Supplementary Materials for

Sequencing seismograms: A panoptic view of scattering in the core-mantle boundary region

D. Kim*, V. Lekić, B. Ménard, D. Baron, M. Taghizadeh-Popp

*Corresponding author. Email: dk696@cornell.edu

Published 12 June 2020, *Science* **368**, 1223 (2020)
DOI: 10.1126/science.aba8972

This PDF file includes:

Materials and Methods

Sdiff Waveforms

Broadband waveform data and synthetics were downloaded from the IRIS Data Management Center (DMC) and Instaseis databases hosted by Syngine (*36*), respectively. We used Mw > 6.5 earthquakes recorded by broadband seismic stations operating during 1990-2018. Instrument response was removed from the raw data. Tangential components of the displacement seismograms were bandpass filtered between 15 to 100 s using a Butterworth filter and windowed around the Sdiff phase (30s before and 100s after the Sdiff phase), as predicted by the 1D Earth model, PREM (11). Manual inspection of the selected Sdiff window was done to only include waveforms with clear Sdiff arrivals. To avoid any other bodywave traffic in the analysis window, we selected only events with hypocenter depth greater than 200 km. We then deconvolved synthetics from data using both time- and frequency-domain deconvolution, which removed radiation pattern effects and highlighted signals not predicted by PREM. We also compared our waveforms to those predicted by other 1D regional models (28, 37) for the Pacific basin (Fig. S2).

The Sequencer

The Sequencer is an algorithm designed to identify the leading trend in a dataset. It uses a graph-based approach to optimally order objects in a set such that the total similarity between all adjacent objects is maximized. The algorithm performs this search over a range of scales and using multiple metrics (the Euclidean Distance, the Kullback-Leibler Divergence, the Monge-Wasserstein or Earth Mover Distance, and the Energy Distance). For each of them, it estimates the level at which a continuous trend is present in the dataset by using the elongation of the corresponding distance matrix minimum spanning tree. It then aggregates information from all relevant scales and metrics by constructing an elongation-weighted distance matrix. The minimum spanning tree of this combined distance matrix provides an ordering of the objects in the dataset that corresponds to the final sequence. This approach is designed to define a view of the data, through the combination of different metrics and scales, that leads to the most elongated manifold. In contrast to the popular dimensionality reduction algorithm t-SNE (*38*, and Fig. S12), the Sequencer algorithm is fully deterministic and, in addition, it is capable of identifying scales or subsets of pixels that contribute to the signature of an underlying trend. The source code is available at http://github.com/dalya/Sequencer and an online version of the algorithm can be used at http://sequencer.org.

For our analysis, the trends were found using the Earth Mover Distance (*39*), and computed on two scales: the full waveform and 1/3 subsets. To emphasize the Sdiff postcursors and suppress apparent waveform differences due to travel-time variations, we aligned the deconvolved waveforms by cross-correlation and removed the mean trace from each waveform before sequencing. Postcursors sorted by the Sequencer become most apparent when we focus on the records after the main Sdiff positive pulse and apply histogram equalization (Fig. S13) to the records prior to inputting them into the Sequencer. The pixel-based histogram equalization here is analogous to applying automatic gain control often used in seismology to achieve optimized but nonlinear normalization. Once sequenced, we applied a running median filter to compute residual waveforms and discard as noisy those waveforms that have root-mean-square amplitudes larger than 1-sigma (Fig. S14).

We assessed Sequencer sensitivity to noise by deploying it on synthetic waveforms constructed from a real, randomly-selected waveform from our dataset contaminated by arbitrary

but realistic noise (e.g., uncorrelated white noise drawn from a normal distribution with standard deviation of the real waveform). We found that when the noise is bandpass filtered between 15 to 100 s (like our data) and possesses a dominant period similar to the duration of the signal being analyzed, the Sequencer detects trends even from noise (Fig. S15). Users must be aware of this fact and cautious when assessing robustness to noise in order to avoid identification of spurious trends.

Significance of Postcursor Waveforms

The order determined by the Sequencer reveals trends in waveform characteristics that allowed us to identify in ~40% of waveforms coherent post-Sdiff arrivals, which we call "postcursors" (Fig. 1B, red box). To quantify the characteristics of these postcursors and how they differ from the other 60% of waveforms, we measured the delay-time and amplitude of the largest post-Sdiff arrivals, relative to the main Sdiff phase. We found that the postcursor waveforms have distinct time-amplitude relationships, including, in the vicinity of Hawaii and Marquesas, significant anti-correlation between the two (Fig. 3), which is not seen in non-postcursor waveforms. Additionally, Fig. 4B shows that the 75^{th} percentile of amplitudes on non-postcursor paths is smaller than the 25^{th} percentile of amplitudes on postcursor paths.

It is possible that deconvolution side-lobes contribute to waveform amplitudes after Sdiff arrival. To quantify this contribution, and rule out deconvolution as the source of the postcursor signals, we carried out transdimensional Hierarchical Bayesian deconvolution (THBD: *40*), which enables amplitude uncertainty to be estimated. We computed ensemble solutions for 4 different Sdiff waveforms, chosen to be most highly correlated with the average postcursor waveform in Hawaii (Fig. S9A) and Marquesas (Fig. S9B and S14), and average of all non-postcursor waveforms (Fig. S9D). The fourth Sdiff waveform is randomly chosen from those showing prominent postcursors near Marquesas (Fig. S9C). We used one million iterations of the THBD algorithm to deconvolve the PREM synthetic from each waveform, discarding the first half as burn-in, and saving every 100^{th} sample to the ensemble. The posterior distributions we obtained demonstrate that postcursor amplitude cannot be an artifact due to deconvolution.

Geographic Patterns of Postcursors

Delay time and amplitude ratio measurements were made to observed postcursors with respect to the main Sdiff phases. These measurements were then assigned to the midpoint of the two D'' (the bottom 150 km of the mantle: *11*) piercing points for a given path (green, Fig. S16). For detected postcursors from the entire northern Pacific in particular, only the last 2500 waveforms (red box; Fig. 1B) in the sequence exhibit a postcursor that can be unequivocally distinguished from noise. To visualize the geographic patterns associated with delay-time and amplitude measurements and to compute bootstrap errors, we grouped postcursor measurements whose midpoints fall within a radius of 100 km of a chosen geographic location. To compute fraction of postcursors at a location, we counted waveforms whose paths travel in D'' and fall within 100 km of that location.

Postcursor measurements were also grouped in 5° bins across the Pacific basin based on the nearest point along their diffracting paths. For each geographical bin, the delay time measurements were binned in 2 s delay time windows, and the mean and its standard error were calculated (e.g., Fig. 4A). Using least squares minimization, we fit a line to the postcursor delay time and log-amplitude measurements averaged within each bin, and computed the corresponding slope and level of confidence (Fig. 3). To avoid overfitting, geographical and time bins containing three or fewer waveforms were excluded from the analysis.

