Auxiliary Material

Supplements 1-4

Klein et al.; 2013GC004858

Supplement 1: Evaluating the validity of photograph-based relative age scale with U-Th radiometric age measurements.

To evaluate the validity of our relative age scale, we draw upon data from a portion of the EPR north of our study area where absolute lava ages have been measured radiometrically and where the seafloor was photographed during lava collection. S*ims et al*. [2002, 2003] report U-Th radiometric ages for lavas collected using DSRV Alvin submersible on- and off-axis along the EPR at 9°48–52'N. Based on a comparison of the radiometric ages and ages calculated on the basis of spreading rate (distance off-axis), they concluded that a large number of off-axis samples were anomalous in that they had erupted more recently than one would predict based on distance off-axis. However, at least three off-axis samples had radiometric ages consistent with their spreading rate age. These three samples, combined with two axial samples from the same study, were used to explore the validity of our relative age scale and estimate absolute age based on seafloor photographs.

Using archived photographs of seafloor near where each of the five samples was collected (Figure A1; *D. Fornari, pers. comm.*), we assigned a relative age rank to each, and compared these to radiometric ages [*Sims et al*., 2003] reported for the five samples. Strong linear correlations are found between our relative ages and U-Th trendline ages (Figure A2). Similarly strong correlations are found between spreading rate age and both relative age and radiometric age.

For the data of *Sims et al*. [2003], linear regression suggests that each unit increase in our relative age rank represents an increase of ~18,000 years in absolute age. Importantly, however, for the 9°50'N samples, as well for samples reported by Waters et al. [2013a,b] our photographic estimates of ages for young seafloor (relative age ranks 1 or 2) appear

to less accurate; as shown in Figure A2, for example, seafloor with relative age ranks 1 and 2 have lavas with U-Th trendline ages reported as zero.

While the strong correlation between relative age rank and radiometric age supports the use of photographic age estimates, the applicability of this relative age-radiometric age calibration to other areas should be approached with caution because the visual characteristics of the seafloor will be subject to local variables, such as sedimentation rate, differences in lava morphology, terrain slope (affecting sediment thickness), tectonism (affecting outcrop integrity), or hydrothermal circulation (affecting weathering rate).

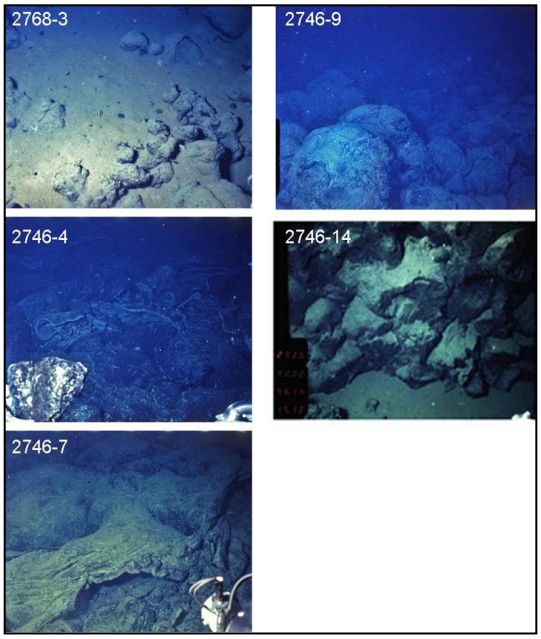

Figure S1. Photographs of seafloor near where samples were collected in the study of the 9°50' N EPR region reported by *Sims et al*. [2003], with additional photographs provided by D. Fornari (*pers. comm.*) These photographs were used to classify the seafloor shown using our relative age scale. The relative ages were then compared to ages estimated on the basis of spreading rate and absolute ages based on U-Th dating (see Figure S2).

