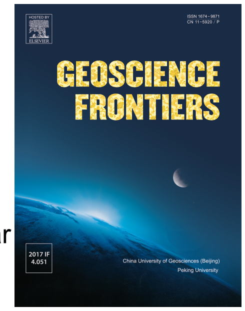


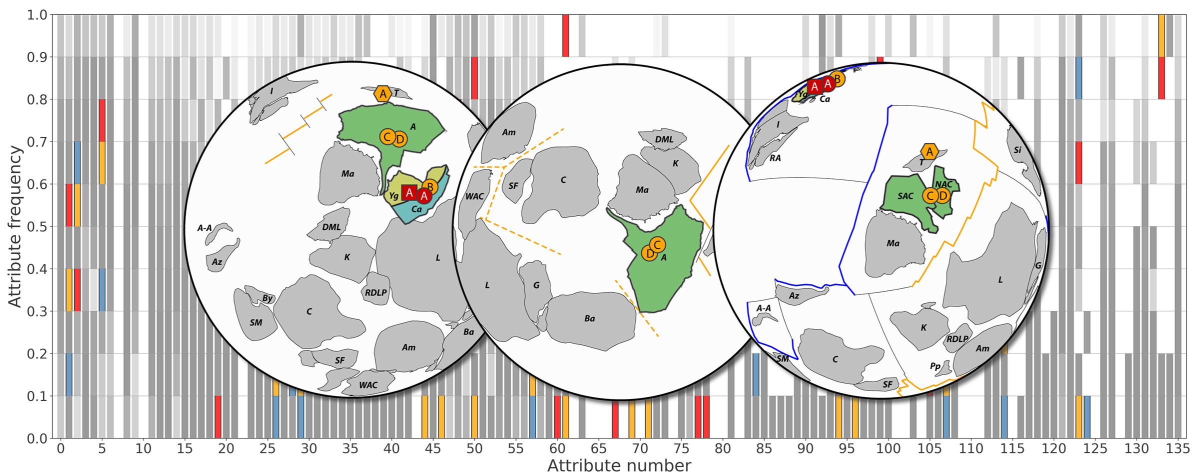
Decoding Earth's plate tectonic history using sparse geochemical data

Michael. G. Tetley, Zheng-Xiang Li, Kara J. Matthews, Simon. E. Williams, R. Dietmar Müller

PII: S1674-9871(19)30090-8
DOI: https://doi.org/10.1016/j.gsf.2019.05.002
Reference: GSF 841

To appear in: *Geoscience Frontiers*

Received Date: 10 October 2018
Revised Date: 18 March 2019
Accepted Date: 9 May 2019


Attribute frequency
Attribute number

Decoding Earth's plate tectonic history using sparse geochemical data

Michael. G. Tetley[a, b, c, 1,*], Zheng-Xiang Li[d], Kara J. Matthews[a], Simon. E. Williams[a], R. Dietmar Müller[a]

[a]EarthByte Group, School of Geosciences, University of Sydney, NSW, Australia
[b]Seismological Laboratory, California Institute of Technology, Pasadena, California, USA
[c]Data61, CSIRO, Australian Technology Park, NSW, Australia
[d]Earth Dynamics Research group, The Institute for Geoscience Research (TIGeR), School of Earth and Planetary Sciences, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia

[1]Present address: Laboratoire de Géologie de Lyon, Université Claude Bernard Lyon 1, France

* Corresponding author: michael.tetley@univ-lyon1.fr

Abstract

Accurately mapping plate boundary types and locations through time is essential for understanding the evolution of the plate-mantle system and the exchange of material between the solid Earth and surface environments**.** However, the complexity of the Earth system and the cryptic nature of the geological record make it difficult to discriminate tectonic environments through deep time. Here we present a new method for identifying tectonic paleo-environments on Earth through a data mining approach using global geochemical data. We first fingerprint a variety of present-day tectonic environments utilising up to 136 geochemical data attributes in any available combination. A total of 38301 geochemical analyses from basalts aged from 5–0 Ma together with a well-established plate reconstruction model are used to construct a suite of discriminatory models for the first order tectonic environments of subduction and mid-ocean ridge as distinct from intraplate hotspot oceanic environments, identifying 41, 35, and 39 key discriminatory geochemical attributes, respectively. After training and validation, our model is applied to a global geochemical database of 1547 basalt samples of unknown tectonic origin aged between 1000–410 Ma, a relatively ill-constrained period of Earth’s evolution following the breakup of the Rodinia supercontinent, producing 56 unique global tectonic environment predictions throughout the Neoproterozoic and Early Paleozoic. Predictions are used to discriminate between three alternative published Rodinia configuration models, identifying the model demonstrating the closest spatio-temporal consistency with the basalt record, and emphasizing the importance of integrating geochemical data into plate reconstructions. Our approach offers an extensible framework for constructing full-plate, deep-time reconstructions capable of assimilating a broad range of geochemical and geological observations, enabling next generation Earth system models.

Keywords:

plate tectonics, geochemistry, geodynamics, supercontinents, rodinia, big data

1. Introduction and background

The global continental configurations since the time of Pangea are relatively well established (Schettino and Scotese, 2005; Torsvik et al., 2008; Stampfli et al., 2013), however continental motions in isolation offer limited insight into the complete tectonic system in operation through deep time. The availability of data describing global seafloor spreading histories has driven the development of self-consistent kinematic reconstructions with continuous plate boundaries, together providing the clearest window into Earth’s tectonic history to date (Gurnis et al., 2012; Seton et al., 2012; Müller et al., 2016). Finding ways to apply this ‘full-plate’ philosophy to periods predating the present-day seafloor to reconstruct the Paleozoic and beyond is at the very frontier of current tectonic research, and continues to present a significant challenge to the global tectonics community (Domeier and Torsvik, 2014; Matthews et al., 2016; Merdith et al., 2017a). A fundamental obstacle lies in the difficulty of identifying the nature and paleo-location of dynamic oceanic tectonic environments associated with plate configurations through time, specifically subduction zones, mid-ocean ridges and hotspot interactions of upwelling plumes and the oceanic crust.

Traditionally, geochemical analyses of igneous rocks, commonly basalt due to its abundance and large environment-specific variation in potentially diagnostic element compositions, are used as a discriminatory tool to identify the tectonic environment within which a given sample formed (Pearce and Cann, 1973; Shervais, 1982; Pearce, 2008). The approach involves evaluating the relationships of typically two or three measured element abundances from a given sample set by plotting them overlaying a suite of discriminatory element ratio fields previously established from geochemical analyses of rocks sourced from known tectonic environments. However, outcomes of such an approach are often ambiguous with the statistical probability of solutions difficult to evaluate. Figure 1A shows the tectonic discrimination diagram of Shervais (1982), derived by evaluating the ratio of measured Ti/V from n = ~500 identified samples. These analyses suggest that volcanic rocks with Ti/V ratios between 10 and 20 are likely sourced from subduction (ARC) systems, volcanic rocks with a Ti/V of between 20 and 50 are associated with mid-ocean ridge (MOR) systems, and volcanic rocks with Ti/V ratios of between 50 and 100 are ocean-island (OIB) hotspot related. To explore the robustness and predictive ability of these models with a larger and more diverse dataset, we evaluate n = 4914 global basalt samples aged 0–5 Ma with measurements for both Ti and V extracted from the EarthChem portal (http://www.earthchem.org), with each sample environment geographically classified using the present-day tectonic configuration of Müller et al. (2016). The resulting Ti/V discrimination diagram produces the same three distinct ratio fields as presented by Shervais (1982), trending from ARC to MOR to OIB as Ti abundance increases. However, when derived from the larger data sample the discrimination fields are systematically shifted towards higher Ti/V ratios as the global dataset contains a greater distribution and dynamic range of measured Ti abundances. The resulting modified discrimination fields with upper and lower bounds calculated by 2σ distribution about each population mean suggest Ti/V ratios between 25.9 and 49.5 represent ARC related rocks, Ti/V ratios of 41.4–70.61 represent MOR, and Ti/V ratios of 61.4–166.3 represent OIB environments. Figure 1B shows the same data points with calculated 0.9 and 0.1 probability contours for each environment, indicating that Ti/V ratio diagrams are unlikely to be able to discriminate between tectonic environment for volcanic rocks with Ti values between ~ 7.5 and 16 and V values between ~ 180 and 360 as all data fields exist within this ratio space. It is also apparent the MOR field almost entirely overlaps with the OIB field, suggesting that only MOR samples with the highest V abundances or the lowest Ti abundances have the potential for identification using this method.

Although powerful and useful tools when applied to well-understood sample sets with unambiguous geochemical signatures, methods that rely on directly comparing ratios from only a small number of geochemical sample data types are limited both in their resolution and discriminatory ability. Such limitations have long been recognized, and a number of more successful and sophisticated alternative approaches have been developed. Statistical methods including linear discrimination analysis (LDA) of raw data and LDA with log-ratio transformations of major-element data (Agrawal et al., 2004; Verma et al., 2006; Verma, 2010) are able to predict the tectonic environments of a small set of randomly-chosen samples from each known environment with reported success rates of ~76%–96% and ~83%–97%, respectively (25 and 100 samples were taken from databases of 2732 and 1159 samples respectively). For each study, databases were constructed with predominantly Pliocene basic and ultrabasic rocks of known tectonic affinity, with selection criteria based on each sample requiring 10 pre-prescribed major-element measurements. An alternative has arisen from the development of semi-automated methods, each utilizing a classification tree (CT) exclusionary filter approach (Vermeesch, 2006a, b). This approach requires a large number of pre-prescribed element and isotopic measurements (up to a maximum of 51), and uses a database of 756 samples of known tectonic affinity (sample ages are not considered) to predict the source tectonic environment from basaltic rocks. In addition, this method requires a set of a priori assumptions of optimal measurement abundances (used to make the decisions within the classification tree), which together with the high number of required measurements per sample limit possible applications. The reported successful tectonic environment identification rate is 89% and 84% for trees requiring 51 and 28 measurements, respectively. From the reported results of LDA and CT, it is clear both these methods have high success rates in predicting 'unknown' tectonic environments provided the a priori assumptions are both sufficiently geologically accurate and objective, the 'unknown' sample environments are known before the experiments in order to evaluate success, and the datasets themselves are filtered to contain only data with all the pre-prescribed geochemical values required to perform the selected analyses.

