONS/VLT-MOONS.html | **WEAVE**  
Multifibre spectrograph being built for WHT  
R ~ 20 000 and R ~ 5000  
800 fibres (switchable R)  
Gaia follow-up (4MOST in the North), extra-galactic science  
Netherlands, UK, Spain, France, Italy  
Project scientist: Scott Trager  
http://www.ing.iac.es/weave/ | **4MOST**  
Multifibre spectrograph to go on VISTA  
R ~ 20 000 and R ~ 5000  
800 + 1600 fibres (sim.)  
Gaia and eROSITA follow-up  
10-20 million LR stars  
1-2 million HR stars  
LR to V ~ 20 w SNR 10/Å  
HR ~ 16.5/17 w SNR of 170/Å  
PI: Roelof de Jong  
http://www.4most.eu |

Cirasuolo et al (SPIE, 2014)  
de Jong et al (SPIE, 2014)
Complementarity

• Many of the surveys are highly complementary
  – North <—> South being an obvious one
  – NIR vs Optical; but not the same stars studied (so far)
  – magnitude range (i.e. GALAH does everything down to V=14, 4MOST almost starts there)

• Cross-calibration not so well developed
WEAVE and 4MOST similar coverage in LRS as SEGUE
Can such precision in abundances be achieved?

- Lets take a step back and consider the limitations
  - stars as time-capsules
    ★ Solar-like stars retain, in their atmospheres, the same composition of elements as the gas cloud from which they formed.
  - but do we measure the “true” values of the elemental abundances?
    ★ Lets look at a few examples that are both discouraging and heartening
Diffusion changes abundance patterns

- Effects of stellar evolution.
- Evidence that selective diffusion occurs in stars at MS and TOP in globular clusters and M67.
- Up to 0.2 dex.

Önehag et al. 2014 A&A 562 A102
Gruyters et al. 2013 A&A 555 A31
Is the trend real?

Fig. 19. Abundance trend plots for the $\alpha$-elements (O, Mg, Si, Ca, and Ti) when only including stars within the temperature interval $5500 < T < 6100$ K (cf. Fig. 7). Typical error bars are shown on the right-hand side of the plots. Note the different scales on the ordinates.

Younger stars do not show this behaviour. Instead there appears to be a rather large scatter in age over the whole metallicity range ($-0.8$ to $+0.4$ dex), i.e., no apparent age-metallicity relation.

8.3.4. $[\alpha/Fe]$ as a proxy for age?

Recently, Navarro et al. (2011) have argued that it is better to relate stars to different populations based on their elemental abundances rather than other properties such as kinematics. This statistic selection based on kinematics causes overlaps between various abundance trends is evident from the nature of that selection process (see Sect. 7), and casts doubt on the reality of distinct trends for different stellar populations. This argument was used, e.g., by Bovy et al. (2012) when they investigated the scale-height of mono-abundance populations (i.e., stars that fall within a narrow range of elemental abundances, e.g., $[\alpha/Fe]$ and $[Fe/H]$) in the SEGUE data-set.

To better understand the formation and evolution of the Milky Way it is very desirable to have stellar ages as well as elemental abundances. Given the overall structure of the elemental abundances and ages observed in the Milky Way (e.g., Edvardsson et al. 1993) it has been suggested that the amount of $\alpha$-enhancement in a star can be used as a proxy for the star's age (Liu & van de Ven 2012). However, age is a very difficult property to derive for most stars (e.g., Soderblom 2010). As our sample contains a fair portion of turn-off and sub-giant stars we are in a position to investigate if old ages are a common feature for all stars with enhanced $[\alpha/Fe]$ in the Solar neighbourhood.

Figure 20b shows that this is indeed the case for stars older than about 8 Gyr and thus that $[\text{Ti}/\text{Fe}]$ can be used as a proxy for age for stars in the sense that young and old stellar populations can be distinguished. Other studies are also finding that various $\alpha$-elements correlate with ages in this sense. For example Ramírez et al. (2013), their Fig. 17, shows the same results as our Fig. 20b, but for $[\text{O}/\text{Fe}]$ as a function of age. However, this result is only valid for dwarf stars in the direct Solar neighbourhood. We do not know if the same is true elsewhere in the Galaxy or indeed recoverable for other stellar evolutionary stages. Bensby et al. (2013) provides data for 58 microlensed dwarf and turn-off stars in the Galactic bulge. These stars, tentatively, show the same trend as the stars in the Solar neighbourhood making it plausible that the connection between $\alpha$-enhancement and age is a property shared by many stellar populations in the Galaxy.

