Contributions to Mineralogy and Petrology (2018) 173:23
https://doi.org/10.1007/s00410-018-1448-3

ORIGINAL PAPER

A cross correlation method for chemical profiles in minerals, with an application to zircons of the Kilgore Tuff (USA)

L. C. Probst[1] · T. E. Sheldrake[1] · M. J. Gander[2] · G. Wallace[3] · G. Simpson[1] · L. Caricchi[1]

Received: 18 September 2017 / Accepted: 14 February 2018

Abstract
Magmatic crystals are characterised by chemical zonation patterns that reflect the thermal and chemical conditions within magma reservoirs in which they grew. Crystals that exhibit similar patterns of zonation are often interpreted to have experienced similar conditions of growth. These patterns of zonation may represent continuous processes such as cooling, or more instantaneous events such as magma injection, and provide an insight into the structure and evolution of a magmatic system, both temporally and spatially. We have developed an algorithm that is objectively able to quantify the similarity within and between suites of magmatic crystals from different samples. Significantly, the algorithm is able to identify correlation that occurs between the interiors of two crystals, but does not extend to the rim, which provides an opportunity to understand the long-term evolution of magmatic systems. We develop and explain the mathematical basis for our algorithm and introduce its application using cathodoluminescence images of zircons from the Kilgore Tuff (USA). The results allow us to correlate samples from two different outcrops that are found over 80 km apart.

Keywords Cross-correlation algorithm · Chemical zonation · Crystal · Magma reservoirs · Volcano · Caldera eruption

Introduction

Zoned magmatic crystals record the evolution and variation of the residual melt chemistry and related intensive parameters within magmatic systems (Davidson et al. 2007; Costa et al. 2009; Humphreys et al. 2012; Cashman and Blundy 2013). Such zoning patterns in crystals are typically revealed using back scatter electron and/or cathodoluminescence imaging on a scanning electron microscope (Ginibre et al. 2002a; Nasdala et al. 2003; Cheng et al. 2017). Characterising the pattern of zonation provides an opportunity to study the temporal evolution of chemical and physical conditions within magmatic systems before eruption (Ginibre et al. 2002b; Wallace and Bergantz 2005; Ginibre et al. 2007). Crystals that share the same magmatic history can exhibit the same pattern of zoning. However, geophysical and geochemical observations suggest that magma reservoirs are heterogeneous and contain magma with diverse chemical and physical properties (Miller and Smith 1999; Gudmundsson 2012; Ellis et al. 2014; Wotzlaw et al. 2015; Myers et al. 2016; Hartung et al. 2017). Consequently, the portions of crystal profiles that are correlated will vary (Fig. 1), depending upon the temporal and spatial heterogeneity of magmatic processes associated with the analysed samples.

One can assume that crystals from the same thin section experienced identical growth conditions upon solidification, and thus will be correlated from the rim (Fig. 1b; Wallace and Bergantz (2005)). Moving inward from the rim, similar zoning profiles are interpreted as representing a shared thermodynamic history of crystal growth. Shared physical history is inferred from the mixing of populations of crystals with similar and dissimilar histories. For example, crystals may originate from identical parental melts but these melts may experience different

Communicated by Mark S. Ghiorso.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00410-018-1448-3) contains supplementary material, which is available to authorized users.

✉ T. E. Sheldrake
thomas.sheldrake@unige.ch

[1] Department of Earth Sciences, University of Geneva, 13 rue des Maraîchers, 1205 Geneva, Switzerland

[2] Section of Mathematics, University of Geneva, 4 rue du Lièvre, 1205 Geneva, Switzerland

[3] Pacific Groundwater Group, Eastlake Ave E, 2377 Seattle, USA

Published online: 26 February 2018

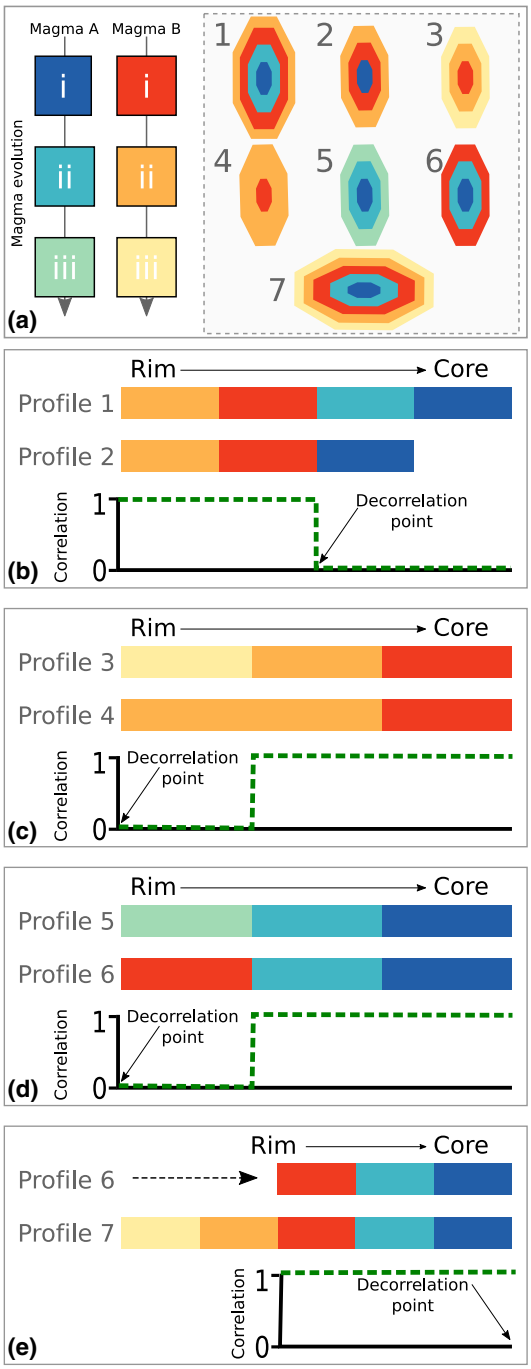

Fig. 1 Correlation scenarios; **a** Two hypothetical magmas and their relative evolutions in composition represented by different colours, alongside 7 zoned hypothetical crystals; **b** scenario where two profiles are correlated continuously from the rim, until a decorrelation point after which they are not correlated (dashed green line); **c**, **d** scenarios where profiles are correlated in the interior, but not continuously from the rim; **e** scenario where the rim of one crystal is correlated with the interior of another crystal, such that the correlation of the profiles begins from the rim of the shifted profile (no. 6)

thermodynamic conditions of growth prior to eruption (Fig. 1c), perhaps stalling at different levels in the crust. Alternatively, dramatic changes in intensive parameters (e.g. due to magma injection) may be localised and not recorded in all crystals (Fig. 1d). Furthermore, the inner portion of one crystal may be correlated with the outer portion of another if, for example, different portions of the same magma body are erupted in multiple events (Fig. 1e).