Waveform Simulations

We tested various scenarios of the lowermost mantle by simulating waveforms using the three-dimensional spectral element method (*41*). The moment tensor solution for each event used in the modeling was obtained from the GlobalCMT catalog (*42*). We explored two types of models: ULVZs and sides of LLVSPs. We systematically varied the height and radius of, and wavespeed reduction in, a cylindrical ULVZ model (Fig. S6). We simulated head-on and edge-on scattering from the sharp sides of the LLSVPs, assuming a 5% Vs contrast (*9*) and large wavelength sinusoidal boundary undulations. For both cases, we used cosine tapers to implement smooth velocity variations across boundaries. Other models used for the forward modeling included PREM, SEMUCB-WM1, Hawaii ULVZ (*4*), a hypothetical Gaussian plume model 600 km in height, and a sharpened SEMUCB-WM1. To sharpen SEMUCB-WM1, we linearly interpolated shear velocity perturbations in the bottom 600km of the mantle and mapped them to an arctangent function of maximum and minimum value of +5% and –5%, respectively. The height of 600 km above the CMB is approximately the maximum height at which the average Vs of LLSVP regions substantially differs from non-LLSVP regions (*18*, *43*).

Differences in Effects of Fast and Slow Anomalies

High velocity anomalies can also scatter seismic waves, and may produce some of the weak-amplitude postcursors (pervasive, Fig. 1C) we observe, since these could be attributable to scattering from a compact anomaly. The normalized Rayleigh scattering coefficient for SH waves is unity (*44*), since our datasets are predominantly in the forward scattering regime. Therefore, an appropriately located high-velocity anomaly could produce postcursors similar to those from a low velocity anomaly (Fig. 4B).

However, generating the large amplitude postcursors, particularly beneath Hawaii and Marquesas, requires low-velocity compact anomalies based on two main reasons. First, low velocity anomalies can amplify the postcursors by focusing. Previous work has shown that focusing of seismic energy by a slow anomaly may be necessary to explain the large postcursor amplitude and their moveout (*4*). Second, waves transmitted through the anomalies and delayed by them play a significant role producing such large amplitude postcursors. Based on our waveform simulations, a low velocity cylindrical structure ~1000 km across and with a 20% Vs reduction can produce postcursors with an amplitude ratio greater than 1 with 25-30 sec time day (Fig. 4A). This is because what started out as the original Sdiff wavefront is delayed to produce postcursors, while what we interpret as the main Sdiff arriving before the postcursors is actually a relatively weaker phase produced by wavefront healing. On the other hand, a similar structure but with a positive velocity anomaly (20% Vs increase) is only capable to produce much smaller amplitude (~0.25) postcursors (Fig. 4A). In this case, what starts out as the original Sdiff wavefront remains the first-arriving Sdiff energy, and the relatively weaker delayed postcursor energy is produced by waves diffracting around the fast anomaly.

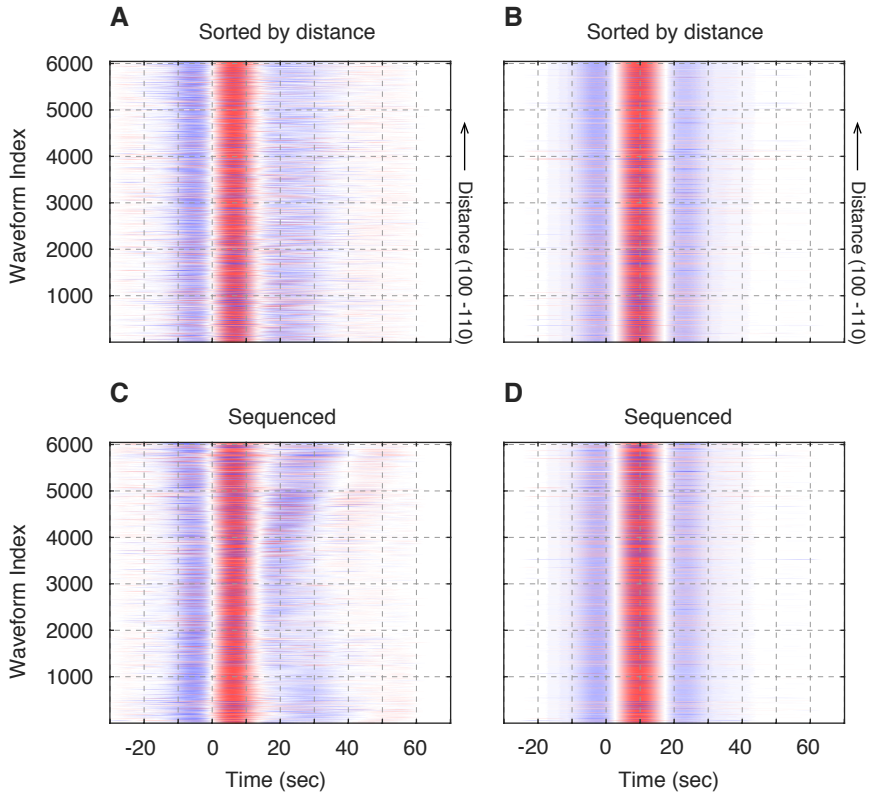

Fig. S1. Raw tangential seismograms (A) and synthetic waveforms computed for PREM (B) sorted by distance. (C-D) Same as A and B but sorted by the Sequencer.