In cases where relatively young, adequately sampled and geologically well understood data are available, these methods demonstrate the best predictive capabilities. However, for investigations into Earth's long term tectonic history, the geological record is rarely sufficiently complete or well understood. Geological data, particularly for times prior to 50 Ma, are both temporally and geographically sparsely sampled (see Supplementary Fig. S2). Of these sparse data, the quantity and type of available geochemical measurements per sample are highly variable, rendering the use of methods with strict input criteria such as Ti/V discriminatory diagrams and statistical methods like LDA and CT, unsuitable for deep time tectonic studies where data are fewest and most spatially and temporally inconsistent. To directly address both this limitation of the available data and the subsequent analysis limitations of most previous approaches, one of the primary aims of this study is the development of a robust method able to tolerate inconsistent data. This approach provides a practical method able to analyze any sample regardless of the number of type of attribute measurements available.

2. New approach to an old problem

Building on this research, we explore the design and application of a new and highly flexible method for identifying the source tectonic environments of sparse basaltic rock data of entirely unknown origin incorporating a significantly wider and variable range of potentially discriminatory attributes without the need for a priori assumptions, prescribed sampled attributes or consistency of measured quantities between samples. In this new approach, we focus on utilizing the structures or 'fingerprints' present within a freely available large basalt geochemistry database to construct data models representative of the first-order tectonic environments ARC, MOR and OIB. Like the fingerprint analogy, each environment model possesses a unique data pattern (Fig. 3), a blueprint that can be used to identify the source tectonic setting when compared with patterns of unknown basalt samples. The dataset was generated using the entire EarthChem Portal database (http://www.earthchem.org) as of July 2015. A total of n = 894,439 individual samples were processed for data quality, assessing each for valid ages, labelling, sample site coordinates and consistent measurement units. Any data that could not be corrected, failed any criteria, or could not be converted to SI units were discarded. From the remaining data, a total of n = 97,952 basalt samples with ages ranging from 1000–0 Ma were identified by their respective EarthChem "ROCK NAME" label and extracted from the database. Tectonic environment data fingerprint models were built using all available basalt data aged 5–0 Ma (n = 38,301). Sample data were geographically assigned the one of three first-order tectonic environment labels of "MOR" for mid-ocean ridges (n = 18,213), "ARC" for subduction zones (n = 1858), and "OIB" for oceanic hotspot related upwellings (n = 7891) by comparing sample site locations with classification polygons derived from known present-day tectonic environment geometry and distribution (Müller et al., 2016). For each environment model, the EarthChem dataset contains up to a total of 136 possible sample discriminatory attributes, comprising of a combination of major and minor element measurements and element ratios. In order to analyze the sample data structure and not the individual geochemical measurement values all samples were normalized using feature scaling making values non-dimensional. In order to use the most representative and robust samples sets for training, only samples with attribute values between the 2.5th and 97.5th percentiles (representing a 2σ distribution) in the environment model training dataset (5–0 Ma) were included. To build a given tectonic environment model, available normalized attribute data from each labelled sample are sorted into 10 equal attribute magnitude bins, generating a frequency distribution for sample attribute occurrences for the given model within the dataset. A diagnostic weighting function is calculated for each model to isolate model attributes with the greatest discriminatory ability (i.e. entirely unique attributes or common attributes found to have unique magnitudes in a single model), positively weighting model discriminatory attribute and negatively weighting common or non-diagnostic attributes. The resulting function weights diagnostic attributes to comprise 50% of the total model fit score, with all non-diagnostic attributes making up the remaining model fit. To classify a basalt sample of unknown tectonic affinity, the individual sample attribute data is normalized using feature scaling using the defined the 5–0 Ma model 2σ distribution. The normalized model attribute data is then cross-referenced with all available environment models, returning a goodness-of-fit score for each attribute based on the match of the data structure of the unknown sample and the data structure of the given tectonic environment model. A maximum individual attribute score of 10 represents a perfect match with a given attribute highest frequency bin and a score of 1 represents a match with the lowest frequency bin. A total model fit is returned for all given environment models, and is the weighted sum of each available attribute fit score. A prediction confidence estimate is calculated for each total model fit using the number of attributes present in the unknown model compared to the number of attributes present in the given tectonic environment model and is weighted by the unknown sample fit to discriminatory attributes. As multiple samples exist at the same geographic localities combined with the use of rigid-plate reconstructions that do not account for deformation processesand significant reconstruction uncertainty for the Neoproterozoic, labelled sample predictions of congruent age and sample site are averaged using a 5º global mesh grid, producing a spatio-temporally averaged predicted sample set of *n* = 1561 from 1000–5 Ma.

To evaluate robustness and predictive ability, first-order tectonic fingerprint models were evaluated in two ways. Cross validation was performed on the 5–0 Ma dataset of *n* = 38,301 labelled samples used to build the fingerprint models. A total of 1000 independent validation tests were performed where the 5–0 Ma data were split into two sets; a randomly sampled training set consisting of 70% of the data (*n* = ~26,800), and a testing set consisting of the remaining 30% (*n* = ~11,500). For each validation test new first-order models were built for MOR, ARC, and OIB environments from the given validation training set, then used to new make environment predictions on the given validation testing set. For validation, a 2º global mesh grid was used for geographical averaging as opposed to the 5º global mesh grid used in the case study as the 5–0 Ma data is sampled in active present-day tectonic environments and is subsequently of higher spatial sampling precision. Resulting predictions were then compared against the original 5–0 Ma training set labels. From 1000 random cross validation tests, first order models predicted the present-day labelled 5–0 Ma training data at a mean success rate of 77.8% with a 2σ standard deviation of 1.45. The distribution of cross validation test success rates is shown in Fig. 2. The second evaluation of the method was to benchmark predictions of all available 'unknown' basalt data of *n* = 11,468 aged 30–5 Ma against labels for the same data points geographically classified by a given plate model (Muller et al., 2016). The plate model classification labelling process for the 30–5 Ma data was identical to the process used to label the 5–0 Ma training data in the main study and predictions were made using the full 5–0 Ma training dataset (*n* = 38,301). The full set of test data aged from 30–5 Ma had an overall mean prediction success rate of 73.2%, consistent with the results of the cross validation tests, with individual success rates of 84.4% for 10–5 Ma, 69.0% for 15–10 Ma, 69.6% for 20–15 Ma, 66.7% for 25–20 Ma, and 78.6% for 30–25 Ma.

Each resulting first-order fingerprint (FP) model contains up to a maximum of 136 discriminatory data attributes to describe the geographically classified environment (Fig. 3).

The predictive ability of the method shows significant tolerance and robustness to imperfect data as the method does not rely on inter-attribute correlation. Though potentially reducing the predictive ability of the method when analyzing highly consistent datasets, the independent nature of the proposed attribute analysis makes this method most practically applicable to deep time tectonic investigations where data are most spatio-temporally sparse and inconsistent. Using this new method, a sample of unknown origin containing any number or type of attributes can be classified by evaluating the available data attributes present in the sample against those present in the models (with predictive confidence proportional to the number and type of sample attributes present in the sample relative to the total number of attributes present in the model). In comparison with the case described in Section 1, using FP models without the restriction of requiring both Ti and V abundance measurements, it is possible to build tectonic discrimination models with almost 8 times the number of individual sample data from the same EarthChem dataset (n = 38,301 as opposed to n = 4914 data points). In contrast to traditional methods which evaluate the relationships between geochemical abundances to make predictions, this approach allows us to analyze the structure of the dataset itself to identify the source environment characteristics of a given sample. The discriminatory data structures in the models created in this study (as shown in Fig. 3) allow us to visualize and compare the unique data attribute characteristics of each first-order tectonic environment. Figure 3A–C show the data structure for MOR (103 attributes from 18213 samples), ARC (94 attributes from 1858 samples), and OIB (102 attributes from 7891 samples) models, respectively. In each plot, the x-axis represents a given non-dimensional sample attribute (full attribute list can be found in Supplementary Table S2), and the y-axis represents the 10 normalized histogram bins, with color opacity representing the frequency of data occurrence within a given bin. Cells marked in black represent the bin mode, that is the bin with the highest frequency occurrence for the given non-dimensional attribute. Each identified environment model from the 0–5 Ma training set produces a unique attribute 'fingerprint', showing very different data availability, distributions and patterns of highest frequency cells between models (Fig. 2A–C). It is this frequency structure that is used to ‘map’ the data of a given tectonic environment based on the training set and be used to identify unknown samples.

3. Tectonic discrimination

As the bulk geochemical composition of basalt formed within each first-order tectonic environment is very similar, there is a critical need to isolate key diagnostic attributes that can differentiate between given environments. Consistent with the assumptions behind the development of the traditional two- and three-dimensional discrimination diagrams, the tectonic environment FP models presented in this study seek to identify all key diagnostic attributes present in each environment model, that is attributes and associated values that are unique to a single model providing a robust discriminatory mechanism. Unidentified sample data found to contain any number of these key attributes are positively weighted towards the given environment identification, with sample data containing only weakly discriminatory attributes negatively weighted. Fig. 3D shows Fig. 3A–C plotted in grayscale and stacked with only identified model key diagnostic attribute values shown in color, demonstrating the uniqueness of each model attribute distribution. The full list of identified key diagnostic attributes for each tectonic environment can be found in Table 1. This approach presents a unique and flexible multi-variable tool to rapidly identify the first-order source environment of rocks of unknown tectonic affinity, in particular for samples that are either spatially or temporally inconsistently sampled and do not have all the specific measurements required by traditional limited-variable methods. Ti is identified as a key diagnostic attribute in n = 38,301 basalt samples, in agreement with traditional discrimination methods, increasing through bins 2, 3, and 4 for ARC, MOR and OIB respectively. However, V is not identified as a key diagnostic attribute in any model, rendering the 2D comparison of Ti/V like that of Shervais (1982) useful but potentially non-unique. Zr is identified in all models as diagnostic (bins 0, 1, and 3 for ARC, MOR and OIB respectively), and like the trend reported in Pearce and Cann (1973), the models suggest MOR basalts display low Zr and low-medium Ti (bins 0.0–0.1 and 0.3–0.4), OIB have low-medium Zr and medium Ti (bins 0.3–0.4 and 0.4–0.5), however, ARC models generally have both low Zr and low Ti (bins 0.0–0.1 and 0.2–0.3). The ternary plots presented in Pearce and Cann (1973) are not replicated in this study as neither Y nor Sr are identified as strongly diagnostic from the overall dataset.