For two crystals that experienced the same growth conditions immediately prior to eruption or cooling to the solidus, a decorrelation point in their profiles can be identified, beyond which they are not correlated (Fig. 1b; Wallace and Bergantz 2005). However, the principle of a decorrelation point does not hold for the alternative scenarios where the inner portions of crystal profiles are correlated, or if the inner portion of one crystal is correlated with the outer portion of another crystal (Fig. 1c–e). Here, we develop a cross correlation procedure for these alternative scenarios, which will be particularly suitable for the correlation of crystals of the same eruption or intrusion, which are not necessarily from the same sample. The ability to correlate the internal regions of crystals will also be useful when analysing a time-series of samples or specimens collected at different outcrops, to elucidate the temporal evolution of crystal populations within a magmatic system. We present the geological basis for the cross correlation algorithm before describing the mathematical procedure that is adopted in this study. We finally provide an example application of the procedure presented here, using zircons from the Kilgore Tuff (Idaho, USA).

Correlation of crystal profiles

Signals, such as the chemical variability within a magmatic crystal, can be quantitatively compared using different mathematical strategies including linear correlation methods (Wallace and Bergantz 2005), wavelet-based methods (Wallace and Bergantz 2002) or two-dimensional reconstructions of crystal profiles (Cheng et al. 2017). Here we aim to further develop the linear correlation approach, using the principles of cross-correlation (Galton 1888; Pearson 1894; Yarlagadda 2010), which have been applied in other fields for pattern recognition, such as the correlation of DNA sequences (Chu et al. 2004). This linear correlation method produces a correlation coefficient that quantifies the similarity between portions of two signals from different crystals. The cross-correlation factor provides the best alignment for which the correlation between the signals is optimised (Yarlagadda 2010).

Sample orientation

Chemical traverses in minerals or images carrying chemical information are often collected in two-dimensional sections of rock samples. As a consequence, the orientation of crystals is generally random and unknown (Ginibre et al. 2002a; Wallace and Bergantz 2004). The lack of exposure of the inner portions of a crystal, or differences in the orientation of two crystals, reduces the length of usable profiles and changes the distance over which chemical variations are observed (Fig. 2).

Figure 2 shows several possible crystal transects together with the corresponding profiles. An ideal transect (b) is centred and oriented perpendicularly to the crystal boundary. Non-ideal transects like a, c and d suffer from different problems. Exposure limited to the outer portions of a crystal results in the total or partial loss of chemical information contained in the inner portion of a crystal (Fig. 2a, d). Profiles collected along sections cut at an angle to the crystallographic axes are stretched (Fig. 2c, d). While obviously there is no solution for the lack of information for the inner part of a crystal, a stretching factor can be applied to the profiles to account for mis-orientation between pairs of crystals (Wallace and Bergantz 2004, 2005).

Zoning type and analytical spatial resolution

In general, zoning of magmatic crystals can either reflect bulk changes in composition and temperature within a magmatic system or may reflect temporary and local changes in the residual melt around a crystal (Ginibre et al. 2002b; Wallace and Bergantz 2005; Davidson et al. 2007; Rubin et al. 2017). Consequently, the frequency of the signal that is being correlated is related to the distance over which significant chemical variations occur, and thus will depend on the mineral phase that is being analysed and the magmatic processes that are investigated (Pearce and Kolisnik 1990). Hence, when shortening a profile, the minimum stretching factor will be limited by the ability to resolve the zoning feature of interest. For example, if an original profile resolves a feature with 8 pixels (or points) and 4 are required to resolve it, the shortening is limited to 50%. However, this will also depend upon the sampling rate (pixels or analysis point per micron/millimeter), which may vary between images. Furthermore, the minimum window over which profiles can be correlated, and the wavelength of any filters that are used to smooth crystal profiles will be limited using a similar logic. In general, at last one full cycle of chemical variation is required for the meaningful correlation of profiles.

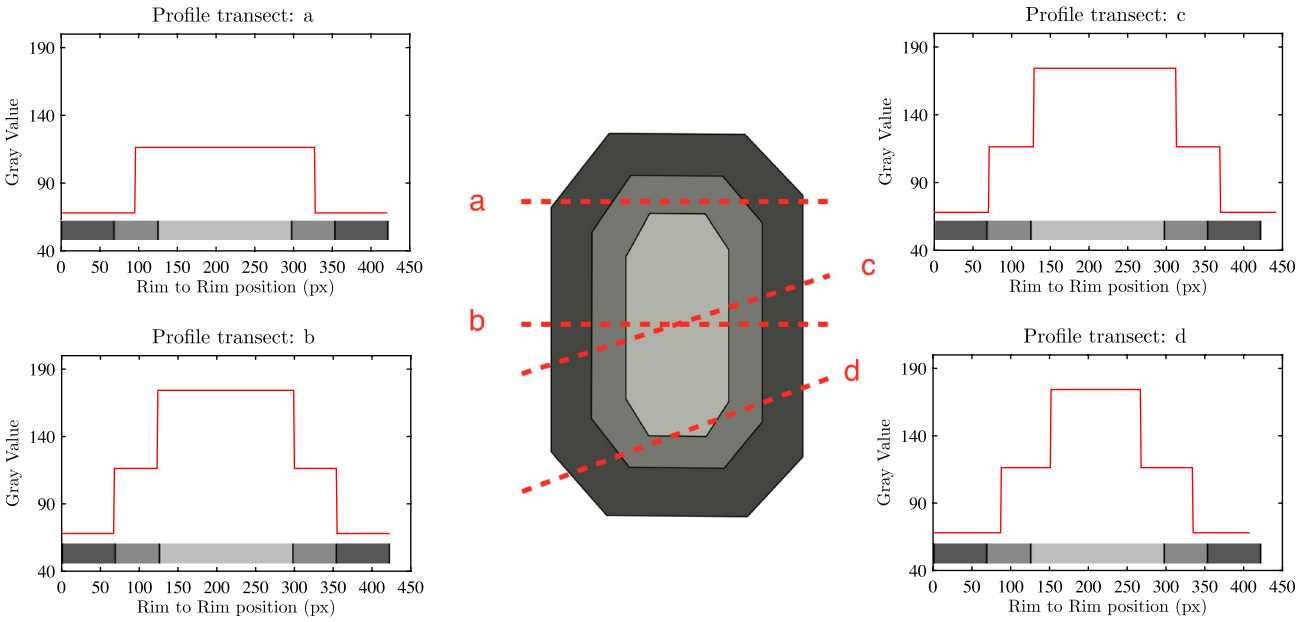

Fig. 2 Profiles for four different transects/sections from an individual crystal. At the bottom of the different graphs representing the profiles is a slice of the crystal corresponding to an ideal transect for comparison. Top, left: transect *a* is off-centered but with the correct orientation. As expected, the grey value is constant for almost the entire length of the transect. There is an information loss in the core region compared to the ideal profile *b*. Bottom, left: transect *b* is ideal, both centered and correctly oriented. Profile *b* is in a perfect accordance with the coloured band at the bottom. Top, right: transect *c* is centered but misoriented. Profile *c* is similar to the profile *b* but stretched. Bottom, right: transect *d* is off-centered and misoriented. The second part of profile *d* is the same as in profile *c* because they have the same orientation