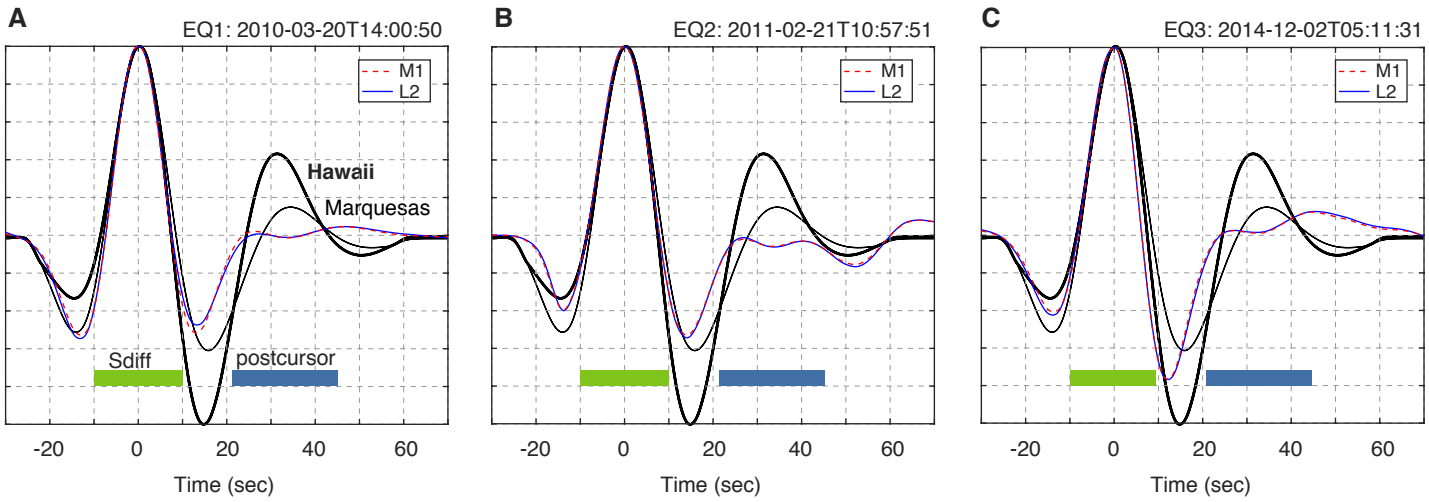

Fig. S2. (A-C) Comparison of average postcursor waveforms from Hawaii and Marquesas against average synthetic waveforms computed using two regional 1D models for the Pacfic basin (*28*, *37*). Three events used in the waveform simulation are shown in Fig. S8. At both hotspots, strong postcursors (indicated by blue horizontal bar) are apparent 30-40s after the main Sdiff arrival (green) for all three events, which are not predicted by the 1D models.

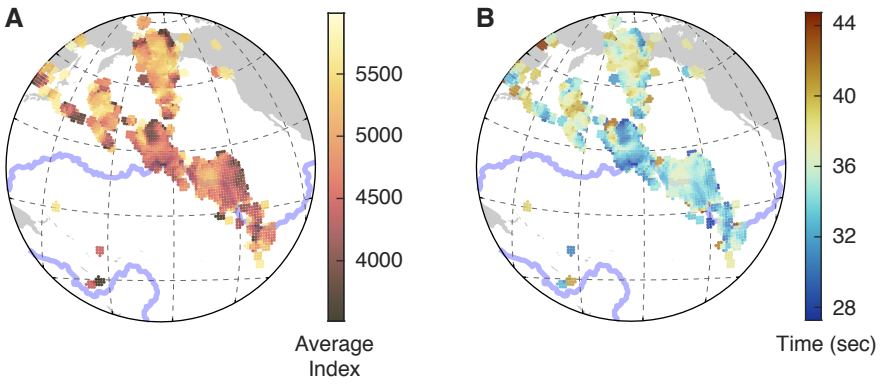

Fig. S3. Stack of average sequencing index and postcursor delay time relative to main Sdiff arrival averaged in 1° bins. The geographic extent of the Pacific LLSVP (*18*) is shown (light blue contour).

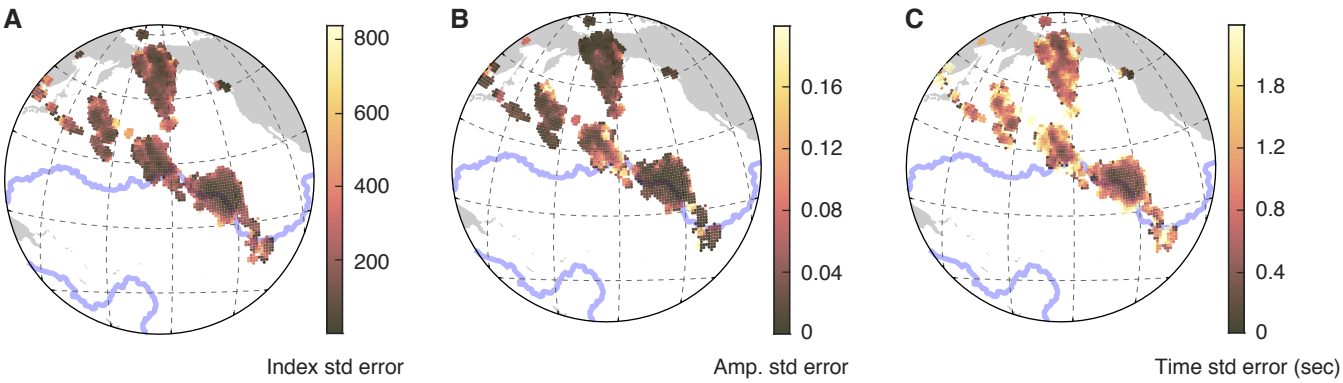

Fig. S4. Standard errors of the postcursor (A) sequencing index, (B) amplitude and (C) delay time relative to the main Sdiff arrival. The geographic extent of the Pacific LLSVP (*18*) is shown (light blue contour).

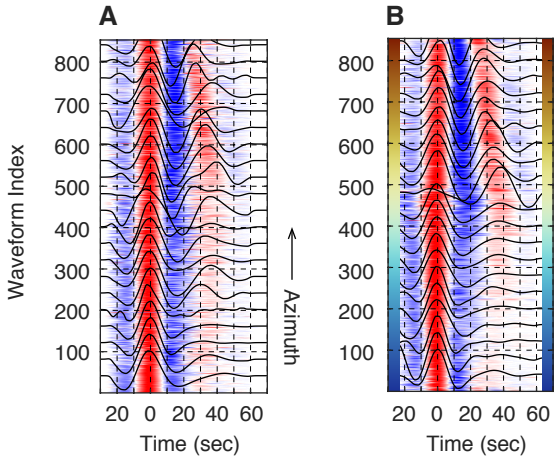

Fig. S5. Deconvolved Sdiff waveforms with turning points within 20° of Hawaii sorted by (A) azimuth and (B) the Sequencer.