4. Case study: Supercontinent formation and breakup

For more than 40 years, alternative models of the tectonic behaviour of pre-Pangea Earth have been published, suggesting a wide variety of interpretations of both the available geological data and the developing understanding of the supercontinent cycle itself (Piper et al., 1976; Bond et al., 1984; Dalziel, 1991; Moores, 1991; Torsvik et al.,1996; Dalziel, 1997; Meert and Torsvik., 2003; Pisarevsky et al., 2003; Collins and Pisarevsky, 2005; Li et al., 2008, 2013; Evans, 2009; Johansson, 2014, Merdith et al., 2017a). These models can be divided into 3 broad model classes, with one model from each class used in this case study. The most common class is referred to as the ‘continental drift’ type in which models are primarily focussed on the evolution of continental configuration through time and contain very little explicit plate boundary location or geometry information (Li et al., 2008). The second class of models is an augmentation of the traditional continental drift approach, producing ‘hybrid’ models primarily focussed on continental behaviours, but also containing predicted non-continuous boundary evolution information (Evans, 2009). The third class represents the most recent set of published models, namely ‘full-plate’ models. These models attempt to predict both continental and plate boundary evolution information and produce globally self-consistent predictions as the model operates as a ‘closed’ system (Gurnis et al., 2012; Merdith et al., 2017a). Although a significant evolution in development of tectonic reconstructions, the prediction of specific boundary environment types and evolution in deep time full-plate models remains challenging as the primarily data constraint of paleomagnetism does not contain explicit boundary information, and supporting geological data are limited. The FP algorithm was applied to n = 1547 unclassified dated samples labelled as basalt taken from the EarthChem portal. Samples were all aged between 1000 Ma and 410 Ma (representing only those rocks not included in the fingerprint models) in an attempt to self-consistently identify the tectonic environment from within which a given sample formed and evaluate the boundary predictions against those in a range of published pre-Pangea tectonic reconstructions. The resulting suite of 56 first-order predictions are listed in Table 2 and shown in Fig. 3D.

The relationship between China and Australia forms a key component of the Rodinian 'core' prior to breakup, with both the location and age of appearance of the Yangtze and Cathaysia blocks within Rodina varying greatly between published reconstructions. These differences reflect a critical divergence and uncertainty in the published interpretations of South China geology between ca. 1000Ma and 700 Ma. However, to practically discriminate alternative Rodinia configurations during this period is difficult, as there are few observations constraining this aspect of Rodinia's configuration. Previously, this problem has been assessed via plate kinematic data extracted from a range of published Rodinia-Gondwana transition reconstructions together with paleomagnetic data to evaluate the competing broader-scale Australia-Laurentia configurations during this period. This was achieved by comparing motion path geometries and plate velocities to identify configurations providing optimal kinematic behaviours (Merdith et al., 2017b). In our case-study, we apply tectonic environments predicted using the FP algorithm to evaluate contrasting configurations of the Australian block relative to the Yangtze and Cathaysia blocks from 1000–720 Ma from three alternative plate reconstructions of Rodinia; (i) Li et al. (2008), hereby referred to as L2008 (Fig. 3A), (ii) Evans (2009), hereby referred to as E2009 (Fig. 3B), and (iii) Merdith et al. (2017a), hereby referred to as M2017 (Table 2, Fig. 3C, and Supplementary Figs. S1 and S2).

Developing a method to consistently and objectively evaluate the fit of contrasting time-dependent plate model geometries using the tectonic environment predictions listed in Table 2 presents a significant challenge. As each class of model present different levels of component detail, such as the inclusion of continuous plate boundaries, or plume location predictions, models described in this case study were analyzed using the following simplified framework: (1) if a tectonic environment prediction for a given time is derived from samples located on a present-day continental block, then only models explicitly defining the given continental block at the given sample age will be considered, 2) where possible, all predicted tectonic environments are directly compared spatially with explicitly defined plate model topology geometries, and 3) if plate model topologies are not explicitly defined, where possible we consider the motion of individual blocks relative to neighboring blocks (either divergent or convergent), together with the location of the prediction site within the context of the surrounding model configuration.

L2008 suggests both the Yangtze and Cathaysia blocks have formed by 1100 Ma and are partially separated from each other by a subduction system as Cathaysia is connected to Laurentia at this time. Both move progressively southward from a relatively high latitudinal position of ~60ºN following the path of Laurentia from 1100 Ma through to 900 Ma, with complete South China Block amalgamation occurring between 1000 Ma and 900 Ma. Alternatively, M2017 suggests a significantly more dispersed continental configuration at 1000 Ma, with the Yangtze Block not considered in this model prior to ~850 Ma, and Cathaysia at a latitude of ~30ºN, straddled to the south by a subduction system and located almost antipodally to the forming Rodinia core. Similar to L2008, the E2009 model implies both Yangtze and Cathaysia are present at least by 1070 Ma and are connected via an inferred orogenic belt. However, unlike L2008 the South China Block is not centrally located within Rodinia in model E2009, but instead at the outer southeastern boundary approximately antipodal to the L2008 location, progressively moving northward from a latitude of ~60ºS to ~15ºS.

Tectonic environment predictions suggest subduction related basalts were forming at both ~980–970 Ma and ~950–940 Ma on the western and northwestern margins of the Yangtze block. This prediction is consistent with the suggestion of long-lived subduction (existing prior to 1100 Ma and ending between 1000 Ma and 900 Ma) outboard of the eastern, northern and western boundaries of the Yangtze Block as found in L2008, and the inferred environments surrounding South China in the E2009 model on the outer edge of Rodinia. The predicted presence of subduction adjacent to Cathaysia during this time is also in agreement with M2017, though drawing meaningful conclusions for this model is limited for times when the Yangtze Block is not properly considered. The next two predictions between ~910 Ma and ~890 Ma, although intraplate in nature, demonstrate good fits with to the MOR environment model indicating a possible plume or upwelling-related magmatic source. The first between ~910 Ma and ~900 Ma predicts either mid-ocean ridge or hotspot-related magmatic activity (prediction fits are within 3% of each other are treated as non-definitive) on the eastern margin of Yangtze, followed by a hotspot prediction between ~900 Ma and 890 Ma located on the northwestern margin of Yangtze. Both sites are located on present-day Yangtze; therefore M2017 cannot resolve these features. The configuration of the South China Block during this period in the E2009 model does not contradict these predictions; however, South China remains at the southeastern margin of Rodinia at this time and does not provide any explicit evidence for plume-related environments. Predictions of plume-related rocks appearing at the northwestern margin of Yangtze are consistent with predictions of precursory magmatism sourced from the proposed superplume in L2008, potentially indicating the initial stages of Rodinia breakup (Li et al., 2008). Between ~830 Ma and 810 Ma, during the period of protracted breakup of Rodinia, three subduction related predictions are made on the central Yangtze block close to the Yangtze-Cathaysian suture. At this time, the South China block is completely landlocked within the core of Rodinia in the L2008 model. Therefore it is uncertain how this series of basalts with an arc-related signature of this age could be found in this region, as South China is both fully amalgamated and significantly inboard of the eastern Rodinian margin at this time. However, this prediction cannot exclude the scenario that the signature being detected by the environment models could be an inherited signature from rocks related to the subduction outbound of Yangtze and its estimated cessation at ~900 Ma. Equally E2009 is unable to provide an explanation for the presence of subduction related rocks at this location, apart from the general inference that a subduction girdle may have existed surrounding Rodinia (Li et al., 2008). For this period, M2017 suggests the recently fully amalgamated South China block (from ~850 Ma onwards) is located significantly northwest of the Rodinia core and closely bound to the west by a subduction zone which is consistent with the predictions (Fig. 3C). As few observations exist to constrain both India and the Yangtze-Cathaysian system during the Neoproterozoic, the prediction of subduction-related basalts continuing to form along the margin of Yangtze and Cathaysia as late as ~820 Ma suggests South China was likely still forming, and a more complex suite of subduction systems may have been active in this region at this time. Temporally concurrent with the formation of these subduction-related rocks forming within the Yangtze-Cathaysia boundary, for the following ~30–40 million years a long series of either mid-ocean ridge or plume-related basalts are predicted to form within the South China block (primarily within Yangtze). Beginning with two first-order mid-ocean ridge (upwelling) predictions forming along the present-day southwestern Yangtze boundary at ~820 Ma, followed by a hotspot prediction at ~810 Ma in the same region, then at ~800 Ma by two additional predictions of hotspot and a mid-ocean ridge magmatism located in the present-day northeastern Yangtze and Cathaysia blocks, respectively, and finally at ~780 Ma a plume signature predicted within the central South China Block (Fig. 3A–C). During this period, three mid-ocean ridge related igneous signatures are also predicted in present-day southern Australia, the first at ~820 Ma, and both the second and third at ~800 Ma, all temporally congruent with the timing of equivalent signatures within the South China block. Although explicitly supporting the existence of subduction-related basalts throughout this period, the continued positioning of South China significantly northwest of the Rodinia M2017 does not provide defined predictions directly compatible with any of the mid-ocean ridge or plume-related predictions located in Yangtze or Cathaysia between ~820 Ma and 780 Ma (Fig. 3C). The southern Australian MOR signatures at ~820 Ma is also not supported by the M2017 boundary configuration at this time, but the two later MOR predictions at ~800 Ma support the M2017 configuration with initiation of the Proto-Pacific Ocean separating Laurentia from Australia, Antarctica, North China and Tarim (Fig. 3C). Dependent on the uncertainty of the constraints used to nominate the beginning of Rodinia core breakup in M2017, precursory upwellings associated with the initiation of spreading may account for the slightly older mid-ocean ridge predictions of this study in southern Australia at ~830 Ma and ~820 Ma. Alternatively, the configuration of South China relative to Australia presented in L2008 from ~820 Ma through to ~780 Ma is consistent with the predictions of this study (Fig. 3A). The generation of mantle upwelling-related rocks appearing simultaneously within both South China and southern Australia at this time appear to resemble the result of a radial dyke swarm-like feature centered between South China, Australia, and Mawson (Li et al., 2008*)*. The final hotspot prediction within the South China block at ~780 Ma coincides precisely with the initiation of Rodinia breakup in L2008 (Fig. 3A), resulting in Australia-Mawson and Laurentia both beginning separation from South China as a result of a newly formed triple-ridge junction in the Proto Pacific Ocean. The final prediction of a mid-ocean ridge-related basalt at ~720 Ma in central Cathaysia does appear to be supported by the continued presence of a suggested superplume beneath South China in L2008; however the configuration presented in M2017 at this time, although not excluding the possibility, does not provide any explicit explanation for this prediction. Throughout this period, E2009 predicts the continued location of South China at the southwestern boundary of Rodinia from ~820 Ma to 780 Ma (Fig. 3B), a supercontinent location more typically associated with subduction systems (Li et al., 2008; Li et al., 2013; Merdith et al., 2017a), and does not provide explicit prediction or motion evidence to support upwelling within South China during this period.