Significance of correlation

Stretching, smoothing and cutting of profiles may result in artificial correlation (Wallace and Bergantz 2004). Consequently, a significance level (SL) is required for the correlation coefficient, above which correlation is statistically significant. Correlations between zoning profiles or segments of zoning profiles are considered statistically significant when the correlation coefficient is above a significance level (SL_{TRUE}) that rejects the null hypothesis that the data series are unrelated. This significance level will depend on the variance, frequency distribution, sampling frequency, and spatial distribution of zoning features in the mineral system of interest (Wallace and Bergantz 2004). Hence, the significance level will be determined by the geological application and mineral phase that is being correlated. Consequently, no general formula or a priori value for the significance level exists. One approach is to, therefore, estimate SL_{TRUE} using a Monte Carlo method to create artificial zoning profiles generated by an algorithm that could be adapted to the aforementioned parameters (Wallace and Bergantz 2004).

To check the validity of SL_{TRUE}, an upper bound for the significance level can be estimated by correlating profiles from the same crystal (Self Correlation Coefficient; SL_{SCC}). If the value of SL_{SCC} is lower than SL_{TRUE}, then this suggests that the cross correlation procedure cannot identify correlation within an individual crystal, and so any correlations that are greater than SL_{TRUE} cannot be distinguished from random crystal populations. However, when $SL_{SCC} > SL_{TRUE}$, correlations can be judged statistically significant, and the greater the difference between SL_{TRUE} and SL_{SCC} the more likely it is that populations may be identified by profile comparison. Due to anisotropy in the growth of crystals, the value of SL_{TRUE} is unlikely to be 1, and depending on how the anisotropy varies between crystals, single or multiple crystals may be required to estimate SL_{SCC}.

Cross-correlation algorithm

The correlation of two profiles occurs between a reference profile, which is not stretched, and a variable profile, which is stretched. Previous approaches have focused on the identification of a decorrelation point, where the variable profile is interpolated according to a series of stretching factors, and an optimum stretching factor is chosen as the one that maximises the distance from the rim of the reference profile for which the correlation is significant (Fig. 1b; Wallace and Bergantz 2004, 2005). The cross-correlation algorithm developed below, however, is based on maximising the correlation between two profiles using an expanding window approach, rather than the identification of a decorrelation point. By maximising the correlation between two profiles, the correlation may not be complete from the interior of two profiles to the outer rim, but it allows correlation to be identified within the interior of two crystals, even if it does not extend to the outer portions of the profiles (Fig. 1c–e). To enable this procedure, rather than compare two profiles by changing the stretching factor, we repeatedly compare the reference profile and variable profile across a series of windows. Furthermore, using the concept of cross correlation, we introduce the idea that the correlation of two profiles may be optimised by displacing or shifting one of the profiles (Fig. 1e). The principles of the algorithm are described in more detail below. The algorithm is written using Fortran 90 and run using the R statistical software interface (R Development Core Team 2008), with the code provided in the supplementary material.

Algorithm description

The algorithm works by comparing (in a loop system) two profiles at once: a reference profile of length l_f, represented by the vector $\mathbf{f} \in \mathbb{R}^{l_f}$, with a variable profile of length l_g, represented by the vector $\mathbf{g} \in \mathbb{R}^{l_g}$. For each profile the user chooses a minimal length window (α) and a maximal window length (β) beyond which the program stops searching for correlation. The value of α will depend on the distance between the zoning features of interest (as discussed in "Zoning type and analytical spatial resolution") and β is chosen to improve computational efficiency. These limits are selected as a percentage of the total length of the profiles, and thus α and β do not depend on the specific lengths l_f and l_g. However, the lower bounds ($\alpha\ l_f$;$\alpha\ l_g$) are rounded down to the nearest whole integer, and the upper bounds ($\beta\ l_f$;$\beta\ l_g$) rounded up to the nearest whole integer. For the lower bounds, they must be at least larger than the minimum distance observed between the zoning features of interest.

The algorithm consists of two parts: (1) calculation of the stretching factor, and linear interpolation of the variable profile; and (2) correlation of the reference profile and interpolated variable profile. The procedure is then repeated for each combination of $\mathbf{v}_j = \mathbf{f}(1 : j)$ and $\mathbf{w}_k = \mathbf{g}(1 : k)$ where $j \in [\alpha\ l_f, \beta\ l_f]$ and $k \in [\alpha\ l_g, \beta\ l_g]$.

Stretching factor

The stretching factor (s_g) for the variable profile is calculated as the ratio between the length of the reference profile and variable profile,

$$S(k, j) = \frac{j}{k},$$

for all possible combinations of the reference and variable profiles. The stretching of $\mathbf{w}_k$ is performed by a linear interpolation and we denote the interpolated vector $\mathbf{g}' \in \mathbb{R}^j$, which represents the stretched variable profile $\mathbf{w}_k^j = \mathbf{g}'(1 : j)$.

A wide range of stretching factors can be applied, although the maximum limit should have a physical or geological meaning. Since the stretching factors are used to compensate for the crystal orientation, the real maximal value of the stretching factor will depend on the geometry of each particular kind of crystal. For example, as the aspect ratio between the longest and shortest axes of a crystal increases, so must the range of possible stretching factors.

Correlation coefficient

For each reference and variable profile, the correlation coefficient is computed between the stretched variable profile $\mathbf{w}_k^j = \mathbf{g}'(1 : j)$ and the reference profile $\mathbf{v}_j := \mathbf{f}(1 : j)$, where the stretched window $\mathbf{w}_k^j$ has the same length as the vector $\mathbf{v}_j$. Hence, the correlation is always calculated from the crystal rim of the reference profile and the two profiles have an identical length. We use the sample Pearson correlation coefficient (Pearson 1894) and estimate it using the following equation,

$$r(\mathbf{f},\mathbf{g}') := \frac{\sum_{i=1}^{j}((f_i-\bar{f})(g'_i-\bar{g}'))}{\sqrt{\sum_{i=1}^{j}(f_i-\bar{f})^2\sum_{i=1}^{j}(g'_i-\bar{g}')^2}},$$

where f_i and g'_i provide the value of the reference profile and stretched variable profiles at the ith point for $i \in (1, j)$, $\bar{f} := \frac{1}{j}\sum_{i=1}^{j} f_i$ and $\bar{g}' := \frac{1}{j}\sum_{i=1}^{j} g'_i$ are the mean values of vector **f** and **g'** and j is the length of the reference profile vector to be correlated.