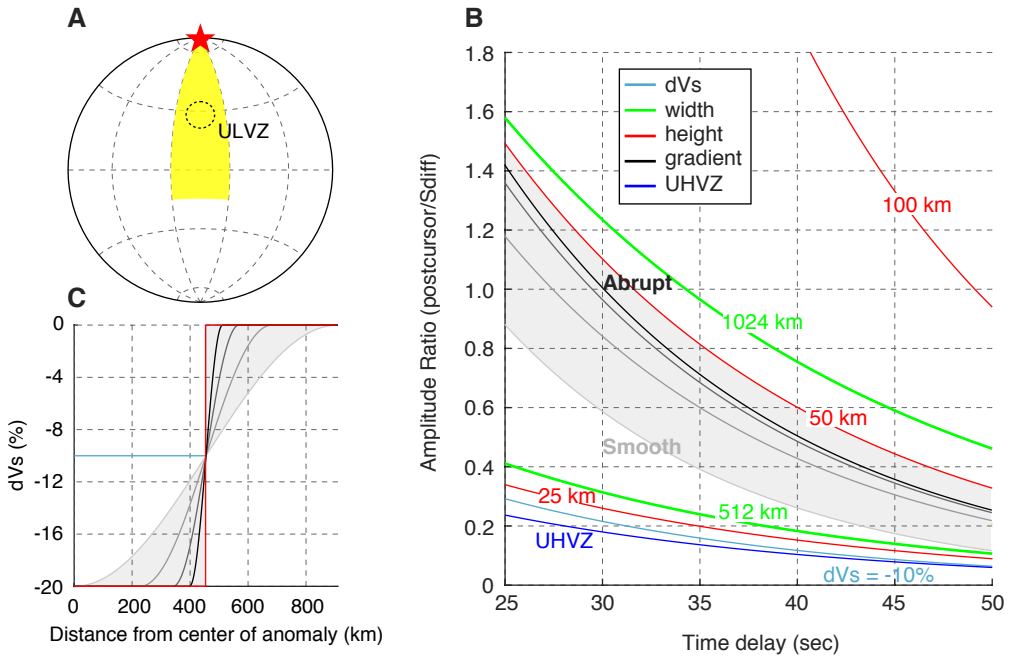

Fig. S6. Effect of width, height, velocity reduction, and velocity gradient on postcusor delay time vs. amplitude trends for a cylindrical velocity anomaly. (A) Illumination geometry. (B) Postcursor delay time and amplitude measured on deconvolved SPECFEM synthetics, with lines labeled according to the parameter being varied with respect to a reference model with: dVs = -20%; width = 910 km; height = 50 km; gradient = none. The ultrahigh-velocity zone (UHVZ) is identical to the reference model, except dVs = +20%. (C) As the gradient parameter increases from none (red), the velocity change across the edge of the cylindrical anomaly becomes smoother.

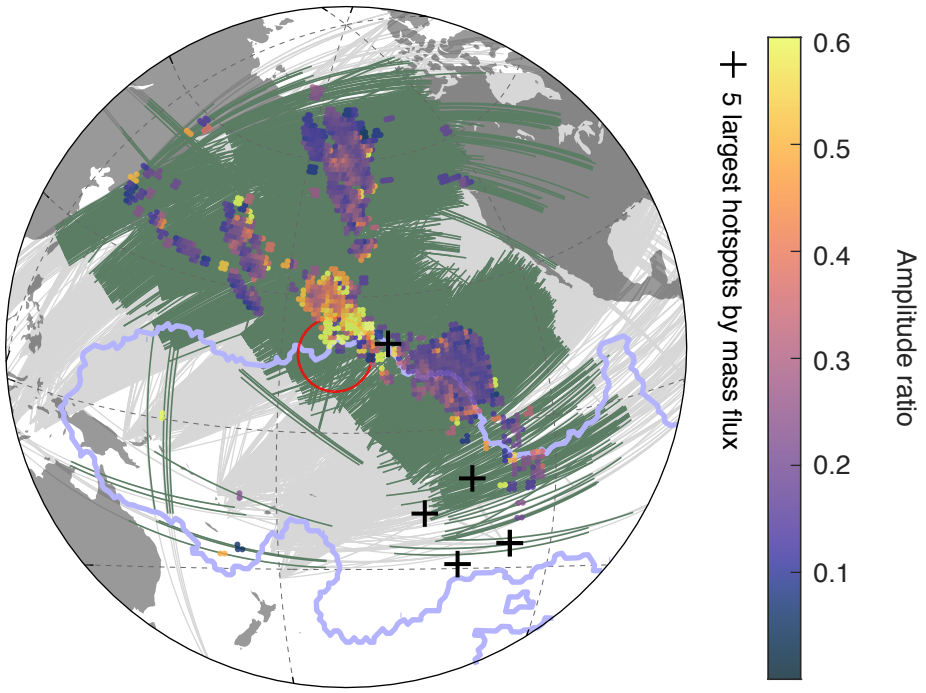

Fig. S7. Sdiff paths in D'' beneath the Pacific basin (dark green). The mega-ULVZ model (*4*) is outlined in red. All other symbols are same as in Fig. 1D.

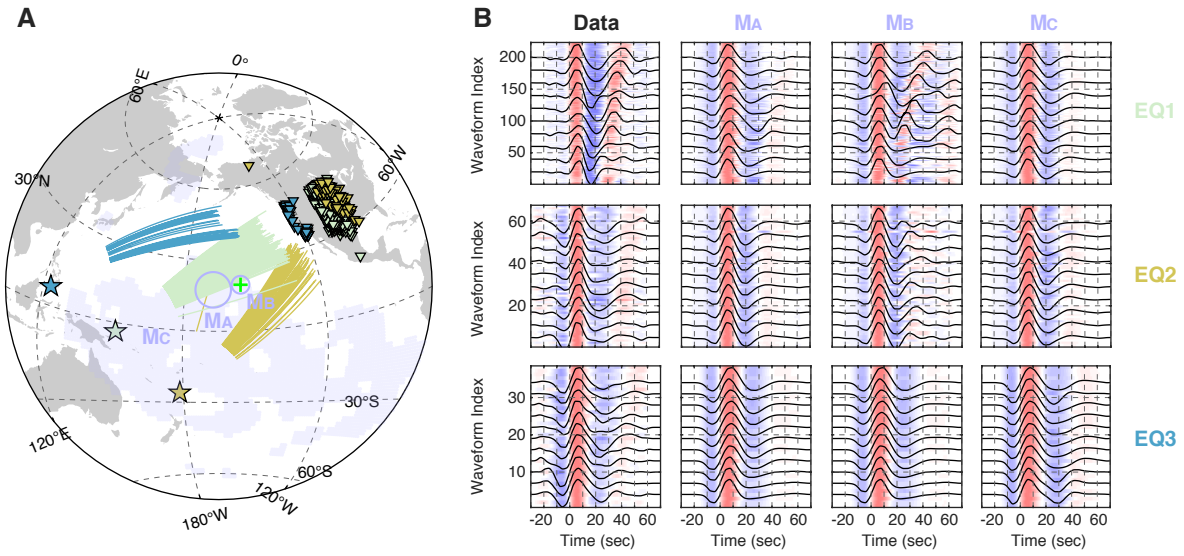

Fig. S8. Comparison of observed waveforms from three earthquakes with their corresponding synthetics computed for three different lowermost mantle structures. M_A is the mega-ULVZ model from (*4*), M_B is a 600 km tall, hypothetical gaussian plume centered beneath the Hawaiian hotspot (green cross), and M_C denotes a sharpened SEMUCB-WM1 (*30*) in the bottom 600 km of the mantle. See (*10*) for details. Earthquakes, stations, and paths are displayed on the map. (B) Sequenced deconvolved waveforms from both data and models aligned on the main Sdiff arrival.