8.4. Metal-rich and $\alpha$-enhanced stars

In Fig. 21 we show the $[\text{Ti}/\text{Fe}]$ abundances trends for all stars in our sample with "good" ages. We find a similar division of the...
• NLTE, 3D effects can be severe, e.g., Ruchti et al. (2013).
Stellar parameters

Bensby et al. 2014 A&A 562 A71
... but abundances work

Comparison Bensby et al (2014) and:

Valenti & Fischer (2005)
140 stars

Reddy et al. (2003,2006)
64 stars

Adibekyan et al. (2012)
168 stars
Uncertainty in understanding and measurement of E(B-V) causes severe limitations for absolute results.

We should

- Spend some time and money on
  - Understanding the stars better or at least
    ★ Obtain suitable corrections that puts stars of different kinds onto the same scales
  - Obtain suitable stellar samples to cross-calibrate the surveys. This means thousands of stars.
    ★ Cannon-Fodder is starting
What about the design of the surveys?

- Surveys will inherently be limited due to selection effects
  - Common examples are when you have a magnitude limited sample (compare next slide)
    ★ Dwarfs and giants sample different volumes
  - Another issue is to define what an observable is
    ★ There is potentially a sliding scale here, but would suggest that we are very careful (compare slide after next and Creevy et al)
Spectral lines as well as SEDs vary significantly with spectral type. Metallicity also plays a major role.
What is an observable?

- This seems to need a better definition or at least when you design your survey you need to think about it
  - parallax, proper motion, observed magnitudes
  - spectra themselves (?)
  - position and velocity?
  - Temperature, log(g), [Fe/H]
  - But is then [Mg/Fe] and observable or a derived parameter?

Rix & Bovy 2013 A&Arv 21 61
• Answers to these questions are likely at the core of how we chose to make progress
So far

Past
RAVE
SEGUE

Now
Gaia-ESO
APOGEE I
GALAH
LAMOST

Soon
4MOST
WEAVE
APOGEE II
PFS

Future
E-ELT MOS
What else?

Resolution
# targets

Starting to step outside MW (PFS)
Beyond

• What do we want beyond the currently operational and planned instrumentation?

• What is more important to make progress?
  ★ Larger samples?
  ★ Smaller σ?
  ★ More Milky Way or other galaxies?
ELT era

• What will the ELT era bring for Galactic Archeology?

• Move fully outside the Milky Way – for the first time allow proper comparisons from spectroscopic studies between MW and other galaxies (N.B. we will be back to Edvardsson et al rather than moving towards Melendez et al.)

• With AO - studies of very crowded regions, such as Bulge and clusters, down to the turn-off and below
Potential ideas

- A single slit spectrograph on an 8-10m, dedicated to following up the surveys [= better $\sigma$]

- A very specialised MOS on an 8-10m; e.g., very high resolution of Eu line (weak lines, hfs, only direct probe of $r$-process) [= better $\sigma$]

- LSST follow-up/variable
Potential ideas

- A single slit spectrograph on an 8-10m, dedicated to following up the surveys [= better $\sigma$]

- A very specialised MOS on an 8-10m; e.g., very high resolution of Eu line (weak lines, hfs, only direct probe of $r$-process) [= better $\sigma$]

- LSST follow-up/variable sky

- But - perhaps better spend our energy on the analysis/stellar understanding
  - Follow-up for asteroseismology?
  - More interferometry?
  - Models?
Type of object

• Mixing different types of tracers can be (very) hard if you aim for high precision.
• This is because (many) analysis methods have been developed in the context of one particular type of star.