From the simple analysis performed above in the case study according to the defined framework, the respective spatio-temporal configurations of the South China and Australian Blocks proposed within the Rodinia reconstruction of Li et al. (2008) appears to demonstrate the greatest consistency with the new paleo-environment fingerprint database, particularly for configurations related to Rodinia formation and breakup. The alternative configurations proposed in the models of Evans (2009) and Merdith et al. (2017a), although demonstrating compatibility with predictions related to early Rodinia formation (E2009) and stable core configurations (M2017), respectively, are less consistent with many of the tectonic paleo-environment predictions throughout the supercontinent cycle. However, these results need to be interpreted in the context of a number of considerations. The first one is the apparent ~850 Ma appearance time of the Yangtze Block in M2017, preventing evaluation prior to this time using tectonic environment predictions. It is acknowledged that as the M2017 model describes Cathaysia associated with subduction systems between 1000 Ma and 850 Ma, the Yangtze Block is likely to be a suprasubduction-related accretionary complex during this period (Cawood et al. 2013, 2017), and subsequently not included as a 'cratonic' block in the model (Merdith et al., 2017a). If taken into account, both the previously unconsidered 'Yangtze' subduction (ARC) predictions at 980 Ma and 950 Ma, respectively, would be consistent with M2017 predictions. The second important consideration in evaluating these results are the time-dependent kinematic implications of each model geometry. A key difference between L2008 and M2017 (E2009 is a unique solution) is the choice of the Australia-Laurentia configuration model, with L2008 adopting a Missing-Link geometry (Li et al. 1995, 2008), and M2017 incorporating an AUSWUS (Australia-Western United States) type configuration (Karlstrom et al., 1999, 2001). Kinematic analyses of each configuration type presented by Merdith et al., (2017b) concluded that during the period of Rodinia break-up ca. 800 Ma, the Missing-Link configuration produces the highest average spreading rates of up to ~90 km/Ma compared with ~57 km/Ma for AUSWUS, the lowest result of the study. For configurations containing a proposed later breakup at c.a. 725 Ma, spreading rates of ~150 km/Ma and ~130 km/Ma were calculated for Missing-Link and AUSWUS respectively. The study also found that motion paths for AUSWUS-based configurations for significantly simpler than those of Missing-Link-based geometries, as the latter require more complex plate motions to meet geological constraints. Although not explicitly considered in this case-study for evaluating alternative Rodinia configurations, these kinematic analyses reinforce the dependency of each configuration on the primary constraints considered, identifying the potential for over-fitting certain constraints at the expense of others.

5. Conclusions

Geochemical analysis is a key instrument in the study of long-term tectonics on Earth. When coupled with auxiliary geological and geophysical datasets able to contribute paleogeographic constraints such as paleomagnetics, it provides the unique ability to isolate the subtle yet highly diagnostic chemical attributes of rock samples which can identify the rock type and source environment. In this paper we demonstrate the limited application, scope and diagnostic ability of published geochemical discrimination methods to accurately identify tectonic source environments from basaltic geochemistry for use as constraints in deep-time tectonic reconstructions without the need for fixed a priori assumptions, highly filtered datasets, and strict input data requirements. Applying a new flexible framework to this long standing problem, from an unfiltered geochemistry database of n = 38,301 basalt samples of Pliocene age or younger, we present a newly derived and robust set of first-order discriminatory tectonic environment models for mid-ocean ridge (MOR), subduction (ARC), and oceanic hotspot (OIB) environments respectively. Using these discriminatory environment models, we analysed a sparse, inconsistent and unfiltered geochemical database of n = 1547 basalt samples of unknown tectonic affinity ranging in age from 1000 Ma and 410 Ma. From this analysis, we present a new suite of 56 identified first-order tectonic paleo-environments spanning the Neoproterozoic, Cambrian, Ordovician and Silurian, together forming a practical dataset directly applicable to both reconstructing new, and evaluating existing models of Rodinia supercontinent amalgamation, stability, and dispersal. To demonstrate this, we analysed the predicted Proterozoic motion histories of the South China and Australian Blocks, together forming a key component within published Rodinia configurations, from three alternative published reconstructions for consistency with the new paleo-environment dataset. From these analyses, the Rodinia reconstruction L2008 of Li et al. (2008) demonstrated the highest degree of both spatial and temporal fit with paleo-environment predictions, with the new dataset in particular informing upwelling or plume-related environments through periods of supercontinent formation and dispersal.

However, the case-study framework also highlighted a lesser degree of fit with subduction environment predictions, specifically related to prediction from samples sourced in present-day South China. Conversely, subduction environment predictions appear more consistent with the configurations presented in M2017 (Merdith et al., 2017a), whereas explicit hotspot-related predictions were not present. Model E2009 (Evans, 2009), although more experimental in its nature, also shows consistency with South China subduction predictions as these blocks maintain positions on the margins of Rodinia throughout the study period, but demonstrates less consistency with the other prediction types. Further, when assessed in the context of the kinematic analyses of key alternative Rodinia configurations as described in Merdith et al. (2017b), L2008 although demonstrating increased fit with the tectonic environment predictions derived from the EarthChem geochemistry database in this study, requires both a more complex and higher velocity plate motion evolution history than that of M2017, highlighting a key consideration in the development of deep-time plate reconstructions.

Acknowledgements

This research was supported by the Science Industry Endowment Fund (RP 04-174) Big Data Knowledge Discovery Project. MGT received additional support from a CSIRO-Data61 Postgraduate Scholarship. ZXL acknowledges the support of the Australian Research Council through a Laureate Fellowship grant (FL150100133). We thank Derrick Hasterok and an anonymous reviewer for their considered and constructive reviews, Craig O'Neill, Peter Cawood, Thomas Bodin and Andrew Merdith for helpful comments that improved the paper, and Jason Ash and Daniel Steinberg for their combined technical support in developing these analyses. This is a contribution to IGCP project 648. Analyses were conducted using the following open source tools: GPlates and pyGPlates (www.gplates.org), Python (www.python.org), and Project Jupyter (www.jupyter.org).

References

Agrawal, S., Guevara, M. and Verma, S. P., 2004 ‘Discriminant Analysis Applied to Establish Major-Element Field Boundaries for Tectonic Varieties of Basic Rocks’, *International Geology Review*, 46(7), pp. 575–594. doi: 10.2747/0020-6814.46.7.575.

Bond, G. C., Nickeson, P. A. and Kominz, M. A. (1984) ‘Breakup of a supercontinent between 625 Ma and 555 Ma : new evidence and implications for continental histories’, *Earth and Planetary Science Letters*, 70, pp. 325–345.

Cawood, P. A. *et al.*(2013) ‘Locating South China in Rodinia and Gondwana: A fragment of greater India lithosphere?’, *Geology*. GeoScienceWorld, 41(8), pp. 903–906. doi: 10.1130/G34395.1.

Cawood, P. A. *et al.* (2017) ‘Reconstructing South China in Phanerozoic and Precambrian supercontinents’, *Earth-Science Reviews*, 3 July. doi: 10.1016/j.earscirev.2017.06.001.

Collins, A. S. and Pisarevsky, S. A. (2005) ‘Amalgamating eastern Gondwana: The evolution of the Circum-Indian Orogens’, *Earth-Science Reviews*. Elsevier, pp. 229–270. doi: 10.1016/j.earscirev.2005.02.004.

Dalziel, I. W. D. (1991) ‘Pacific margins of Laurentia and East Antarctica-Australia as a conjugate rift pair: evidence and implications for an Eocambrian supercontinent’, *Geology*, 19(6), pp. 598–601. doi: 10.1130/0091-7613(1991)019<0598:PMOLAE>2.3.CO;2.

Dalziel, I. W. D. (1997) ‘Neoproterozoic-paleozoic geography and tectonics: Review, hypothesis, environmental speculation’, *Bulletin of the Geological Society of America*. GeoScienceWorld, pp. 16–42. doi: 10.1130/0016-7606(1997)109<0016:ONPGAT>2.3.CO;2.

Domeier, M. and Torsvik, T. H. (2014) ‘Plate tectonics in the late Paleozoic’, *Geoscience Frontiers*. Elsevier Ltd, 5(3), pp. 303–350. doi: 10.1016/j.gsf.2014.01.002.

Evans, D.A.D., 2009. The palaeomagnetically viable, long-lived and all-inclusive Rodinia supercontinent reconstruction. Geological Society, London, Special Publications, 327(1), pp.371–404.

Gurnis, M. *et al.* 2012 ‘Plate tectonic reconstructions with continuously closing plates’, *Computers & Geosciences*, 38(1), pp. 35–42. doi: 10.1016/j.cageo.2011.04.014.

Johansson, Å. (2013) ‘From Rodinia to Gondwana with the “SAMBA” model-A distant view from Baltica towards Amazonia and beyond’, *Precambrian Research*. Elsevier B.V., 244, pp. 226–235. doi: 10.1016/j.precamres.2013.10.012.

Karlstrom, K.E., Harlan, S.S., Williams, M.L., McLelland, J., Geissman, J.W. and Ahall, K.I., 1999. Refining Rodinia: Geologic evidence for the Australia–western US connection in the Proterozoic. *GSA Today*, *9*(10), pp.1-7.

Karlstrom, K. E. *et al.* (2001) ‘Long-lived (1.8-1.0 Ga) convergent orogen in southern Laurentia, its extensions to Australia and Baltica, and implications for refining Rodinia’, *Precambrian Research*. Elsevier, 111(1–4), pp. 5–30. doi: 10.1016/S0301-9268(01)00154-1.

Li, Z-X., Zhang, L. and Powell, C. M. (1995) ‘South China in Rodinia: part of the missing link between Australia- East Antarctica and Laurentia?’, *Geology*. GeoScienceWorld, 23(5), pp. 407–410. doi: 10.1130/0091-7613(1995)023<0407:SCIRPO>2.3.CO;2.

Li, Z.-X. et al., 2008. Assembly, configuration, and break-up history of Rodinia: A synthesis. Precambrian Research, 160(1–2), pp.179–210.

Li, Z.-X., Evans, D. A. D. and Halverson, G. P. (2013) 'Neoproterozoic glaciations in a revised global palaeogeography from the breakup of Rodinia to the assembly of Gondwanaland', *Sedimentary Geology*, 294, pp. 219–232. doi: 10.1016/j.sedgeo.2013.05.016.

Matthews, K.J., Maloney, K.T., Zahirovic, S., Williams, S.E., Seton, M. and Mueller, R.D., 2016. Global plate boundary evolution and kinematics since the late Paleozoic. *Global and Planetary Change*, *146*, pp.226-250.

Meert, J. G. and Torsvik, T. H. (2003) 'The making and unmaking of a supercontinent: Rodinia revisited', *Tectonophysics*, 375(1–4), pp. 261–288. doi: 10.1016/S0040-1951(03)00342-1.

Merdith, A. S. *et al.* (2017a) 'A full-plate global reconstruction of the Neoproterozoic', *Gondwana Research*. Elsevier B.V., pp. 84–134. doi: 10.1016/j.gr.2017.04.001.