A two-dimensional grid is constructed, R(k, j), which is filled by correlation coefficients for each combination of the reference (**f**) and original variable (**g**) profiles. The algorithm searches for the global maximum of the correlation coefficient $\rho(\mathbf{g},\mathbf{f}) := \max_{k,j}(\mathrm{R}_{k,j})$ within this grid, which will define the optimum stretching factor and length of the reference profile.

In the scenario where the optimum correlation may originate at the rim of one crystal but from within the interior of the other crystal (Fig. 1e), a shift (τ) can be applied to the algorithm. The maximum possible value of τ can be set by the user, but the algorithm will only cross correlate signals when the displacement is less than half the length of the sampled reference profile $\frac{j}{2}$. The displacement is performed on the stretched variable profile $\mathbf{g}'$ and can be either negative or positive with respect to the reference profile (i.e. symmetrical around zero). For each value of τ, sections of the vectors **f** and $\mathbf{g}'$ will not have corresponding values on the alternate profiles and so the correlation cannot be calculated. Hence, each of the vectors are cut accordingly. When $\tau > 0$,

$$\mathbf{v}_j := \mathbf{f}(1+\tau : j),$$
$$\mathbf{w}_k^j = \mathbf{g}'(1 : j-\tau),$$

and when $\tau < 0$,

$$\mathbf{v}_j := \mathbf{f}(1 : j+\tau),$$
$$\mathbf{w}_k^j = \mathbf{g}'(1-\tau : j).$$

Mathematical framework

To ensure the consistency and reciprocity of our procedure we examine the symmetry of the results (Wallace and Bergantz 2005). This means that comparing the reference profile **f** with the variable profile **g** will give the same result as using the vector **g** as reference profile and comparing it with vector **f** as the variable profile. To test for symmetry, we run the cross-correlation algorithm using 500 pairs of synthetic profiles, with the results showing that the respective stretching factors are symmetric (Fig. 3) and correlation coefficients identical. This observation is proved rigorously in the supplementing material by a theorem at the level of continuous profile functions, instead of vectors, to exclude approximation errors due to the linear interpolation when the variable profile is stretched. The theorem is based on showing that every window length within the signal is reachable starting from every window length within the other signal. Importantly, to ensure the symmetry in the results we have to set a minimum stretching factor that is the inverse of the maximum stretching factor (i.e. $s_{\min}(k,j) = \frac{1}{s_{\max}(k,j)}$), and so we are only interested in the portion of the correlation matrix where $s_{\min} \le \frac{j}{k} \le s_{\max}$. Thus, in the discrete case the theorem will be symmetrical for the portion of the matrix R(k, j) that satisfies this stretching limit criterion.

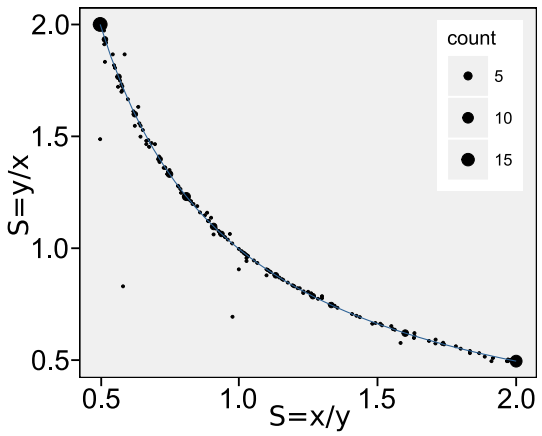

Fig. 3 Pairwise stretching factor for the maximum correlation coefficient for 500 pairs of synthetic profiles, exhibiting the symmetry of the algorithm. The size of the points is relative to the number of pairs that plot for the specific coordinate. The generation of synthetic profiles is based on the method presented by Wallace and Bergantz (2004)

General application

In "Application: Kilgore Tuff" we have chosen a series of non-ideal profiles from zircons to test the performance of the model. However, when applying the algorithm profiles should be chosen to enable optimum conditions to identify significant correlation, by limiting the freedom within the model parameters as much as possible. The maximum correlation coefficient between two profiles will depend on which values are chosen for the model parameters (e.g. maximum stretching factor), which themselves will depend on how the crystals are prepared (e.g. grains mount vs. thin section) and whether profiles are drawn systematically perpendicular to a specific face. As the limits of the model parameters are relaxed (e.g. choosing a wider range of stretching factors) then the maximum correlation coefficient calculated between two profiles will increase, the value of SL_{TRUE} will increase and the ability to distinguish significant correlation will become trickier. Therefore, profiles drawn from the same crystallographic axis from crystals mounted and cut through the centre puts the user in the best possible conditions to apply the algorithm, because the model parameters can be constrained and thus the capacity to distinguish significant correlation increases. Profiles should also be chosen to avoid inclusions and cracks. The methodology can be applied to crystals even from thin sections, but only if the model assumptions (stretching, smoothing, shifting of profiles) are justified and the significance level for the correlation correctly estimated. Combined with other detailed petrography and geochemical analyses, this cross correlation algorithm provides a framework to quantify the amount of crystals sharing some common history.

Application: Kilgore Tuff

We apply the cross-correlation algorithm to zircons from two deposits that are located > 80 km apart at the margins of a large-caldera volcanic system, the Heise Volcanic Field (Idaho, USA). To test whether these two outcrops were produced by the same eruption we correlate grayscale cathodoluminescent profiles of zircons, which provides a bulk compositional record of the thermal and chemical evolution of magma reservoirs, with the additional advantage of being resistant to alteration (Hoskin and Schaltegger 2003). How zircons respond to changes in melt chemistry or temperature is not well constrained, however, if significant correlation is found between zircon zonation patterns from the two outcrops, this would imply that the zoning was formed by a sequence of similar processes, and therefore, that the outcrops are from the same eruption.

Geological setting

The Heise Volcanic Field (HVF) is located in eastern Idaho at the SW limit of the Yellowstone Volcanic Field (Fig. 4a). HVF hosts the volcanic products of the last complete caldera cycle in the Snake River Plain (SRP) preceding the Yellowstone caldera cycle (~ 6.6–4.0 Ma: Ellis et al. 2013; Wotzlaw et al. 2014. The youngest (~ 4.45 Ma) most voluminous eruption (1800 km^3) of the HVF produced the Kilgore Tuff (1800 km^3), which is presently well-exposed over an area of about 20,000 km^2 (Morgan and McIntosh 2005; Watts et al. 2011). The Kilgore Tuff was erupted from two main vents to the north and south of the present-day caldera (Fig. 4a; Morgan and McIntosh 2005). The two deposits we sample are geochemically identical and are identified as both belonging to the Kilgore Tuff (Bindeman et al. 2007; Watts et al. 2011), which is important to test our algorithm.