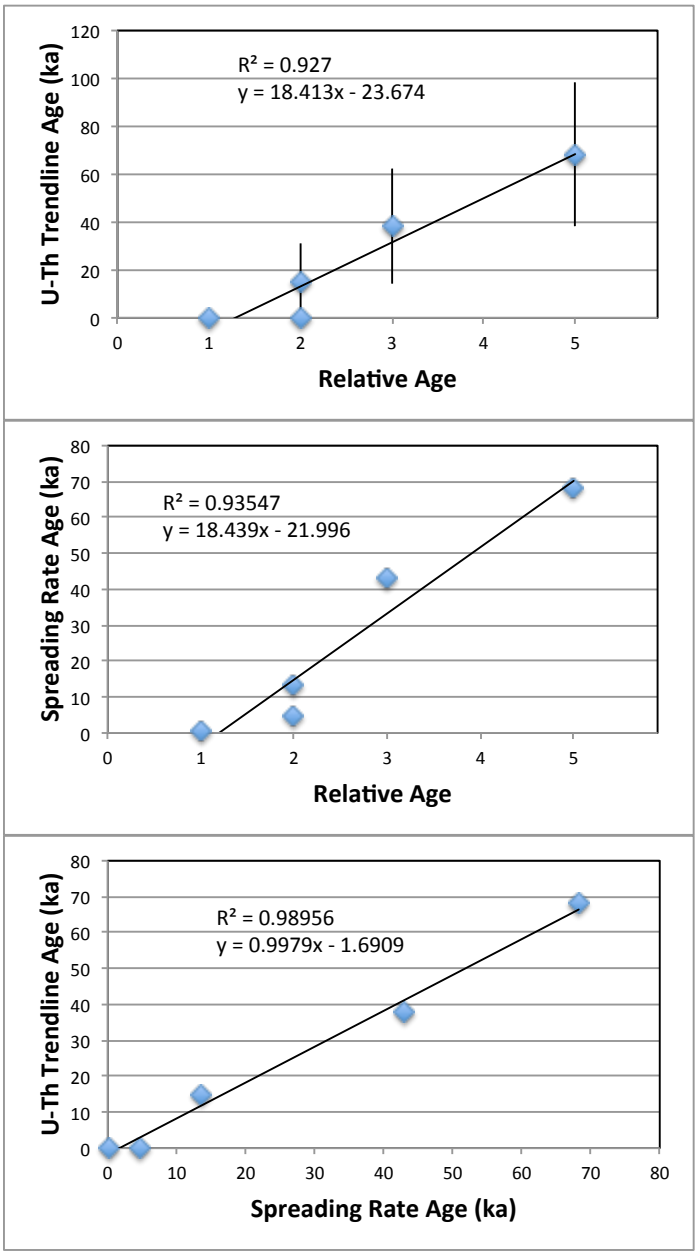

Figure S2. Age data for samples recovered on- and off-axis along the EPR at 9°48–52'N. The five samples shown were judged by *Sims et al*. [2002, 2003] to be not anomalously young, in contrast to other off-axis samples in their study. The three plots show the correlations among relative ages (determined using our relative age scale), spreading rate ages, and U-Th trendline ages [*Sims et al*., 2003].

Supplement 2: Bathymetry and digitizing seafloor geologic features

Two existing data sets from this area were merged to create the bathymetry used in this study. *White et al*. [2006] collected a Simrad EM300 bathymetry swath through the area that was gridded at 30m x 30m, and *Combier et al*. [2008] used seafloor reflections from the ARAD seismic survey to create a 25m x 100m bathymetry grid. We resampled the *Combier et al*. grid using cubic convolution to match the EM300 grid, then overlaid the EM300 data values, and clipped the grid to match the boundaries of the DSL-120a backscatter mosaic. To eliminate sharp depth changes due to slight mismatches at the boundary between bathymetric datasets, we used a weighted average of the values within stripes running along the outer 400m (10%) of the EM300 swath. The grid values were averaged using a weight assigned to each dataset as a linear function of distance to the edge of the EM300 swath from 50% to each dataset at the edge to 100% EM300 at 400m inside.

This bathymetic grid was median filtered using a 2 km^2 moving window, and then subtracted from the original grid removing relief associated with the axial highs and basin, to create a grid of local relief. From the local relief grid we derived several topographic metrics: the mean of aspect in 500 m circular neighborhoods; range of local relief in 500 m circular neighborhoods; standard deviation of surface roughness in 700 m by 400 m neighborhoods; and mean of slope in 500 m circular neighborhoods.

Mounds, fissures and faults observed in backscatter images were identified and digitized as follows (Figure 7). The base of the mounds was found by viewing bathymetry simultaneously with backscatter. Closed contours (5m interval) were used to initially identify a mound and the base was digitized from a break in slope on the backscatter and bathymetry to retain the highest resolution possible. Fissures and faults were manually digitized by tracing the length of a paired bright-dark backscatter reflection in the case of fissures, or a single linear reflector coinciding with a bathymetric gradient for faults.

Supplement 3: Supervised and Unsupervised Automated Seafloor Classification

Supervised and unsupervised classifications were both used for this study in order to compare the supervised result based on predetermined, user-defined classes to the unsupervised result based on user-independent clustering of data values. Unlike typical supervised classifications, no definitive dataset describing different seafloor terrains exists for training the classifier. *Sempere and Macdonald* [1986a] provide the closest approximation to prior definitions by identifying OSC provinces, and so we adapted their provinces to train the supervised classifier. By comparing results from the supervised and unsupervised classifications on the same data, we can determine the extent to which our choice of training data has influenced the outcome (Figure 8a,b).

Unsupervised classifications receive no *a priori* information about what classes exist. In an unsupervised classification, the computer uses an algorithm, in our case the ISODATA clustering algorithm, that examines the raster stack and classifies the cells that meet criteria for similarity into the same class [*Jensen*, 2005]. The user specifies the number of classes that should be contained in the output, and classification runs until a convergence threshold or the maximum number of iterations has been reached. Several unsupervised classifications were run, with different numbers of classes specified at input, to determine the spatial pattern and number of classes that appeared consistently.

