SUPPLEMENTARY MATERIAL FOR:
Evidence of an axial magma chamber beneath the ultraslow spreading Southwest Indian Ridge

Hanchao Jian\textsuperscript{1,2}, Satish C. Singh\textsuperscript{3*}, Yongshun John Chen\textsuperscript{1,3*} and Jiabiao Li\textsuperscript{4}
\textsuperscript{1}Institute of Theoretical and Applied Geophysics, School of Earth and Space Science, Peking University, Beijing 100871, China
\textsuperscript{2}Laboratoire de Géosciences Marines, Institut de Physique du Globe de Paris, 1 rue Jussieu, 75238 Paris Cedex 05, France
\textsuperscript{3}School of Oceanography, South University of Science and Technology of China, Shenzhen 518055, China.
\textsuperscript{4}Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012, China.
*E-mails: johnyc@pku.edu.cn (Y.J.C); singh@ipgp.fr (S.S.)

SUPPLEMENTARY METHODS

Seismic experiment and data
The 3-D seismic dataset covers an about 70 km by 70 km box, and consists of 13 OBS covering the central volcano and a total of 5765 air-gun shots (Li et al., 2015) (Fig. 1b). Each air gun within the 4-airgun array had a volume of 1500 in\textsuperscript{3}. Since the volume was constant for each air gun and thus the bubble sequence was not suppressed, the array provides significant low-frequency energies (3-8 Hz), which are very important to acquire very long-distance records. Shots were fired every 2 minutes (~240 m) along a grid of lines spaced at ~4 km in the 3-D box. The raw data has a sampling interval of 4 ms and was converted to SEG-Y format with a length of 40 s for later analysis.

Full waveform inversion
Unlike travel time tomography that tries to match the travel time information using synthetics from high-frequency approximated ray tracing, full waveform inversion (FWI) matches the waveform using synthetics derived by solving the dynamic wave equations. Since the amplitude information is added, less approximation is involved and no explicit picking is needed (Shipp and Singh, 2002), FWI provides a higher resolution than tomography, especially for imaging low-velocity anomalies that strongly attenuate the traversing seismic energies and thus prevent effective travel time picking (Wang et al., 2014). The largest shortcoming of FWI is the non-linearity problem that can trap the solution into local minima, but it is natural to refine the model by starting from a long-wavelength model constrained by travel time data.

We adopted a 2-D time-domain elastic FWI approach firstly created by Shipp and Singh (2002), which uses a finite difference method (Levander, 1988) for waveform modeling and the adjoint method (Tarantola, 1986) for gradient calculations. The misfit function is defined as:
\[ S = \sum_{s} \sum_{r} \int_{0}^{T} (\mathbf{d}_{\text{syn}} - \mathbf{d}_{\text{obs}})^T \mathbf{W}_{d}^{-1} (\mathbf{d}_{\text{syn}} - \mathbf{d}_{\text{obs}}), \]

where \( \mathbf{d}_{\text{syn}} \) and \( \mathbf{d}_{\text{obs}} \) represent the synthetic and observed seismic waveforms, respectively.

The sample-by-sample differences are summed over the recording time (\( T \)), and all shots (\( s \)) and receivers (\( r \)). The \( \mathbf{W}_{d}^{-1} \) is the data covariance matrix, which we made a diagonal matrix to weight different parts of the data.

The across-axis profile running around the central volcano (Fig. 1b) was chosen to improve the resolution of the LVA image. Before the FWI, the observed data was pre-processed aiming to reduce the inconsistency between the 2-D waveform modelling and the 3-D nature and to enhance SNR over a broad frequency band. We first corrected the crooked geometry by shifting the time axis based on the results from ray tracing in the 3-D tomography model. This correction is accurate enough for the long-offset crustal phases (Pg and PmP). Then we designed a debubbling filter using the approach of Wood et al. (1978) to suppress the bubble energies and thus flatten the amplitude spectra between 3 and 8 Hz. We have not involved the anelasticity in the FWI. Instead we estimated the long-wavelength attenuation structure from the data and prior information, and then compensated for its effect in the data domain.

In the inversion, we have only used the long-offset primary phases in the hydrophone channel (Fig. 2, Fig. DR5), since this part of data has the most linear behavior in FWI and the largest sensitivities to the lower-crustal structures. Within this subset of data, there are basically no S-wave energies and thus it is much less sensitive to the S-wave velocities compared to P-wave velocities. Hence, only the P-wave velocities that were initially extracted from the 3-D tomography model have been iteratively updated to minimize the data misfit, whereas the S-wave velocities were always related to the P-wave velocities according to the Brocher’s regression fit (Brocher, 2005). The finite difference modeling used a grid spaced at 30 m, which can assure the accuracy for propagating P waves with a frequency below 10 Hz.

During the inversion, a Gaussian smoothing filter (750 m horizontally and 210 m vertically) was applied to suppress high-wavenumber artifacts. Besides, we forced the model update within the lower crust (between 4 km and 10.5 km depth below the sea level) to be symmetrical with respect to the ridge axis, which is a generally valid assumption considering the symmetrical spreading.