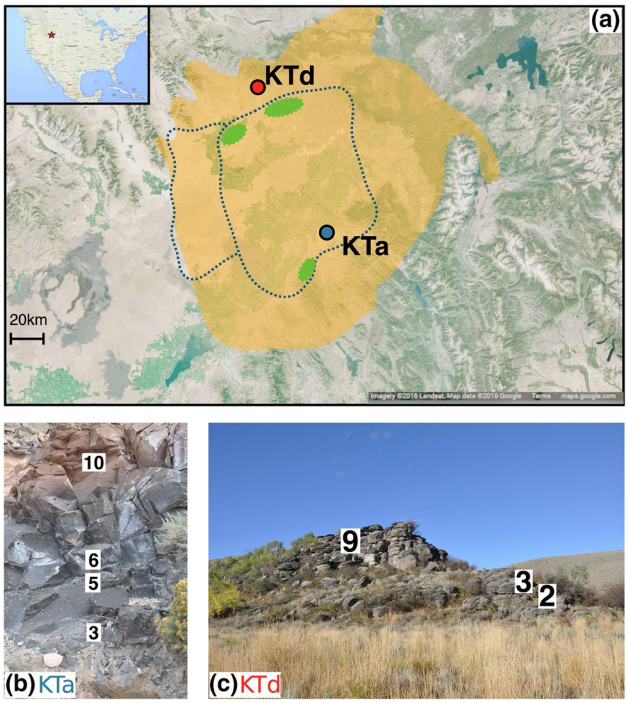

Fig. 4 **a** Map of the Heise Volcanic Field modified from Wotzlaw et al. (2014). Blue dashed line and orange bounded area represent the estimated caldera boundaries and the inferred extent of the Kilgore Tuff, respectively. The blue and red dots are the locations of the two sampled outcrops, *KTa* and *KTd*; **b** sample locations in the *KTa* outcrop, located in the southeastern part of the caldera rim (N 43°31.979, W 111°41.995, 1683 m); **c** sample locations in the *KTd* outcrop, located in the northwestern part of the caldera rim (N 44°21.577, W 112°09.910, 1860 m). Numbers in **b** and **c** represent the vertical location and reference number of samples used in this study

Data acquisition

Samples were collected from welded ignimbrite deposits at two locations on the southeast (*KTa*) and northwest (*KTd*) rim of the Heise caldera (Fig. 4). Zircons were separated from 4 different vertical levels in the *KTa* deposit (Fig. 4b) and 3 different layers in the *KTd* deposit (Fig. 4c) to ensure the ignimbrite deposits were sampled along the stratigraphy. Zircons were separated using heavy liquids, picked, mounted, and polished to expose their interior. Zircons were then imaged at the University of Geneva using a Jeol JSM 7001F Scanning Electron Microscope (SEM) equipped with a cathodoluminescence (CL) detector. Grayscale images of each zircon were generated by the JED 2300 software to characterise temporal variations in the local melt conditions in which the zircons grew (Nasdala et al. 2003). Importantly, the grayscale values of each zircon are not standardised for the chemical content of each zircon. However, because we are interested in comparing the pattern of zonation, this does not affect the results of the statistical analysis.We used a single 1D profile for each crystal, and there was no systematic approach to take the profile. This is not the optimum scenario to correlate profiles from zircon crystals, as discussed in "General application", because we have to choose a wider range of stretching factors, but it enabled us to test the performance of the algorithm.

For each vertical level we imaged ~ 25 zircons. Of these images we selected between 12 and 14 zircons for analysis with the cross-correlation algorithm, removing crystals with broken edges or complex sector zoning. Rim to core profiles were then drawn using the software *ImageJ* (Schneider et al. 2012). Consequently, the cross-correlation method was performed on profiles from a total of 48 zircons from the *KTa* deposit and a total of 41 zircons from the *KTd* deposit.

Cross-correlation parameters

We assume crystals that have a common pattern of zonation experience the same variation of conditions over a prolonged period of time, which implies there can be no displacement ($\tau = 0$) in the correlation of crystal profiles. We set $\alpha = 0.3$, $\beta = 1$, and $S_{MAX} = 3$ meaning $S_{MIN} = {}^1/_3$. We choose $\alpha = 0.3$ to ensure that the shortest profile we analyse is larger than the minimum distance between the zoning features that records changes in bulk melt composition or temperature. Given the shortest profile we analyse is composed of 39 pixels, where each pixel is ~ 0.9 µm, we assume that the shortest zoning feature we consider in our analysis is > 11 pixels, which is ~ 10 µm. Given that oscillatory zoning which records the competition between element diffusion and growth (Pearson 1894) has a lower amplitude than oscillatory zoning that records bulk changes, we can ignore its effect on the results of the correlation coefficient. Finally, we apply a light smoothing filter processed in *ImageJ*, to remove analytical noise, but to ensure the zoning features of interest remain.

Significance level

As discussed in "Significance of correlation", the correlation of mineral profiles requires an assessment of the minimum coefficient, or significance-level (SL_{TRUE}), above which any correlation is significant. To quantify the SL_{TRUE} value for this application we use a Monte Carlo approach (Supplementary material: Wallace and Bergantz 2004), in which 500 random profiles that have a similar length zonation frequency to the profiles in our analysis are generated. We run the correlation algorithm using the same value of α and S_{max} in the analysis of all 89 zircons as this will determine the range of correlation coefficients. The value of SL_{TRUE} is determined as the 95^{th} percentile of the correlation coefficients, which equals 0.90 (Fig. 5). However, the upper value of the maximum correlation coefficient (r_{max}) appears to depend on the percentage of the reference profile that is correlated (Fig. 5), suggesting that SL_{TRUE} may actually be lower when the maximum correlation is observed over longer portions of the two profiles, and a larger value of α is chosen.