Merdith, A. S. *et al.* (2017b) 'Kinematic constraints on the Rodinia to Gondwana transition', *Precambrian Research*. Elsevier B.V., 299, pp. 132–150. doi: 10.1016/j.precamres.2017.07.013.

Moores, E. M. (1991) 'Southwest US-East Antarctic (SWEAT) connection: a hypothesis', *Geology*, 19(5), pp. 425–428. doi: 10.1130/0091-7613(1991)019<0425:SUSEAS>2.3.CO;2.

Müller, R.D. et al., 2016. Ocean basin evolution and global-scale reorganization events since Pangea breakup. Annual Review of Earth and Planetary Science Letters, (April), pp.107–138.

Pearce, J.A. & Cann, J.R., 1973. Tectonic setting of basic volcanic rocks determined using trace element analyses. Earth and Planetary Science Letters, 19(2), pp.290–300.

Pearce, J. A. (2008) 'Geochemical fingerprinting of oceanic basalts with applications to ophiolite classification and the search for Archean oceanic crust', *Lithos*. Elsevier, 100(1–4), pp. 14–48. doi: 10.1016/j.lithos.2007.06.016.

Piper, J. D. A., Beckmann, G. E. J. and Badham, J. P. N. (1976) 'Palaeomagnetic Evidence for a Proterozoic Super-Continent [and Discussion]', *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*. Royal Society, pp. 469–490. doi: 10.2307/74572.

Pisarevsky, S. A. *et al.*(2003) 'Models of Rodinia assembly and fragmentation', *Geological Society, London, Special Publications*. Geological Society of London, 206(1), pp. 35–55. doi: 10.1144/GSL.SP.2003.206.01.04.

Schettino, A. and Scotese, C. R. (2005) 'Apparent polar wander paths for the major continents (200 Ma to the present day): a palaeomagnetic reference frame for global plate tectonic reconstructions', *Geophysical Journal International*, 163(2), pp. 727–759. doi: 10.1111/j.1365-246X.2005.02638.x.

Seton, M. *et al.* (2012) 'Global continental and ocean basin reconstructions since 200Ma', *Earth-Science Reviews*, 113(3), pp. 212–270.

Shervais, J.W., 1982. Ti-V plots and the petrogenesis of modern and ophiolitic lavas. Earth and Planetary Science Letters, 59(1), pp.101–118.

Stampfli, G. M., Hochard, C., Vérard, C., Wilhem, C. and VonRaumer, J. (2013) 'The formation of Pangea', *Tectonophysics*. Elsevier, pp. 1–19. doi: 10.1016/j.tecto.2013.02.037.

Torsvik, T. *et al.* (1996) 'Continental break-up and collision in the Neoproterozoic and Palaeozoic — A tale of Baltica and Laurentia', *Earth-Science Reviews*, 40(3–4), pp. 229–258. doi: 10.1016/0012-8252(96)00008-6.

Torsvik, T. H., Müller, R. D., Van der Voo, R., Steinberger, B. and Gaina, C. (2008) 'Global plate motion frames: Toward a unified model', *Reviews of Geophysics*, 46(3), p. RG3004. doi: 10.1029/2007RG000227.

Verma, S. P., Guevara, M. and Agrawal, S., 2006 'Discriminating four tectonic settings: Five new geochemical diagrams for basic and ultrabasic volcanic rocks based on log — ratio transformation of major-element data', *Journal of Earth System Science*, 115(5), pp. 485–528. doi: 10.1007/BF02702907.

Verma, S.P., 2010. Statistical evaluation of bivariate, ternary and discriminant function tectonomagmatic discrimination diagrams. *Turkish Journal of Earth Sciences*, *19*(2), pp.185-238.

Vermeesch, P. 2006a 'Tectonic discrimination of basalts with classification trees', *Geochimica et Cosmochimica Acta*, 70(7), pp. 1839–1848. doi: 10.1016/j.gca.2005.12.016.

Vermeesch, P., 2006b. Tectonic discrimination diagrams revisited. *Geochemistry, Geophysics, Geosystems*, 7(6).

Figure Captions

Figure 1. Titanium / Vanadium (Ti/V) ratio tectonic environment discrimination diagrams derived from n = 4914 basalt samples taken from the EarthChem portal (http://www.earthchem.org). All samples are aged between 5 Ma and present-day and were classified using the Müller et al. (2016) plate reconstruction. Blue points are samples classified as subduction zone related basalts (ARC), orange points are samples classified as mid-ocean ridge related basalts (MOR), and red points are classified as oceanic hotspot related basalts (OIB). Gray dashed trend lines and associated grey labels represent reference discriminatory bounding ratios of Ti/V = 10, 20, 50, and 100 (Shervais, 1982). (A) Blue, orange, and red trend lines represent updated reference discriminatory bounding ratios representing a 2σ distribution about each classified population mean. The shaded gray regions represent overlap between the discriminatory bounds. (B) Outer blue, orange and red polygons represent 0.9 probability contours and inner polygons represent 0.1 probability contours for each tectonic environment calculated using probability density functions.

Figure 2: Results from 1000 independent cross validation tests using a 2º spatial grid, each performing a 70% / 30% random split of the 0–5 Ma data for training and testing data respectively. Blue bars represent probability (0.0–1.0) of individual success rates. Red dotted line shows the gaussian distribution of cross validation test results.

Figure 3. Tectonic environment data fingerprint models for (A) mid-ocean ridge (colored orange), (B) subduction (colored blue), and (C) oceanic hotspot (colored red). Color opacity gradient indicates normalized attribute bin frequency. Black cells represent maximum frequency bin for associated attribute. White space indicates zero data available for associated attribute. (D) Stacked composite of models A, B and C colored in gray, with individual model attribute maximum frequency bins colored by model type (A, B or C) visualizing the key discriminatory attributes for each FP model. For full attribute number key, see Supplementary Table S2.

Figure 4. Alternative tectonic configurations of Rodinia at 780 Ma as predicted by the three published reconstructions of (A) Li et al. (2008), (B) Evans (2009), and (C) Merdith et al. (2017a). Solid yellow lines represent mid-ocean ridges, with dashed yellow lines representing poorly constrained or inferred mid-ocean ridge predictions. Blue solid lines are subduction zones and solid black lines represente passive / transform boundary segments. Reconstructed first-order tectonic environment predictions aged 780 ± 20 Ma are shown as filled circles at ~800 Ma, filled squares at ~780 Ma, and filled hexagons at ~760 Ma. Predictions are labelled as per ID listed in Table 2 and color-coded by type: blue = ARC, yellow = MOR, and red = OIB. Green filled polygons = Australia, cyan filled polygons = Cathaysia, and yellow filled polygon = Yangtze. A, Australia; A-A, Afif-Abas Terrane; Am, Amazonia; Az, Azania; Ba, Baltica; By, Bayuda; Ca, Cathaysia (South China); C, Congo; DML, Dronning Maud Land; G, Greenland; I, India; K, Kalahari; L, Laurentia; Ma, Mawson; NAC, North Australian Craton; NC, North China; Pp, Paranapenema; Ra, Rayner (Antarctica); RDLP, Rio de la Plata; SAC, South Australian Craton; SF, São Francisco; Si, Siberia; SM, Sahari Megacraton; WAC, West African Craton; Yg, Yangtze (South China). (D) Graphical representation of tectonic environment classification predictions and associated model fits from 1000–420 Ma as listed in Table 2. Calculated fits to all predicted first-order environment models of subduction (ARC), mid-ocean ridge (MOR), and oceanic hotspot (OIB) are shown as filled blue, red and orange circles respectively. Circle size is directly proportional to prediction confidence, with larger circles indicating higher prediction

confidence. Results are divided by reconstruction snapshot age (as presented directly above each cell), with individual prediction ID labels as per Table 2 presented directly below each cell. A full suite of all reconstruction snapshots overlaid with predictions can be found in the supplement.