There are good reasons for this – the stellar spectra are rather different and challenges differ depending on evolutionary phase and metallicity.
Optimizing 4MOST

Line-list wetted for science

Ruchti et al. (in prep.)

Hα, 10 low EP and 10 high EP Fe I
10 Fe II lines
Mg T or Ca 616.2
Abundance analysis

- Traditional so called “fine analysis” (Drake 1992) based on $W_\lambda$.
- Full spectrum synthesis.
- Finding best matching template spectrum.
- Get stellar parameters from independent source, and only analyse the atomic/molecular lines.
- ...
- ...
Gaia Benchmark stars – sheep/goats

Fig. 2. Difference between the metallicity obtained by each node and the mean literature value (see Sect. 2). Stars are ordered by effective temperature. Different symbols correspond to the different methods, which are indicated in the legend.

Fig. 3. Metallicity (upper panel) and microturbulence velocity (lower panel) obtained by different methods for each GBS as a function of temperature. Black dots correspond to the values of $v_{\text{mic}}$, also obtained from the GE Sr relation of Bergemann and Hill.

As in previous figures, we illustrate the difference in metallicity as a function of GBS in order of increasing temperature in the upper panel. In the lower panel of Fig. 4, we plot together the stars observed with the same instrument. Different instruments are separated by the dashed line. The value of the spectral resolution before convolution is indicated at the top of the figure.

It is interesting to comment on the result of $\psi_{\text{Phe}}$, which has the lowest original resolution and is the coldest star, because it shows the greatest difference. In the case of the LUMBA method, the synthetic spectra produced by SME need to have a given resolving power, which is set to be constant along the entire spectral range. In the original spectra, this is not completely...
We need “standards”

• This has woefully been lacking in spectroscopic analysis.
• Yes, “everyone” analysis the sun, perhaps Arcturus or μLeo – but this covers a very limited set of stellar properties.
• The Gaia-Benchmark stars are a first, but not final, step in the right direction.
• In addition, the proper establishing of equatorial spectroscopic fields should be a high priority (could include open clusters).
A wider plan?

- Not only do we need common stars and clusters.
  - Reduced spectra need to be available for all to analyse.
  - Analysis codes should be well documented in the public domain and also, preferably, in some form of “open source”.
  - Analysis codes should be kept on version control and DR-codes kept.
For now...

• Do not mix apples and oranges – make life simple!

• Make the analysis self-consistent
  - Differential studies – useful for many things, absolute values not necessarily needed.

• (arbitrary) re-calibrations of derived parameters/abundances
  - Leaves us not knowing what we are actually looking at – will not be reproducible/compatible/possible to compare with (certainly not with other methods).
• An ultimate goal of all the large on-going and up-coming surveys will be to join the data in order to achieve a much bigger data-set

• This requires that the results are reproducible such that
  – Each survey can analyse both their own and other spectra for the same star(s), i.e. spectra are publicly available.
  – Results are fully documented and reproducible.
Sheep and goats