We initially selected an 11-class output from the unsupervised classification, which produced incoherent patterns of mixed pixel values within larger blocks of only a few classes. As the number of classes was reduced, larger blocks of pixels consisting of the same value appeared. When these classes were combined, large areas containing only one class were generated, and we used these as the final result for comparison to the classes produced from the supervised classification. The unsupervised classification of texture and residual bathymetry produced only four classes (Figure 8b).

A supervised classification uses specified data (e.g., terrain areas) chosen by a user to describe a particular signature, to train the computer to classify on the basis of the

different characteristics. The computer then uses these characteristics to assign terrain classes to every pixel in the study area, based on seafloor texture from side-scan, and seafloor relief, slope and aspect from bathymetry in our case, to identify other seafloor areas with matching characteristics. Areas of interest (AOIs) are polygons created manually that only contain points representing a single class based on *a priori* understanding of the classes that the user wants the classifier to find. We created five AOIs derived from the five provinces defined by *Sempere and Macdonald* [1986a] for the supervised classification. AOIs were drawn approximately centered within and to half the total length and width of the provinces in *Sempere and Macdonald* [1986a]. Maximum likelihood calculates the probability that a cell belongs to each class from maximum likelihood probability statistics and assigns the cell to the class with the highest probability.

The supervised classification also found four specific provinces (Figure 8a), despite the fact that the classifier was trained on five different AOIs, corresponding to the five provinces of *Sempere and Macdonald* [1986a]. This illustrates the fact that tectonically defined provinces may contain a mixture of distinct seafloor terrains (see, for example, the complex mix of terrains found in the tectonically defined "eastern ridge tip." Names for each province were chosen based on the gross morphologic characteristic of each province: Smooth, Hummocky, Fissured, and Mounded provinces.

To produce the final terrain classification maps (Figure 8), we applied a boxcar mode filter over a 7x7 cell (210 x 210 m) area. This removed isolated individual or small groups of pixels from within larger areas of a different class. Two iterations of this filter were applied. The border of the final map was trimmed by one filter width (210 m) to eliminate edge effects due to the window size applied during data processing.

Supplement 4: Accuracy Assessment of terrain classification

In order to assess the internal consistency of the supervised classification method, an accuracy assessment was performed by comparing the output cell values from the classifier within each AOI to the value we assigned to that AOI. The AOIs were chosen

somewhat *ad hoc* from the areas picked by *Sempere and Macdonald* [1986a] to define a supposedly uniform terrain signature. It was unclear at the outset if these AOIs would be appropriately distinctive; by comparison, traditional training data for a supervised classification generally represent well-agreed-upon categorical features (e.g. streets, buildings, forests). The purpose of this accuracy assessment is to determine to what extent each AOI contains a unique terrain signature. If the AOI represents a unique and distinct terrain province, then the classifier should reproduce the uniform value of the cells chosen by the user for that AOI. Unlike traditional accuracy assessments, it does not provide an estimate of the accuracy of the classification of regions outside of AOI.

To perform the accuracy assessment, we used the direct output from the supervised classifier before the mode filter smoothing was applied. User's and producer's errors were calculated from the class chosen automatically by the supervised classifier and our pre-assigned AOI class (Table D1). The Kappa coefficient uses the number of cells classified correctly within each AOI and the total number of cells classified that represent a specific province that are found over the entire classified image. Kappa value of 0.79 was obtained for the supervised classifications, which falls at the boundary between "strong" and "very good" agreement between the predefined areas and the classifier output (*Jensen*, 2005). The result indicates that the classifier consistently reproduces the classes chosen to represent distinct terrains in four out of five areas, with the fifth area being consistently identified as one of the other four.

The high Kappa value and close correspondence between the supervised and unsupervised classifications indicates the four provinces are robust representations of the local terrain types on a landscape scale. The precision of the boundaries of the provinces is a function of the size of the analysis window, which is 1 x 2 km in this study. A kilometer level of precision seems appropriate for this study since the tectonic and magmatic processes attributed to creating these provinces are such that transitions would be likely to occur on a similar scale.

Province Cells	Basin AOI	East Limb Tip AOI	West Limb AOI	West Limb Tip AOI	East Limb AOI	*TOTAL*
Basin	48325	0	15	1609	1	49950
East Limb Tip	-	-	-	-	-	0
West Limb	0	36797	56906	1530	288	95521
West Limb Tip	171	6169	1143	68973	6154	82610
East Limb	2	173	0	1600	122851	124626
TOTAL	48498	43139	58064	73712	129294	*352707*

Table S1. Contingency table created using the output from the supervised terrain classification (rows) and the user-defined areas of interest (AOI) (columns).