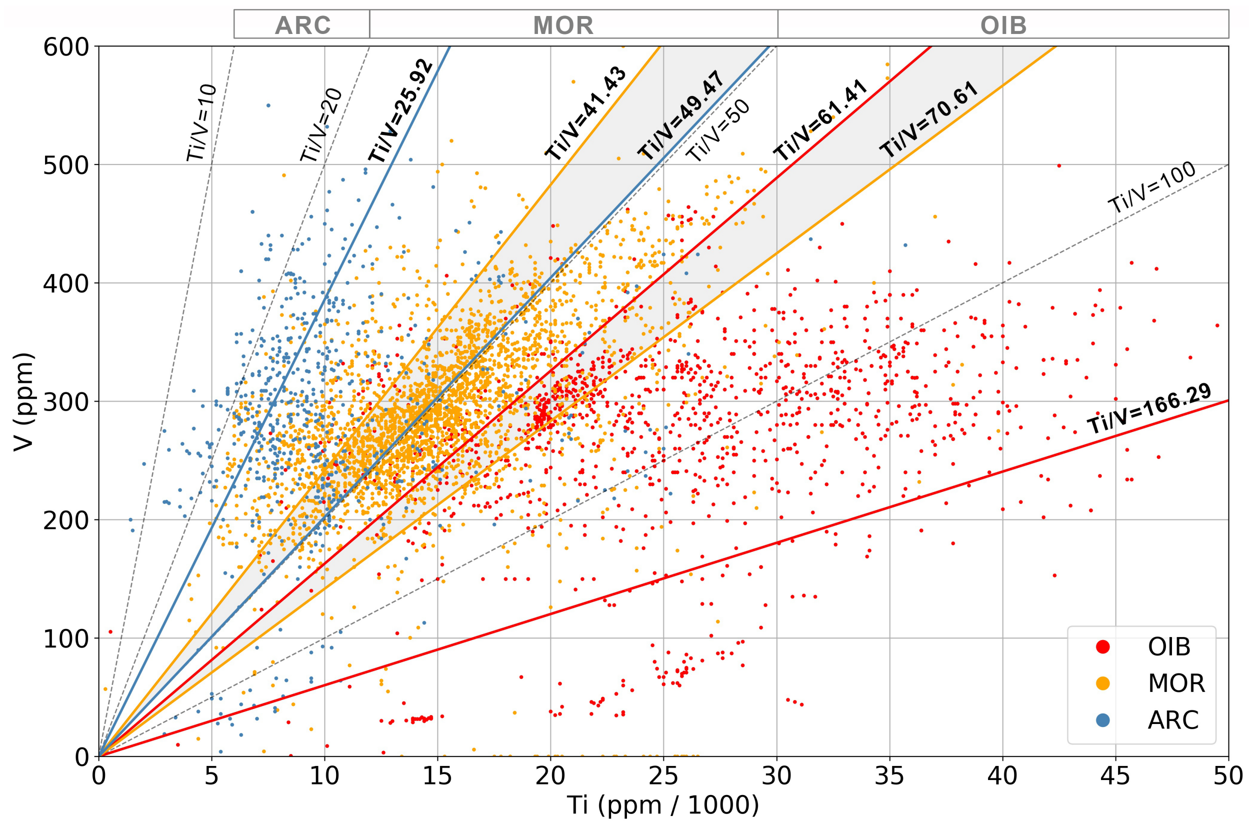
Table captions

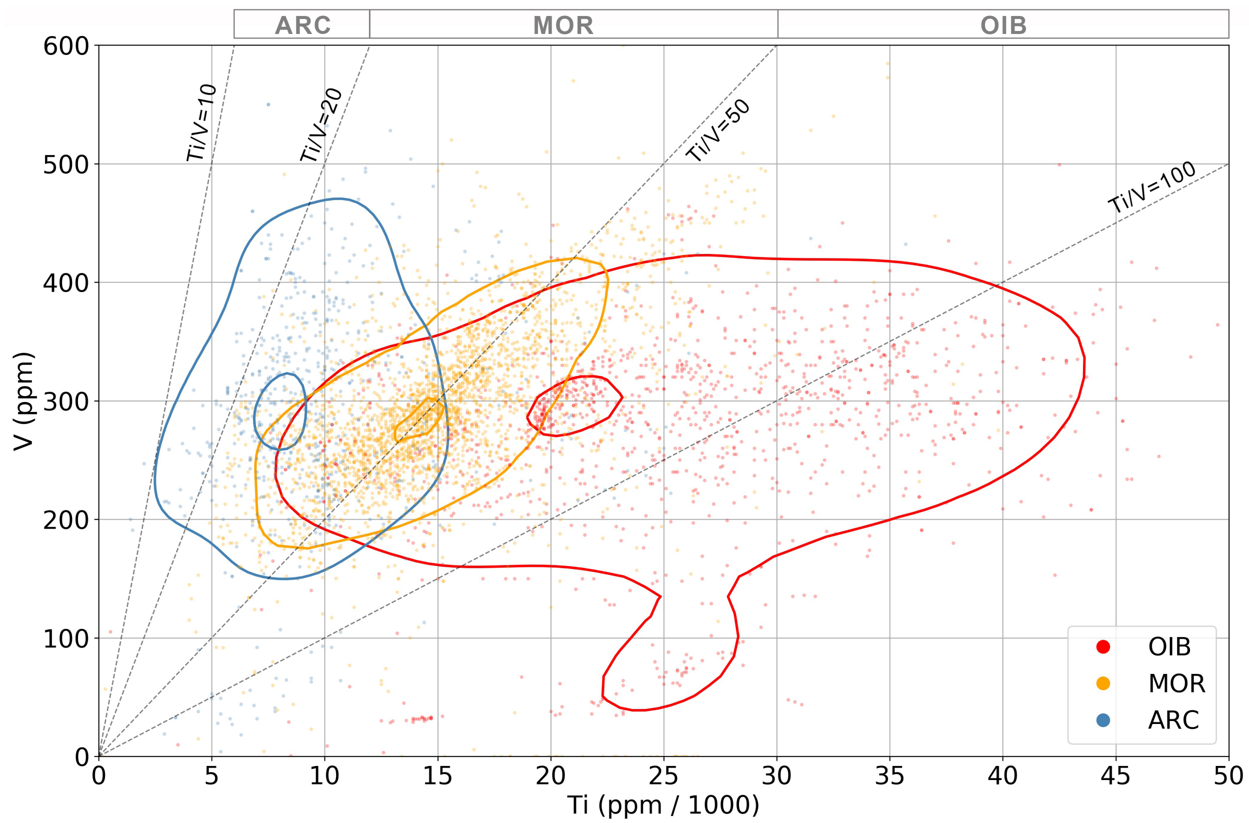
Table 1. Identified key discriminatory attributes and their non-dimensional frequency magnitudes for each first-order environment model. ARC model calculated from *n*= 1858 samples, MOR model calculated from *n* = 18213 samples, and OIB model calculated from *n* = 7891 samples. Freq. = Normalized data frequency bin, Att no. = Model attribute ID number, Att ref. = Reference sample data measurement name taken from EarthChem Portal. Full Att no. and Att ref. listed in Supplementary Table S2.

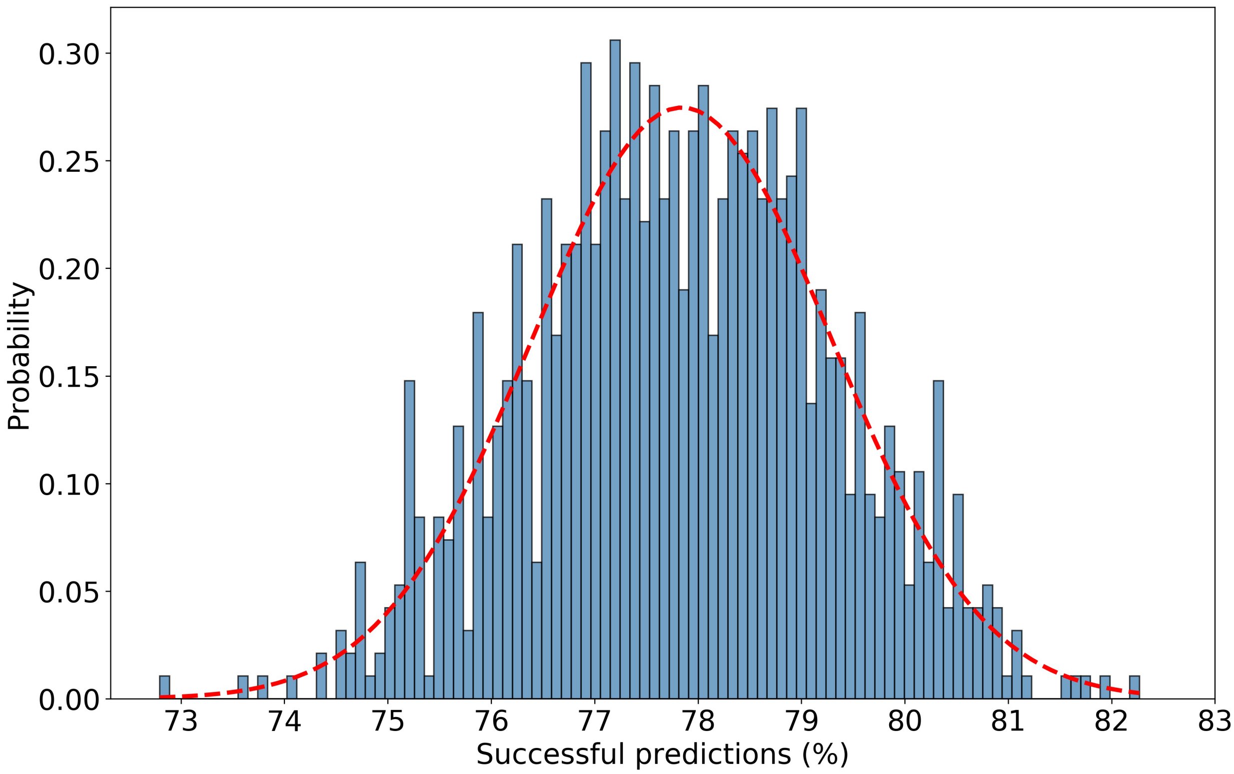
Table 2. First-order tectonic environment predictions from 1000–410 Ma grouped into 10 Ma age bins. ID = Data point ID.*, ARC, MOR, OIB fit%* = calculated percentage fit of the given sample against each environment model. Bold values indicate best fit model. Italic values indicate multiple results within threshold of 3%. ARC, MOR, OIB conf. = calculated confidence parameter for each prediction, Site lat = present-day sample site latitude, Site lon = present-day site longitude.

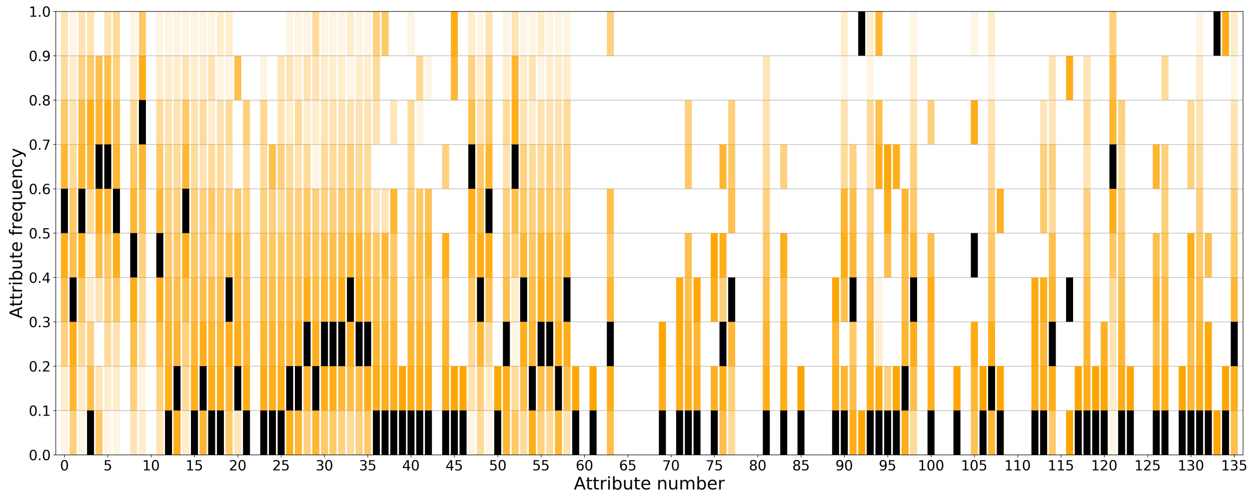
	ARC		MOR		OIB	
Freq.	**Att no.**	**Att ref.**	**Att no.**	**Att ref.**	**Att no.**	**Att ref.**
0.9 - 1.0			92	Nd-143/Nd-144	61	Pa-231
0.8 - 0.9	2	Al_2O_3			50	Fe
	92	Nd-143/Nd-144	9	CaO	99	Nd ε
	123	Al				
0.7 - 0.8			47	Sc	5	FeO
			52	Mn	114	In
0.6 - 0.7	52	Mn			1	TiO_2
	87	Pb-206/Pb-204			9	CaO
			5	FeO	46	Mg
					71	U-238/Pb-204
					123	Al
0.5 - 0.6	5	FeO	19	Cr_2O_3	52	Mn
	9	CaO	63	U-234/U-238	56	Cu
	47	Sc	91	Th-230/U-238	97	P
0.4 - 0.5	50	Fe	2	Al_2O_3	27	Eu
	56	Cu	33	Tm	28	Gd
	86	Pb-207/Pb-204	77	Lu-176/Hf-177	47	Sc
	88	Pb-208/Pb-204	105	Te	48	Ti
	106	Pt			53	Co
					58	Ga
					63	U-234/U-238
					69	Th-232/Pb-204
					89	Sn
0.3 - 0.4	44	K	32	Er	26	Sm
	58	Ga	48	Ti	29	Tb
			53	Co	57	Zr
			56	Cu	107	Hf
			114	In		
			116	Pb-210/Ra-226		
0.2 - 0.3	46	Mg	1	TiO_2	2	Al_2O_3
	48	Ti	27	Eu	32	Er
	53	Co	28	Gd	33	Tm
	77	Lu-176/Hf-177	29	Tb	37	Be
			58	Ga	44	K
			107	Hf		
0.1 - 0.2	27	Eu	26	Sm	91	Th-230/U-238
	28	Gd	57	Zr	105	Te
	29	Tb	76	Xe-129/Xe-132	106	Pt
	32	Er	89	Sn	117	Ag
	33	Tm	97	P		
	37	Be				
	84	Sr-87/Sr-86				
	94	Ra-226/Th-230				
0.0 - 0.1	1	TiO_2	37	Be	19	Cr_2O_3
	26	Sm	44	K	60	Ra-226
	57	Zr	46	Mg	67	Th-232/U-238
	89	Sn	61	Pa-231	78	Hg
	97	P	69	Th-232/Pb-204		
	107	Hf	71	U-238/Pb-204		
	114	In	94	Ra-226/Th-230		
	124	Be-10/Be-9	96	I		
			106	Pt		
			117	Ag		
			123	Al		