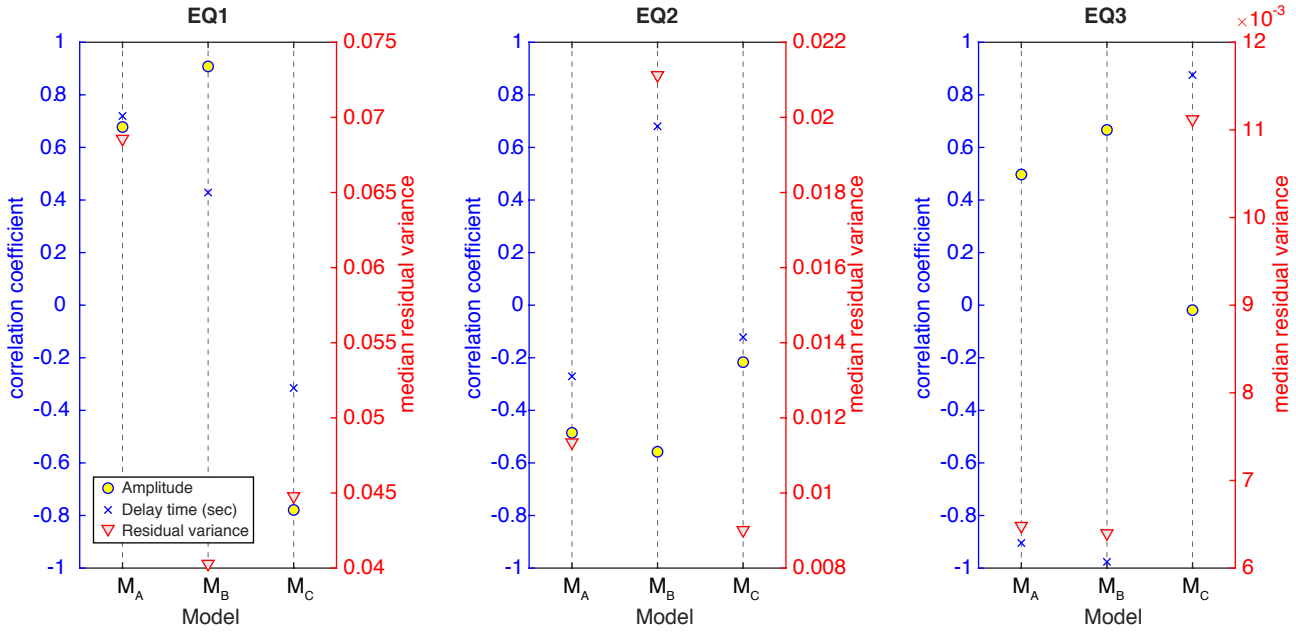

Fig. S9. Model fits for data vs. synthetic waveforms in Fig. S8. For each earthquake, we plot the correlation coefficients and residual variance (triangles) between data and synthetics comptued for each model. Correlation coefficients are computed separately for postcursor amplitude (circles) and delay time (crosses).

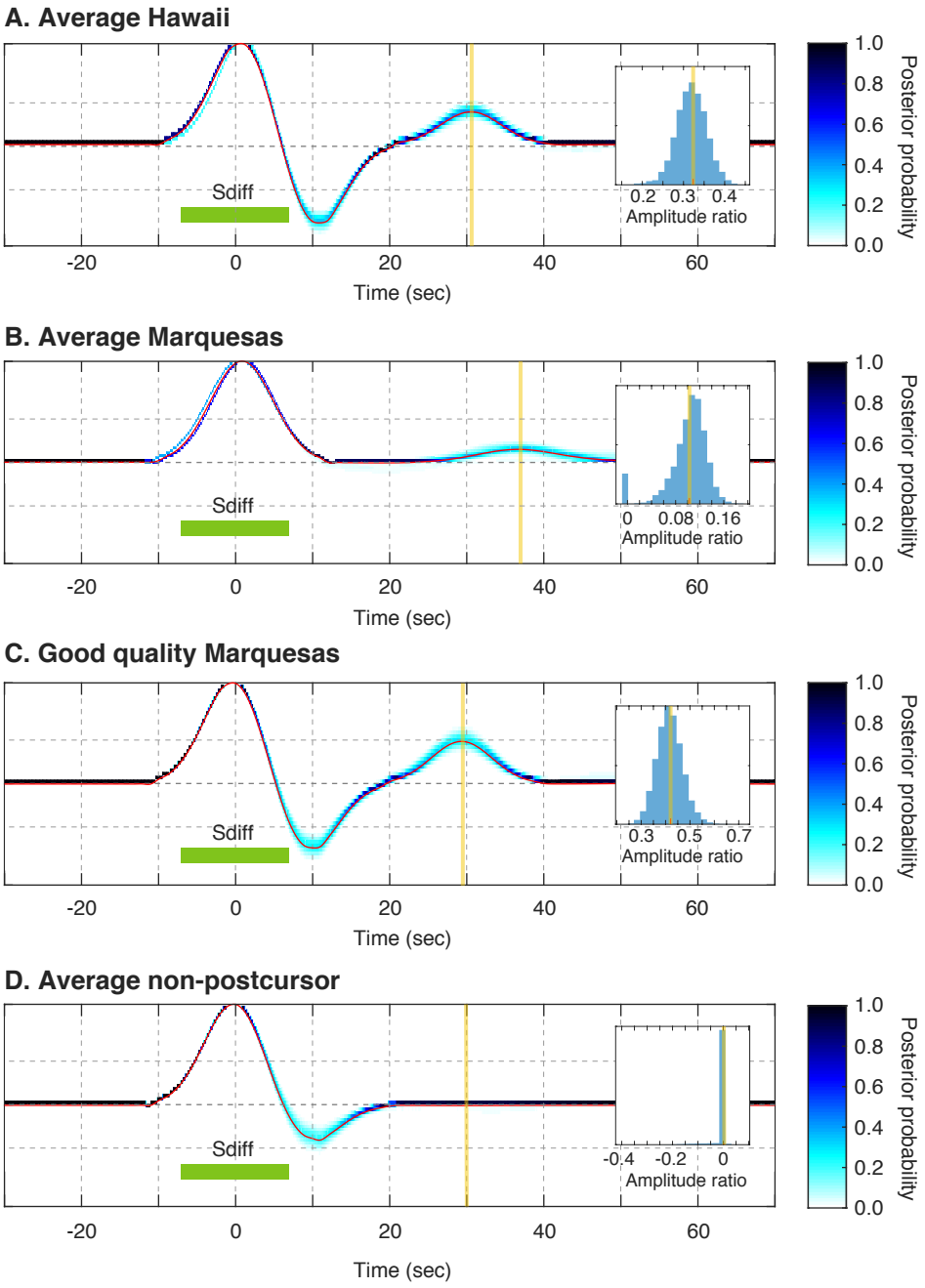

Fig. S10. Quantitative assessment of observed postcursors amplitudes. Ensemble solutions for deconvolved Sdiff waveforms and their associated uncertainties in (A) Hawaii, (B-C) Marquesas, and (D) those that do not show postcursors. One million iterations were carried out for each waveform using a transdimensional hierarchical Bayesian deconvolution (*40*). Yellow vertical line indicates the time at which the post-Sdiff amplitude is observed, and the inset shows the corresponding conditional probability density for amplitude.