To test the validity of SL_{TRUE} we correlate four profiles from an individual crystal and use the lowest value of the correlation coefficient between the four profiles to estimate SL_{SCC} (Fig. 6a, b). We choose just one crystal to estimate SL_{SCC}, which reflects the general level of anisotropy within zircons in the Kilgore Tuff. Once again, we run the

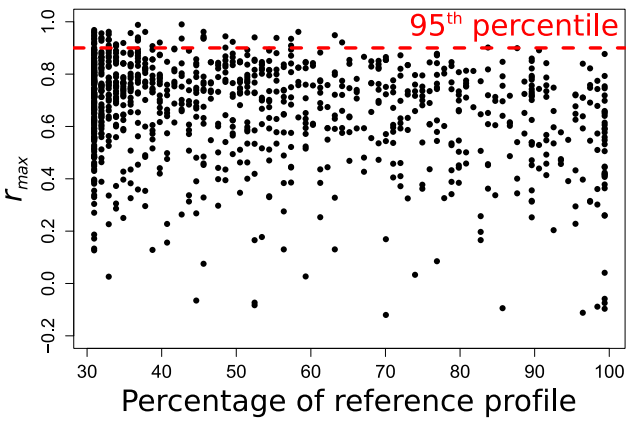

Fig. 5 Maximum values of the correlation coefficient from correlation of 500 pairwise profiles using a Monte Carlo approach. The 95th percentile is highlighted by the red dashed line. The maximum coefficient is plotted against the percentage of the reference profile that is correlated, with the upper bound in the coefficient decreasing as the percentage of the reference profile increases. For profiles that are identical in length, the percentage of the profile would be equivalent to the window size

correlation algorithm using the same value of α and S_{max} used for the analysis of the entire set of 89 zircons.

Whilst the general shape of each of the four profiles is similar, due to the resolution of the image and crystal anisotropy (i.e. variations in growth and dissolution), the algorithm is not able to perfectly (i.e. coefficient of 1) correlate all profiles, and for some pairwise combinations the maximum correlation coefficient does not identify correlation over the whole of the profile (Fig. 6c, d; Table 1). The lowest value of the correlation coefficient is 0.94 (Table 1), which we take to be our SL_{SCC}. Importantly, the value of $SL_{SCC} > SL_{TRUE}$, indicating that correlations between the 89 real profiles that have a coefficient larger than or equal to 0.90 are significant.

Table 1 Correlation coefficients for the four profiles in Fig. 6. As expected, the correlation between profile pairs is symmetrical

	1	2	3	4
1	1	0.95	0.95	0.95
2	0.95	1	0.96	0.94
3	0.95	0.96	1	0.96
4	0.95	0.94	0.96	1

Results and discussion

The maximum correlation coefficient is based on correlating a reference profile and stretched variable profile from the rim (Fig. 7). In some cases two profiles are correlated from the rim (Fig. 7b) whereas in other cases it appears that the correlation occurs in the interior of the crystals (Fig. 7c). However, the algorithm still identifies the correlation between these two profiles as significant. Currently, the identification of what parts of the profiles are correlated is performed qualitatively, based on the correlation coefficient for different lengths of the stretched variable profile (Fig. 7b, c). Currently it is not feasible to compare all pairwise combination, but future development of the algorithm would be to automatically identify which part of the profiles are correlated and to ultimately characterise different families of crystals (e.g. Wallace and Bergantz 2005) to aid magmatic interpretations and guide further research.

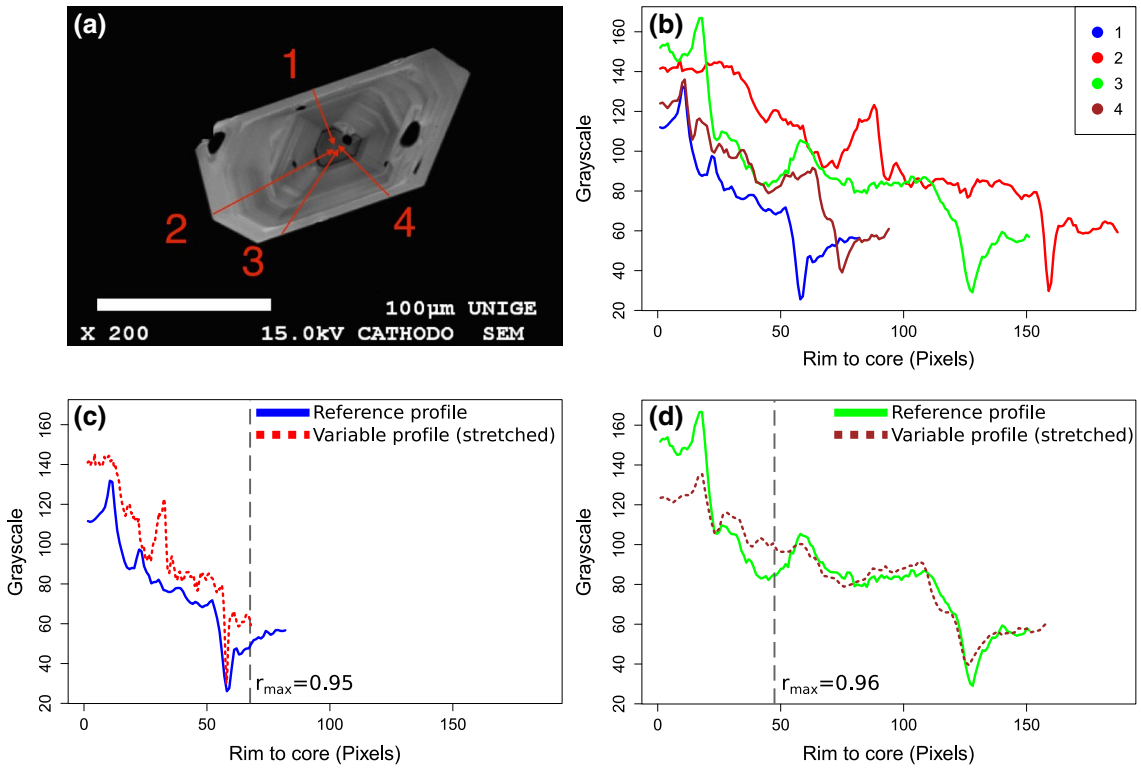

Fig. 6 Correlation of four profiles from an individual zircon. **a** CL image with the four profiles identified; **b** Rim to core grayscale values for the four profiles; **c**, **d** Comparison of rim to core grayscale values between a reference profile, and the corresponding stretched variable profile that produces the maximum correlation coefficient. Profile colours are the same as in Fig. 6b and the dashed vertical line indicates the length of the reference profile window at which the maximum correlation coefficient occurs. Due to irregularities in the growth of the crystal on each of the different faces (e.g. the irregular peak observed on profile 2), no two profiles are perfectly correlated. The absolute value of each of the profiles is not identical, due to anisotropy in the CL image, but does not affect the results of the cross correlation