In order to examine the resolvability of the sparse dataset, we performed extensive synthetic tests using the real-data geometry and source signature, and a synthetic AMC model. We created the synthetic AMC model by adding an AMC-like anomaly (Fig. DR4a) and a constant velocity contrast of 0.75 km/s across the Moho to the tomography model (Fig. 2a). The input data was generated from the synthetic AMC model with the same finite difference modeling as in the FWI. The inversion was carried out using the same time and offset windowing and inversion strategies as in the real-data inversion, with no Moho velocity contrast in the starting model (Fig. DR4a). The result (Fig. DR4b and c) indicates that the AMC-like anomaly can be reconstructed with a good vertical resolution, but smeared out laterally due to the limited OBS density. Consequently, the dimension of the anomaly is over-
estimated, whereas its amplitude is under-estimated. Besides, due to the limited frequency band and aperture of the input data, an artifact of positive velocity update exists above the AMC. However, the amplitude of this positive update is less than 1/3 of that of the negative update atop the AMC.

Besides, in the synthetic FWI test, the sharp velocity variation across the Moho interface is partly reconstructed within -10 and 10 km distance, due to the presence of strong PmP reflections. However, PmP reflections are simultaneously affected by the lower-crustal velocities and Moho reflector structures (depth and contrast). The trade-off between these structures is difficult to be eliminated because of the sparse data coverage within the lower crust. In addition, since the data does not contain significant Pn energies that travelling below the Moho interface, the long-wavelength structure across the Moho is not constrained. Consequently, the mantle velocity would be under-estimated, and a negative update immediately above the Moho always accompanies the positive update of mantle velocities.

In the real-data FWI, the Moho structure is inaccurate in the starting model, as the 3-D tomography provided no information in the Moho contrast at all and the accuracy of Moho depth could also be low because of the large picking uncertainties of PmP arrival times. These inaccuracies might be mapped to the lower crust through PmP wave paths, which can be observed in the synthetic FWI test (Fig. DR4b). To identify this effect in the real-data FWI, we ran inversions from various starting models where the Moho depth and contrast had been perturbed. By comparing their model updates, we found that the AMC anomaly is constant. Some artifacts may distribute along the PmP wave paths, especially within 2 km above the Moho interface. In this paper, we show the result of the real-data FWI using the starting model without Moho velocity contrast. Despite a good data fit for both Pg and PmP arrivals from the final model, we only focus the interpretation on the AMC structure.

REFERENCES CITED


**Figure DR1.** Bathymetric map showing locations of data and model profiles. OBSs are shown as circles, and those shown in Fig. DR2 are labelled with numbers over filled colours. The across-axis shooting profiles running over the five numbered OBS are shown with the corresponding colours. The thick dashed black square marks the location of the horizontal slice of 3-D tomography model shown in Fig. DR3e and f. The rest is the same as in Fig. 1.

**Figure DR2** (On next page). OBS data and travel time picking. Receiver gather for OBS a, 34, b, 35, c, 32 and d, 31 from the corresponding across-axis shooting profile, as shown in Fig. DR1. The red and black curves indicate the manually picked Pg and PmP arrivals used in the 3-D tomography, respectively. A water-path correction, a water-wave muting, a band-pass filter between 3-8 Hz and a time reduction with a velocity of 6.5 km/s were applied to all data sections. Note the time axis is different with that in Fig. 2 and Fig. DR4 due to the additional water-path correction. One can easily see the reduced amplitude of Pg resembling shadow zones (blue ellipses), which are followed by strong PmP reflections within the gathers (a and c) from the profile running over the central volcano (the green profile in Fig. DR1). However, there is no such feature for the record (b and d) from the other profile, suggesting the localized anomaly in the lower crust beneath the central volcano.
Figure DR3. Synthetic FWI test. The geometry and staring model are the same as used in the real-data inversion (Fig. 2a). The input data was generated from a model that was created by adding an AMC-like anomaly (a, true anomaly) and a constant velocity contrast of 0.75 km/s across the Moho to the starting model. The FWI was terminated when no significant data misfit reduction could be achieved. The final recovered anomaly relative to the starting model is shown in b (recovered anomaly). The green triangles mark the OBS locations. Note that the colour scales are different in a and b. Depth profiles of the anomalies beneath the ridge axis are compared in c. Note that the anomaly shown here was calculated with respect to the starting model, rather than the reference model shown in Fig. 3c. The recovered AMC anomaly is smeared out laterally, and thus has a larger dimension but smaller amplitude compared to the true anomaly, whereas the vertical resolution is good. Due to the limited frequency band and aperture of the input data, an artifact of positive velocity update exists above the AMC. However, the amplitude of the positive update is less than 1/3 of that of the negative update atop the AMC.
Figure DR4. Data fit. Long-offset waveforms of wide-angle refractions (Pg) and reflections (PmP) for the OBS used in the FWI: a, OBS 34 and b, OBS 32, whose locations are shown in Fig. DR1, whereas that for OBS 40 is shown in Fig. 2. In each three-row panel, top: pre-processed real data; middle: synthetic data for the starting model and bottom: synthetic data for the full waveform inverted model. The red curves mark the first-arrivals predicted with a ray tracing in the 3-D tomography model. The time is reduced with a velocity of 6.5 km/s. Both Pg and PmP are better fitted after the inversion as the AMC’s top and Moho contrast have been built up.
Figure DR5. Global seismic observations and numerical predictions of the depth to the top of AMC as a function of spreading rate (after Chen and Lin, 2004). The solid and dashed curves are model predictions with relatively low and high conductive cooling effect that is analogous to hydrothermal cooling in upper crust. Additional data is from the Lucky Strike segment (LS) (Singh et al., 2006).