Comparison of Stellar Parameters of 18 Sco

<table>
<thead>
<tr>
<th>$T_{\text{eff}}$ (K)</th>
<th>Error (K)</th>
<th>log $g$ (dex)</th>
<th>Error (dex)</th>
<th>[Fe/H] (dex)</th>
<th>Error (dex)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>5823</td>
<td>6</td>
<td>4.45</td>
<td>0.02</td>
<td>0.054</td>
<td>0.005</td>
<td>This work</td>
</tr>
<tr>
<td>5816</td>
<td>4</td>
<td>4.45</td>
<td>0.01</td>
<td>0.053</td>
<td>0.003</td>
<td>Ramírez et al. (2014b)</td>
</tr>
<tr>
<td>5824</td>
<td>5</td>
<td>4.45</td>
<td>0.02</td>
<td>0.055</td>
<td>0.010</td>
<td>Monroe et al. (2013)</td>
</tr>
<tr>
<td>5810</td>
<td>12</td>
<td>4.46</td>
<td>0.04</td>
<td>0.05</td>
<td>0.01</td>
<td>Tsantaki et al. (2013)</td>
</tr>
<tr>
<td>5831</td>
<td>10</td>
<td>4.46</td>
<td>0.02</td>
<td>0.06</td>
<td>0.01</td>
<td>Meléndez et al. (2012)</td>
</tr>
<tr>
<td>5817</td>
<td>30</td>
<td>4.45</td>
<td>0.13</td>
<td>0.05</td>
<td>0.05</td>
<td>da Silva et al. (2012)</td>
</tr>
<tr>
<td>5826</td>
<td>5</td>
<td>4.45</td>
<td>0.01</td>
<td>0.06</td>
<td>0.01</td>
<td>Takeda &amp; Tajitsu (2009)</td>
</tr>
<tr>
<td>5840</td>
<td>20</td>
<td>4.45</td>
<td>0.04</td>
<td>0.07</td>
<td>0.02</td>
<td>Meléndez et al. (2009)</td>
</tr>
<tr>
<td>5848</td>
<td>46</td>
<td>4.46</td>
<td>0.06</td>
<td>0.06</td>
<td>0.02</td>
<td>Ramírez et al. (2009a)</td>
</tr>
<tr>
<td>5818</td>
<td>13</td>
<td>4.45</td>
<td>0.02</td>
<td>0.04</td>
<td>0.01</td>
<td>Sousa et al. (2008)</td>
</tr>
<tr>
<td>5834</td>
<td>36</td>
<td>4.45</td>
<td>0.05</td>
<td>0.04</td>
<td>0.02</td>
<td>Meléndez &amp; Ramírez (2007)</td>
</tr>
<tr>
<td>5822</td>
<td>4</td>
<td>4.451</td>
<td>0.006</td>
<td>0.053</td>
<td>0.004</td>
<td>Weighted mean from the literature</td>
</tr>
</tbody>
</table>

Metallicity decreases, from >0.2 dex to –2.7 dex

Beware of the sheep and sheep-goat effect

Gustafsson (2004)  
Carnegie Observatories Astrophysics Series 4. Eds. McWilliam & Rauch

Smiljanic et al. 2014 arXiv1409.0568  
Not only apples and oranges different methods give different results too

Fig. 2. Difference between the metallicity obtained by each node and the mean literature value (see Sect. 2). Stars are ordered by effective temperature. Different symbols correspond to the different methods, which are indicated in the legend.

Fig. 3. Metallicity (upper panel) and microturbulence velocity (lower panel) obtained by different methods for each GBS as a function of temperature. Black dots correspond to the values of $v_{\text{mic}}$ also obtained from the GE Sr relation of Bergemann and Hill.

As in previous figures, we illustrate the difference in metallicity as a function of GBS in order of increasing temperature in the upper panel. In the lower panel of Fig. 4, we plot together the stars observed with the same instrument. Different instruments are separated by the dashed line. The value of the spectral resolution before convolution is indicated at the top of the figure.

It is interesting to comment on the result of $\psi_{\text{Phe}}$, which has the lowest original resolution and is the coldest star, because it shows the greatest difference. In the case of the LUMBA method, the synthetic spectra produced by SME need to have a given resolving power, which is set to be constant along the entire spectral range. In the original spectra, this is not completely

A133, page 9 of 27

$T_{\text{eff}}$ (K)
Apples and oranges

• Mixing different types of tracers can be (very) hard if you aim for high precision.
• This is because (many) analysis methods have been developed in the context of one particular type of star.
In summary, systematic differences between the stars and the Sun could arise due to the …

(i) analysis techniques (equivalent widths vs. spectrum synthesis)
(ii) stellar parameters
(iii) adopted grid of model atmospheres
(iv) treatment of line formation (LTE vs. NLTE)
(v) adopted gf-values
(vi) adopted line lists
(vii) spectral resolution
(viii) signal-to-noise ratio
(ix) problems with the spectrograph
(x) adopted solar spectrum (sky, Moon, moons of other planets, asteroids, solar atlas)
(xi) data reduction
(xii) determination of the continuum
(xiii) blends
(xiv) equivalent width measurements
(xv) adopted solar abundances
σ ≈ 0.04 dex around the mean

Better data (S/N, larger λ-range, # lines)


Bensby et al. 2014 A&A 562 A71
And stars move