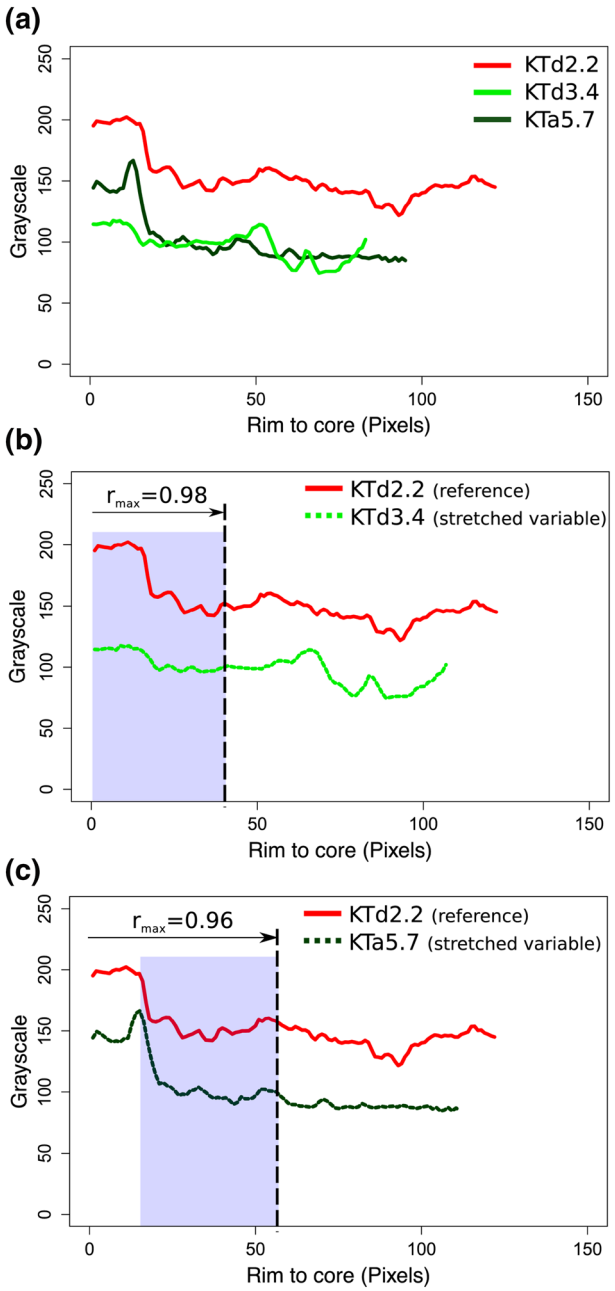

Fig. 7 Correlation of three zircon profiles: **a** original profiles; **b** comparison between two zircons from KTd; **c** comparison between a zircon in KTd and a zircon in KTa. In **b** and **c**, the maximum correlation is identified. The purple shaded area represents the lengths of the stretched variable profile that have a correlation coefficient that is greater than the significance level. This identifies that correlation may extend from the rim (**b**) or may only occur in the interior of two crystals (**c**)

Zircons are correlated within each deposit, and between both deposits (Table 2). The correlation of zircons between the two outcrops indicates zircons in both deposits experienced similar magmatic processes, which is in agreement with published data showing that both outcrops belong to the Kilgore Tuff (Bindeman et al. 2007; Watts et al. 2011). The level of correlation within each deposit is high, with 46 of 48 (~96%) zircons in KTa correlating with at least one other zircon in the same deposit, and 37 of 41 (~90%) zircons in KTd correlating with at least one other zircon in the same deposit. Between the deposits, however, the level of correlation varies, with 45 of 48 (~94%) zircons in KTa correlating with at least one other zircon in KTd, and only 33 of 41 (~80%) zircons in KTd correlating with at least one other zircon in KTa. This suggests that the patterns of zonation and that are observed in zircons in KTa are also observed in zircons in KTd, but not all the patterns of zonation that are observed in zircons in KTd are observed in zircons in KTa. Specifically, it suggests there are zircon populations within KTd that are distinct with respect to the patterns of zonation observed in KTa. If we assume that these different zonation patterns represent different magmatic histories it suggests that a proportion of the zircons in KTd record a magmatic history that is not registered by zircons in KTa. However, such an interpretation requires further geochemical analyses and examination of all pairwise correlations to understand whether profiles are correlated from the rim or within the interior of the crystal (e.g. Fig. 7).

Table 2 Number of zircons that are significantly correlated with at least one other zircon in either the *KTa* or *KTd* deposits

	KTa (48 total)	*KTd* (41 total)
KTa	46/48	33/41
KTd	45/48	37/41

Conclusions

We have developed an algorithm that can compare and identify similar zoning patterns in different portions of magmatic crystals. This algorithm has been applied to the Kilgore Tuff, where zircons in two deposits over 80km apart have been correlated. Consequently, where volcanic deposits are found over a large spatial scale, the cross-correlation of crystal profiles provides an opportunity to statistically determine whether they are related (e.g. from the same eruption). For example, this could be at different sides of a large volcanic caldera, or in dispersed tephra in lake or deep sea sediments.

Thus far we have used the global maximum coefficient to identify correlation between two profiles, based on satisfying a significant level of correlation. The choice of significance level will be crucial when interpreting the results to understand magmatic processes in more detail, and to ultimately constrain the level of correlation between two crystal profiles, so that different families of crystals can be identified. Currently, the location of correlation is based on

an expanding window approach, but examining the results of algorithm has allowed the qualitative identification of crystals that are correlated in the interior, but not at the rim (Fig. 7c). In the future, the location of correlation within crystal profiles may be quantitatively defined by identifying local maxima within the correlation matrix produced by the algorithm. This would allow correlation from the rim to be better constrained, and the identification of correlation in the innermost portions of crystals that may be too small to locate using a global maxima approach. Such an approach may also overcome the issue of the dependence between the maximum correlation coefficient and the percentage of the reference profile that is correlated (Fig. 5).

This technique offers the possibility to identify similarities and differences in crystal zoning patterns rapidly and objectively using an automated algorithm. Hence, when the cost of image production is low the algorithm provides a cost-effective method to interrogate large numbers of crystals. In turn this provides an excellent basis on which to select minerals for more time consuming and expensive geochemical analyses or more intensive age dating techniques.

Acknowledgements This study is dedicated to the memory of Line, a beautiful mind and friend that we will always carry in our hearts. This study was supported by the SNSF Grant 200021_162503 and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement no. 677493). We would to thank Emmanuelle Ricchi for providing the cathodoluminescence images of zircons. We would also like to thank two anonymous reviewers for their comments, which helped improve the manuscript.