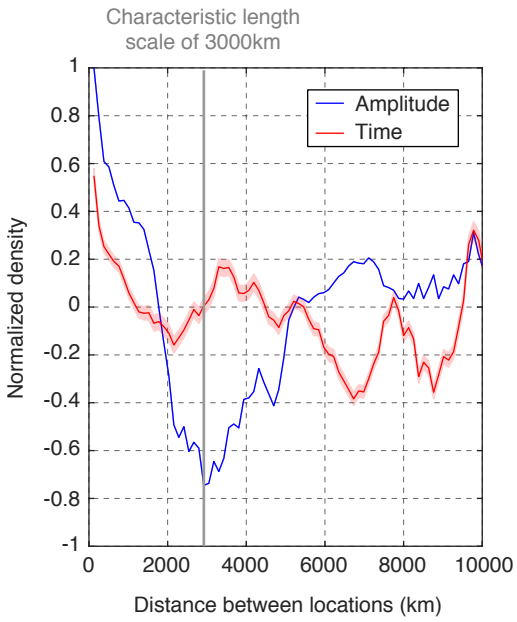

Fig. S11. Two-point correlation functions for amplitude (blue) and delay time (red) of postcursors shown in Fig. 1B. The correlation function for postcursor amplitude (blue line) is smallest at ~3000 km, corresponding to the characteristic length scale of the spatial pattern of postcursor amplitude variation across the Pacific basin.

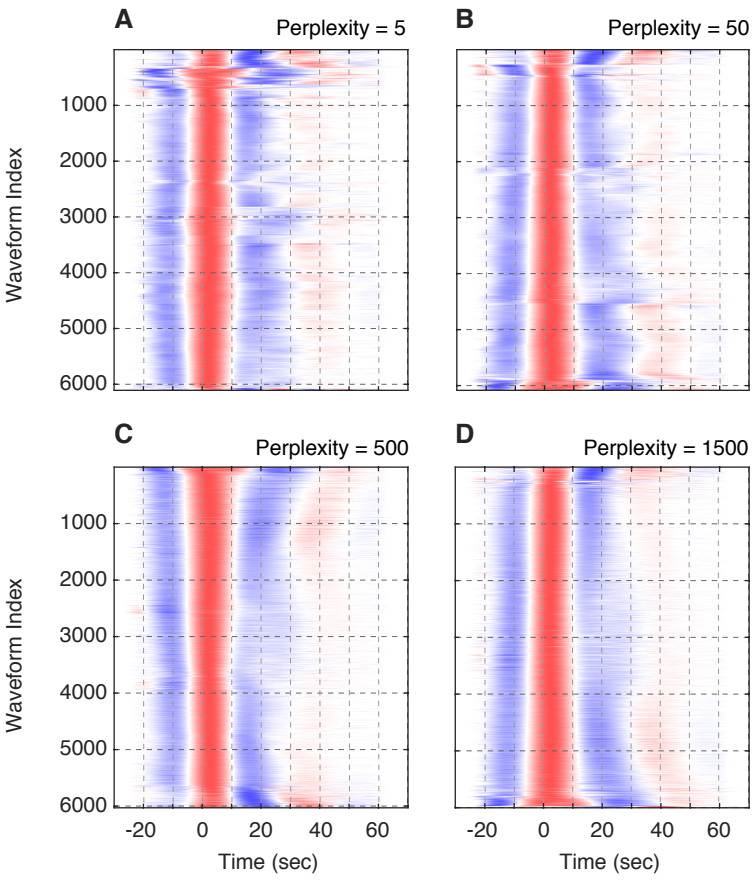

Fig. S12. Waveforms in Fig. 1B sorted by the t-SNE algorithm (*38*). Note that using the Sequencer is superior in detecting a clear trend of the postcursors (Fig. 1B). Although postcursors get more apparent with larger values of perplexity (A-D), the exact indices cannot be reproduced due to the stochasticity in t-SNE.

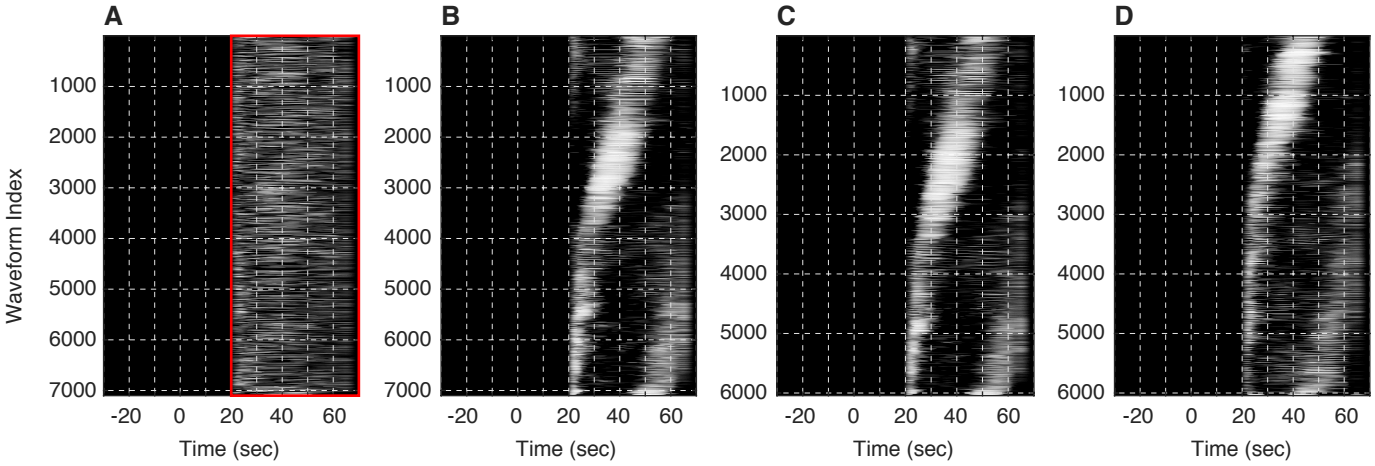

Fig. S13. Histogram equalized waveforms. (A) Historgram-equalized data after the main Sdiff positive pulse (red box) are used as the initial input for the Sequencer. (B) Same as A but sequenced. (C) Same as B but with outliers discarded according to procedure in Fig. S14. (D) Final sequence determining the order of waveforms shown in Fig. 1B.