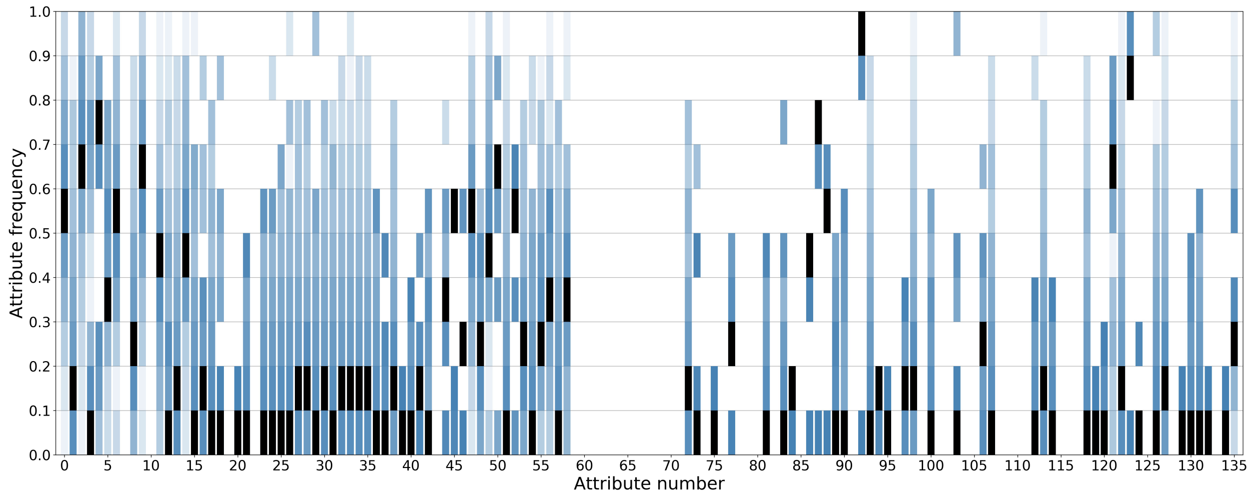
Age (Ma)	ID	ARC fit %	MOR fit %	OIB fit %	ARC conf.	MOR conf.	OIB conf.	Site lat	Site lon
1000-990	A	66.89	**71.45**	68.43	15.4	14.26	14.5	47.1	-84.7
	B	*77.81*	***78.4***	65.54	30.76	28.16	28.3	72.76	-80.5
980-970	A	66.3	64.55	**70.17**	41.45	38.92	39.38	60	136
	B	**67.64**	57.95	63.22	27.39	26.79	27.11	24.59	102.07
950-940	A	**80.93**	73.19	57.15	33.42	30.64	30.75	33	107.6
910-900	A	74.09	***79.31***	*77.37*	37.89	35.11	35.47	26.9	101.57
900-890	A	57.84	58.17	**62.83**	25.84	23.49	23.75	33	107.6
830-820	A	72.35	**79.96**	68.66	28.61	26.65	26.89	-30.64	139.13
	B	**80.11**	66.99	51.33	40.24	37.63	37.27	28.58	112.34
	C	**83.28**	74.87	68.71	41.19	38.13	38.58	25.72	109.87
820-810	A	**78.54**	69.6	57.17	28.67	26.93	26.89	28.59	112.33
	B	64.18	***68.56***	*65.9*	35.48	33.8	34.03	32.5	107
	C	69.31	**75.03**	69.78	43.6	41.34	40.24	32.5	105
	D	77.43	**83.24**	74.67	26.1	23.83	24.05	-30	133
810-800	A	66.95	62.44	**75.29**	35.88	33.58	34.05	29.16	102.8
800-790	A	64.82	61.34	**69.45**	35.84	33.23	34.01	27.63	117.89
	B	73.93	***76.9***	*74.54*	36.21	32.96	33.31	29.97	120.2
	C	73.91	**82.23**	72.54	50.87	45.97	46.58	-25.23	131.51
	D	75.45	**79.68**	72.37	47.75	43.13	42.43	-23.75	134.11
780-770	A	*59.38*	55.59	***61.16***	25.33	25.99	26.8	27.53	110.72
760-750	A	*64.37*	***67.1***	61.57	13.68	12.71	13	41.58	86.86
750-740	A	62.87	66.33	**79.58**	46.27	44.41	45.16	-25.05	123.76
720-710	A	*56.37*	***58.98***	*58.85*	17.21	16.45	15.71	23.4	111.72
620-610	A	***68.7***	*66.21*	62.87	27.38	25.59	26.23	45.5	-64.1
	B	57.94	61.63	**66.43**	20.21	19.12	18.71	58	6
600-590	A	*73.61*	***74.13***	*71.14*	19.83	18.27	18.38	42.22	-70.88
600-590	B	**74.91**	71.28	57.25	10.51	10.33	10.44	36.39	-78.98
580-570	A	**69.78**	62.25	65.28	29.81	27.92	28.21	22	29
560-550	A	55.61	62.07	**69.58**	17.43	16.21	16.12	46.03	-71.64
550-540	A	57.36	57.92	**63.27**	23.44	21.56	21.9	29.9	35.1
	B	*58.77*	***61.16***	*59.2*	18.88	17.11	17.81	30.63	35.5
540-530	A	***67.22***	*65.93*	60.48	26.29	25.3	25.41	48.1	-68.5
	B	***73.1***	*72.87*	61.87	38.77	35.35	35.3	-17.27	128.72
520-510	A	*64.91*	***67.15***	*66.65*	22.12	21.05	21.14	-30.71	142.04
	B	**72.01**	66.16	57.88	32.24	30.21	30.73	38	94
	C	*69.58*	***71.26***	61.41	38.36	35.07	34.94	-27.05	125.16
	D	*69.32*	***70.3***	65.18	18.89	17.51	17.68	45.39	-66.22
	E	57.3	69.35	**73.39**	22.74	20.51	22.77	39.21	-112.95
	F	67.48	69.31	**73.19**	29.11	25.77	27.75	37.08	-77.7
	G	66.02	69.04	**76.74**	44.61	40.48	45.2	36.64	-81.73
500-490	A	45.44	42.33	**49.62**	24.65	26.55	27.14	46.63	-70.98
	B	**85.68**	75.78	56.62	32.09	29.5	29.83	48.55	-56.65
480-470	A	*72.37*	***73.47***	66.78	22.2	20.78	21.06	46.6	-60.5
	B	*63.85*	62.77	***65.95***	30.12	27.24	28.39	46.63	-68.63
470-460	A	76.58	**85.16**	68.68	41.45	38	38.57	44.3	-69.32
	B	**73.57**	67.6	55.88	10.76	9.81	9.92	45.4	-71.9
	C	***72.28***	*70.41*	62.81	35.57	32.6	35.08	46.53	-67.93
	D	65.61	*70.97*	***71.35***	29.03	27.02	27.12	49.92	-55.83
	E	*76.3*	***77.35***	61.61	22.69	21.78	21.34	-24.53	-66.47
	F	*81.11*	***80.25***	70.04	46.3	43.3	43.36	-32.35	-69.18
	G	*76.27*	***78.31***	72.95	17.56	16.6	16.73	40.42	-76.47
	H	75.76	**83.93**	80.91	9.68	9.67	10.44	35.65	-80.1
460-450	A	*75.73*	***75.99***	64.17	30.78	28.23	29.4	46.54	-68.59
430-420	A	*76.31*	***78.31***	70.7	33.09	29.79	32.45	44.48	-68.01
	B	*68.8*	66.23	***70.59***	38.89	37.64	38.03	48	-65
420-410	A	62.94	62.12	**67.61**	35.58	34.86	36.43	48.29	-65.3

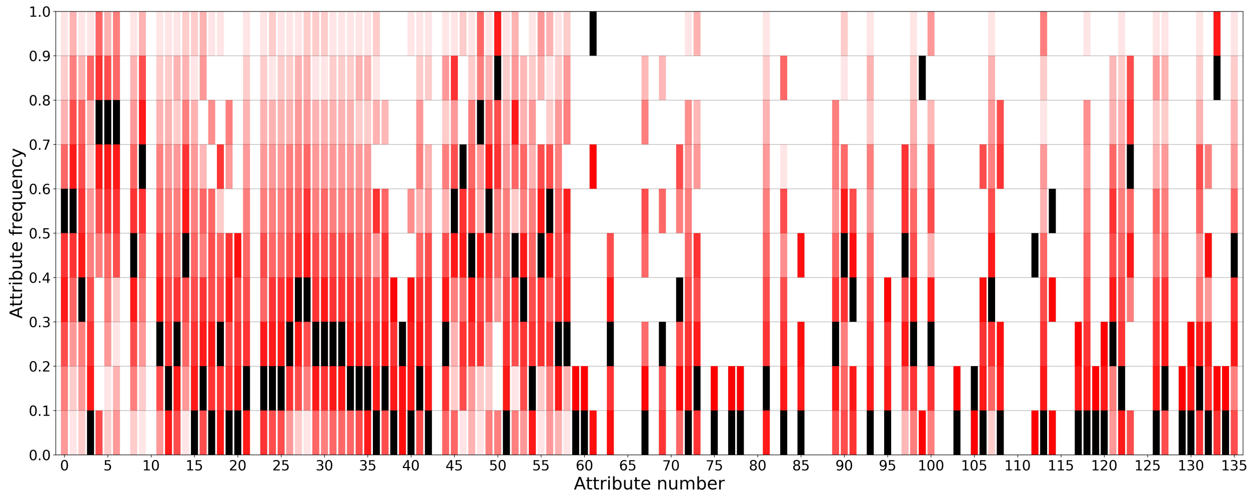

ARC
MOR
OIB
Ti/V=10
Ti/V=20
Ti/V=25.92
Ti/V=41.43
Ti/V=49.47
Ti/V=50
Ti/V=61.41
Ti/V=70.61
Ti/V=100
Ti/V=166.29
V (ppm)
Ti (ppm / 1000)
OIB
MOR
ARC