References

Bindeman IN, Watts KE, Schmitt AK, Morgan LA, Shanks PWC (2007) Voluminous low δ18O magmas in the late miocene heise volcanic field, idaho: implications for the fate of yellowstone hotspot calderas. Geology 35(11):1019. https://doi.org/10.1130/G24141A.1

Cashman K, Blundy J (2013) Petrological cannibalism: the chemical and textural consequences of incremental magma body growth. Contrib Mineral Petrol 166(3):703729. https://doi.org/10.1007/s00410-013-0895-0

Cheng L, Costa F, Carniel R (2017) Unraveling the presence of multiple plagioclase populations and identification of representative two-dimensional sections using a statistical and numerical approach. Am Mineral 102(9):1894–1905. https://doi.org/10.2138/am-2017-5929CCBYNCND

Chu KH, Qi J, Yu ZG, Anh V (2004) Origin and phylogeny of chloroplasts revealed by a simple correlation analysis of complete genomes. Mol Biol Evol. https://doi.org/10.1093/molbev/msh002

Costa F, Coogan LA, Chakraborty S (2009) The time scales of magma mixing and mingling involving primitive melts and melt-mush interaction at mid-ocean ridges. Contrib Miner Petrol 159:371387. https://doi.org/10.1007/s00410-009-0432-3

Davidson JP, Morgan DJ, Charlier BLA, Harlou R, Hora JM (2007) Microsampling and isotopic analysis of igneous rocks: Implications for the study of magmatic systems. Annu Rev Earth Planet Sci 35:273311. https://doi.org/10.1146/annurev.earth.35.031306.140211

Ellis BS, Wolff JA, Boroughs S, Mark DF, Starkel WA, Bonnichsen B (2013) Rhyolitic volcanism of the central Snake River Plain: a review. Bull Volcanol. https://doi.org/10.1007/s00445-013-0745-y

Ellis BS, Bachmann O, Wolff JA (2014) Cumulate fragments in silicic ignimbrites: the case of the Snake River Plain. Geol. https://doi.org/10.1130/G35399.1

Galton F (1888) Correlations and their measurement, chiefly from anthropometric data. Proc R Soc Lond 45:135–145

Ginibre C, Kronz A, Wörner G (2002a) High-resolution quantitative imaging of plagioclase composition using accumulated backscattered electron images: new constraints on oscillatory zoning. Contrib Miner Petrol 142(4):436448. https://doi.org/10.1007/s004100100

Ginibre C, Wörner G, Kronz A (2002b) Minor- and trace-element zoning in plagioclase: implications for magma chamber processes at Parinacota volcano, northern Chile. Contrib Miner Petrol 143:300315. https://doi.org/10.1007/s00410-002-0351-z

Ginibre C, Wörner G, Kronz A (2007) Crystal zoning as an archive for magma evolution. Elements 3(4):261266. https://doi.org/10.2113/gselements.3.4.261

Gudmundsson A (2012) Magma chambers : formation, local stresses, excess pressures, and compartments. J Volcanol Geotherm Res 237–238:19–41. https://doi.org/10.1016/j.jvolgeores.2012.05.015

Hartung E, Caricchi L, Floess D, Wallis S, Harayama S, Kouzmanov K, Chiaradia M (2017) Evidence for residual melt extraction in the takidani pluton, central japan. J Petrol 58(4):763–788. https://doi.org/10.1093/petrology/egx033

Hoskin PWO, Schaltegger U (2003) The composition of zircon and igneous and metamorphic petrogenesis. Rev Miner Geochem 1:2762. https://doi.org/10.2113/0530027

Humphreys MCS, Edmonds M, Plail M, Barclay J, Parkes D, Christopher T (2012) A new method to quantify the real supply of mafic components to a hybrid andesite. Contrib Miner Petrol 165:191215. https://doi.org/10.1007/s00410-012-0805-x

Miller DS, Smith RB (1999) P and S velocity structure of the Yellowstone volcanic field from local earthquake and controlled-source tomography. J Geophys Res. https://doi.org/10.1029/1998JB900095

Morgan LA, McIntosh WC (2005) Timing and development of the Heise volcanic field, Snake River Plain, Idaho, western USA. Bull Geol Soc Am 117(3–4):288–306. https://doi.org/10.1130/B25519.1

Myers ML, Wallace PJ, Wilson CJN, Morter BK, Swallow EJ (2016) Prolonged ascent and episodic venting of discrete magma batches at the onset of the Huckleberry Ridge supereruption, Yellowstone. Earth Planet Sci Lett. https://doi.org/10.1016/j.epsl.2016.07.023

Nasdala L, Zhang M, Kempe U, Panczer G, Gaft M, Andrut M, Ploetze M (2003) Spectroscopic methods applied to zircon. Rev Miner Geochem 53:427463. https://doi.org/10.2113/0530427

Pearce T, Kolisnik A (1990) Observations of plagioclase zoning using interference imaging. Earth Sci Rev 29:9–26. https://doi.org/10.1016/0012-8252(0)90024-P

Pearson K (1894) Contributions to the mathematical theory of evolution. Philos Trans R Soc Lond 185:71–110

R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. ISBN 3-900051-07-0

Rubin AE, Cooper KM, Till CB, Kent AJR, Costa F, Bose M, Gravley D, Deering C, Cole J (2017) Rapid cooling and cold storage in a silicic magma reservoir recorded in individual crystals. Science 356(6343):1154–1156. https://doi.org/10.1126/science.aam8720

Schneider C, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9(7):671–675. https://doi.org/10.1038/nmeth.2089

Wallace GS, Bergantz GW (2002) Wavelet-based correlation (WBC) of zoned crystal populations and magma mixing. Earth Planet Sci Lett 202:133145. https://doi.org/10.1016/S0012-821X(02)00762-8

Wallace GS, Bergantz GW (2004) Constraints on mingling of crystal populations from off-center zoning profiles: a statistical approach. Am Miner 89:6473. https://doi.org/10.2138/am-2004-0109

Wallace GS, Bergantz GW (2005) Reconciling heterogeneity in crystal zoning data: an application of shared characteristic diagrams at Chaos Crags, Lassen Volcanic Center, California. Contrib Miner Petrol 149:98112. https://doi.org/10.1007/s00410-004-0639-2

Watts KE, Bindeman IN, Schmitt AK (2011) Large-volume rhyolite genesis in caldera complexes of the Snake River Plain: insights from the Kilgore Tuff of the Heise volcanic field, Idaho, with comparison to Yellowstone and Bruneau-Jarbidge rhyolites. J Petrol 52:857890. https://doi.org/10.1093/petrology/egr005

Wotzlaw JF, Bindeman IN, Watts KE, Schmitt AK, Caricchi L, Schaltegger U (2014) Linking rapid magma reservoir assembly and eruption trigger mechanisms at evolved yellowstone-type supervolcanoes. Geology 42(9):807. https://doi.org/10.1130/G35979.1

Wotzlaw JF, Bindeman IN, Stern RA, D'Abzac FX, Schaltegger U (2015) Rapid heterogeneous assembly of multiple magma reservoirs prior to yellowstone supereruptions. Sci Rep 5(14):026. https://doi.org/10.1130/G35979.1

Yarlagadda RKR (2010) Analog and digital signals and systems. Springer Science + Business. Media. https://doi.org/10.1007/978-1-4419-0034-0