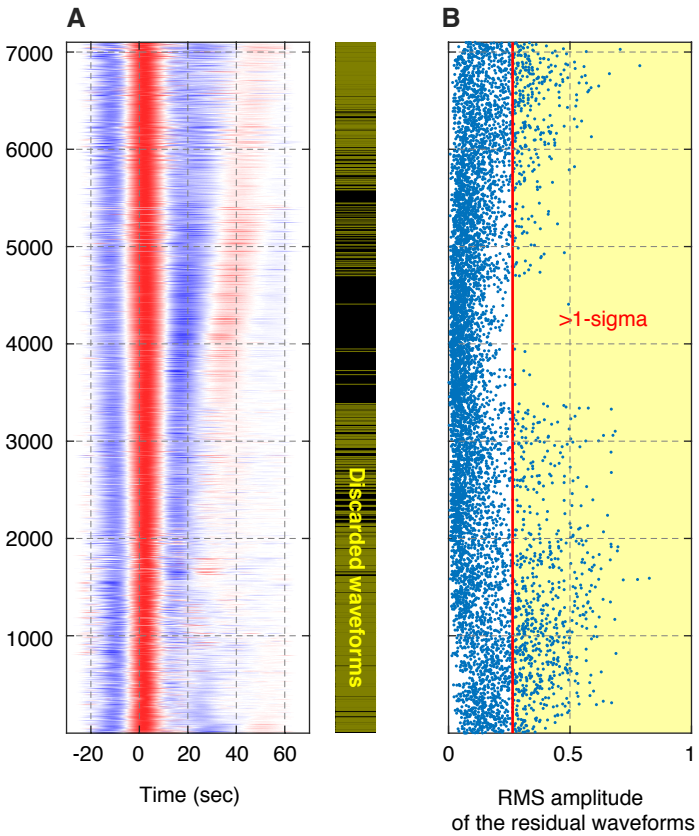

Fig. S14. Discarding outliers. (A) Sequenced Sdiff waveforms prior to discarding noisy data. Horizontal yellow bars indicate waveforms that are discarded. (B) RMS amplitudes of the residual waveforms with respect to a moving median of the sequenced waveforms. Red vertical line is 1-sigma RMS amplitude and waveforms to its right are discarded.

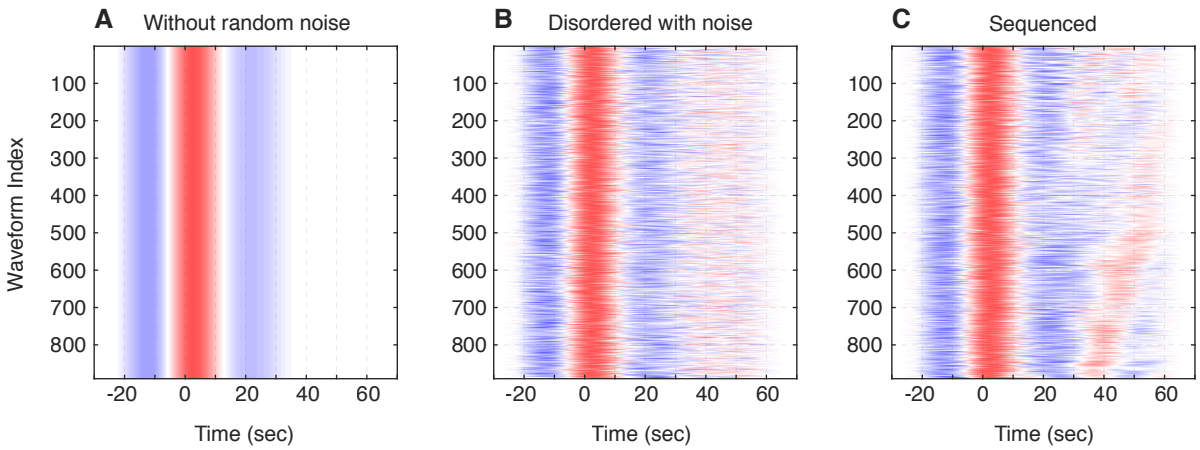

Fig. S15. (A) Randomly-selected waveform from Sdiff dataset. (B) Uncorrelated white noise added to A. (C) Same as B but sorted by the Sequencer.

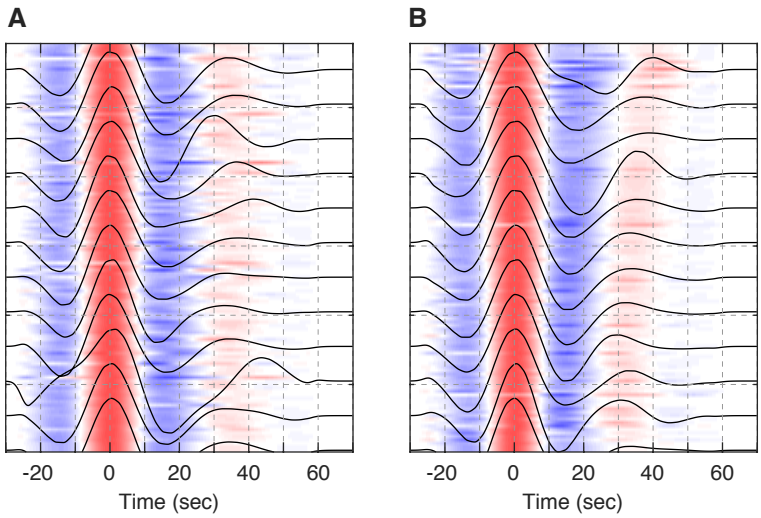

Fig. S16. Postcursors in the Marquesas region. Deconvolved Sdiff waveforms with turning points within 5° of Marquesas sorted by (A) distance and (B) the Sequencer.