ARC
MOR
OIB
Ti/V=10
Ti/V=20
Ti/V=50
Ti/V=100
V (ppm)
Ti (ppm / 1000)
OIB
MOR
ARC
0
100
200
300
400
500
600
0
5
10
15
20
25
30
35
40
45
50

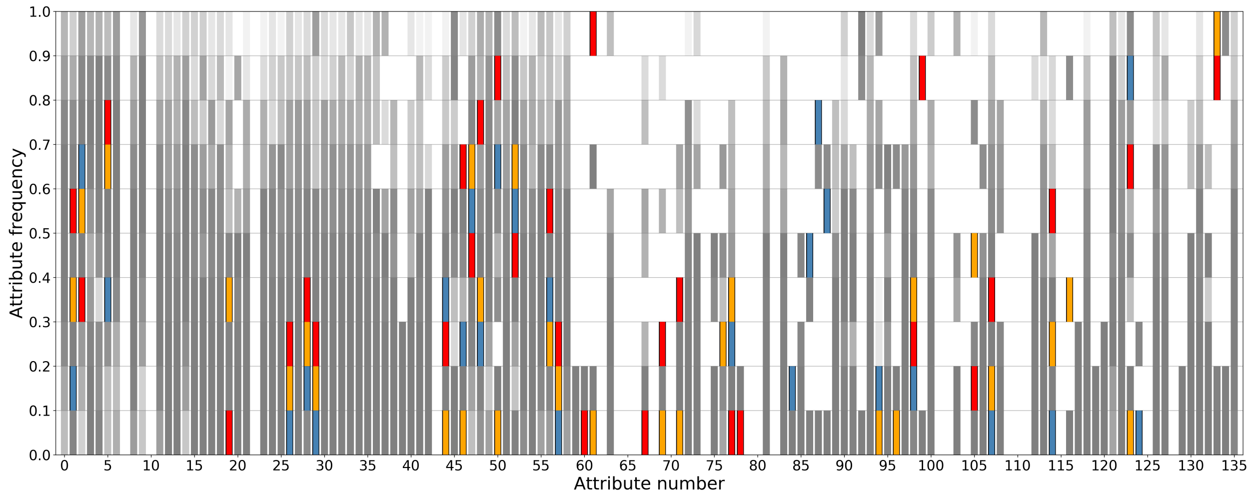

Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
73
74
75
76
77
78
79
80
81
82
83
Successful predictions (%)

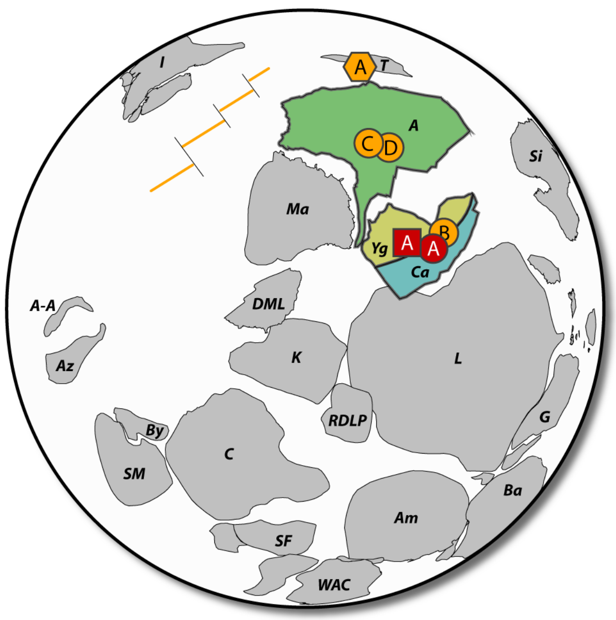

Attribute frequency
Attribute number
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135


Attribute frequency
Attribute number
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135

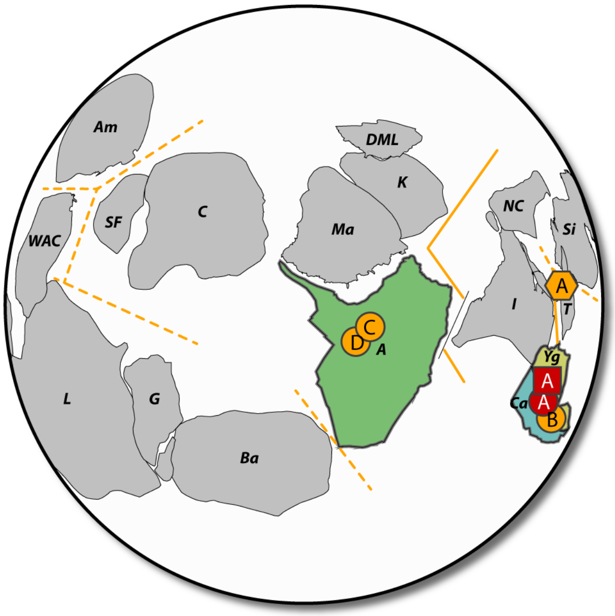


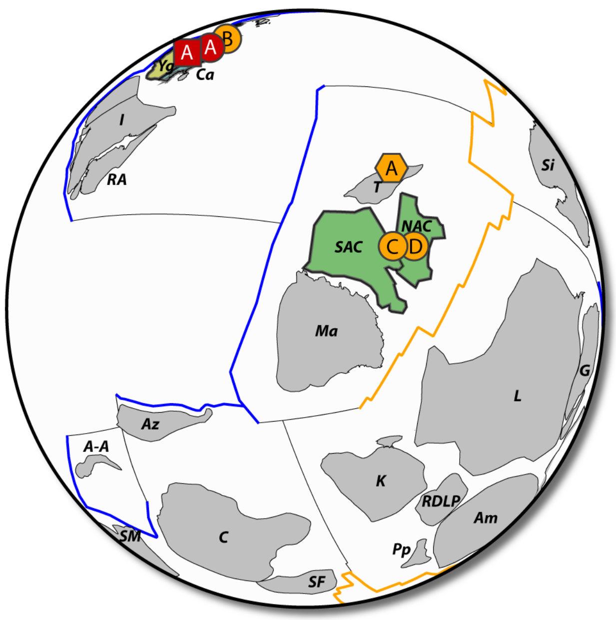
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Attribute frequency
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135
Attribute number

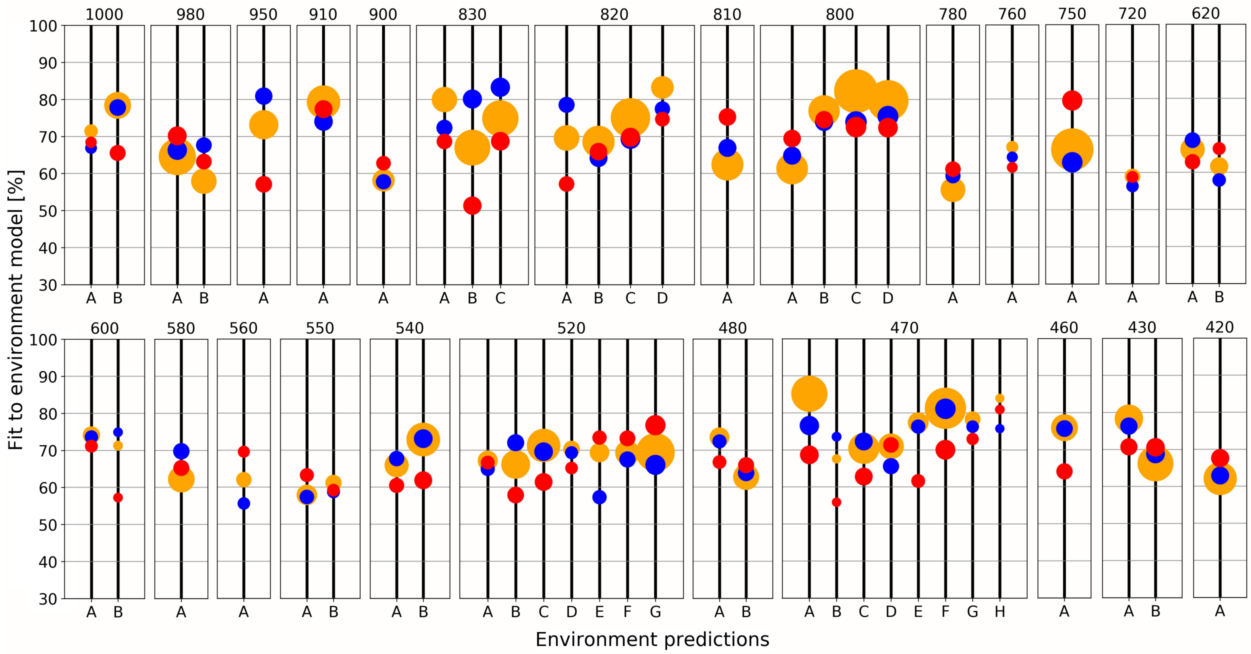

Attribute frequency
Attribute number
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135



I
A
T
A
C
D
Si
Ma
B
A
A
Yg
Ca
A-A
DML
K
L
Az
G
By
RDLP
C
SM
Ba
Am
SF
WAC


Am
DML
K
C
SF
WAC
Ma
NC
Si
I
T
A
C
D
A
Yg
A
Ca
A
B
L
G
Ba


A
A
B
Yo
Ca
I
RA
A
T
Si
NAC
SAC
C
D
Ma
L
G
Az
A-A
K
RDLP
Am
SM
C
Pp
SF



1000
980
950
910
900
830
820
810
800
780
760
750
720
620
600
580
560
550
540
520
480
470
460
430
420
Fit to environment model [%]
Environment predictions

- New method leveraging big data to characterize first-order tectonic environments
- Identification of 115 key discriminatory geochemical attributes in basaltic rocks
- Tectonic fingerprints used to evaluate and constrain supercontinent cycles