References and Notes

1. A. M. Dziewonski, B. A. Romanowicz, Deep Earth seismology: An introduction and overview. *Treatise on Geophysics* **1**, 1–28 (2015). doi:10.1016/B978-0-444-53802-4.00001-4

2. T. Lay, Q. Williams, E. J. Garnero, The core–mantle boundary layer and deep Earth dynamics. *Nature* **392**, 461–468 (1998). doi:10.1038/33083

3. E. J. Garnero, S. P. Grand, D. V. Helmberger, Low P-wave velocity at the base of the mantle. *Geophys. Res. Lett.* **20**, 1843–1846 (1993). doi:10.1029/93GL02009

4. S. Cottaar, B. Romanowicz, An unusually large ULVZ at the base of the mantle near Hawaii. *Earth Planet. Sci. Lett.* **355–356**, 213–222 (2012). doi:10.1016/j.epsl.2012.09.005

5. L. Wen, P. Silver, D. James, R. Kuehnel, Seismic evidence for a thermo-chemical boundary at the base of the Earth's mantle. *Earth Planet. Sci. Lett.* **189**, 141–153 (2001). doi:10.1016/S0012-821X(01)00365-X

6. K. Yuan, B. Romanowicz, Seismic evidence for partial melting at the root of major hot spot plumes. *Science* **357**, 393–397 (2017). doi:10.1126/science.aan0760 Medline

7. A. To, Y. Fukao, S. Tsuboi, Evidence for a thick and localized ultra low shear velocity zone at the base of the mantle beneath the central Pacific. *Phys. Earth Planet. Inter.* **184**, 119–133 (2011). doi:10.1016/j.pepi.2010.10.015

8. A. To, Y. Capdeville, B. Romanowicz, Anomalously low amplitude of S waves produced by the 3D structures in the lower mantle. *Phys. Earth Planet. Inter.* **256**, 26–36 (2016). doi:10.1016/j.pepi.2016.04.001

9. A. To, B. Romanowicz, Y. Capdeville, N. Takeuchi, 3D effects of sharp boundaries at the borders of the African and Pacific Superplumes: Observation and modeling. *Earth Planet. Sci. Lett.* **233**, 137–153 (2005). doi:10.1016/j.epsl.2005.01.037

10. Materials and methods are available as supplementary materials.

11. A. M. Dziewonski, D. L. Anderson, Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981). doi:10.1016/0031-9201(81)90046-7

12. D. Baron, B. Ménard, Black hole mass estimation for Active Galactic Nuclei from a new angle. *Mon. Not. R. Astron. Soc.* **487**, 3404–3418 (2019). doi:10.1093/mnras/stz1546

13. Y. Xu, K. D. Koper, Detection of a ULVZ at the base of the mantle beneath the northwest Pacific. *Geophys. Res. Lett.* **36**, L17301 (2009). doi:10.1029/2009GL039387

14. D. Sun, D. Helmberger, V. H. Lai, M. Gurnis, J. M. Jackson, H. Y. Yang, Slab control on the northeastern edge of the mid-Pacific LLSVP near Hawaii. *Geophys. Res. Lett.* **46**, 3142–3152 (2019). doi:10.1029/2018GL081130

15. S. Cottaar, V. Lekic, Morphology of seismically slow lower-mantle structures. *Geophys. J. Int.* **207**, 1122–1136 (2016). doi:10.1093/gji/ggw324

16. E. J. Garnero, A. K. McNamara, S. H. Shim, Continent-sized anomalous zones with low seismic velocity at the base of Earth's mantle. *Nat. Geosci.* **9**, 481–489 (2016). doi:10.1038/ngeo2733

17. S. Yu, E. J. Garnero, Ultralow velocity zone locations: A global assessment. *Geochem. Geophys. Geosyst.* **19**, 396–414 (2018). doi:10.1002/2017GC007281

18. V. Lekic, S. Cottaar, A. Dziewonski, B. Romanowicz, Cluster analysis of global lower mantle tomography: A new class of structure and implications for chemical heterogeneity. *Earth Planet. Sci. Lett.* **357–358**, 68–77 (2012). doi:10.1016/j.epsl.2012.09.014

19. Y. He, L. Wen, T. Zheng, Seismic evidence for an 850 km thick low-velocity structure in the Earth's lowermost mantle beneath Kamchatka. *Geophys. Res. Lett.* **41**, 7073–7079 (2014). doi:10.1002/2014GL061249

20. M. S. Thorne, E. J. Garnero, G. Jahnke, H. Igel, A. K. McNamara, Mega ultra low velocity zone and mantle flow. *Earth Planet. Sci. Lett.* **364**, 59–67 (2013). doi:10.1016/j.epsl.2012.12.034

21. R. Maguire, J. Ritsema, M. Bonnin, P. E. van Keken, S. Goes, Evaluating the resolution of deep mantle plumes in teleseismic traveltime tomography. *J. Geophys. Res. Solid Earth* **123**, 384–400 (2018). doi:10.1002/2017JB014730

22. S. W. French, B. Romanowicz, Broad plumes rooted at the base of the Earth's mantle beneath major hotspots. *Nature* **525**, 95–99 (2015). doi:10.1038/nature14876 Medline

23. C. Zhao, E. J. Garnero, M. Li, A. McNamara, S. Yu, Intermittent and lateral varying ULVZ structure at the northeastern margin of the Pacific LLSVP. *J. Geophys. Res. Solid Earth* **122**, 1198–1220 (2017). doi:10.1002/2016JB013449

24. Q. Williams, E. J. Garnero, Seismic evidence for partial melt at the base of Earth's mantle. *Science* **273**, 1528–1530 (1996). doi:10.1126/science.273.5281.1528

25. W. L. Mao, H. K. Mao, W. Sturhahn, J. Zhao, V. B. Prakapenka, Y. Meng, J. Shu, Y. Fei, R. J. Hemley, Iron-rich post-perovskite and the origin of ultralow-velocity zones. *Science* **312**, 564–565 (2006). doi:10.1126/science.1123442 Medline

26. T. Lay, Deep Earth Structure – Lower Mantle and D". *Treatise on Geophysics* **1**, 619–654 (2015). doi:10.1016/B978-044452748-6/00022-5

27. S. Ni, E. Tan, M. Gurnis, D. Helmberger, Sharp sides to the African superplume. *Science* **296**, 1850–1852 (2002). doi:10.1126/science.1070698 Medline

28. T. Lay, J. Hernlund, E. J. Garnero, M. S. Thorne, A post-perovskite lens and *D″* heat flux beneath the central Pacific. *Science* **314**, 1272–1276 (2006). doi:10.1126/science.1133280 Medline

29. C. Zhao, E. J. Garnero, A. K. McNamara, N. Schmerr, R. W. Carlson, Seismic evidence for a chemically distinct thermochemical reservoir in Earth's deep mantle beneath Hawaii. *Earth Planet. Sci. Lett.* **426**, 143–153 (2015). doi:10.1016/j.epsl.2015.06.012

30. S. W. French, B. A. Romanowicz, Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophys. J. Int.* **199**, 1303–1327 (2014). doi:10.1093/gji/ggu334

31. A. K. McNamara, A review of large low shear velocity provinces and ultra low velocity zones. *Tectonophysics* **760**, 199–220 (2019). doi:10.1016/j.tecto.2018.04.015

32. C. D. Williams, S. Mukhopadhyay, M. L. Rudolph, B. Romanowicz, Primitive helium is sourced from seismically slow regions in the lowermost mantle. *Geochem. Geophys. Geosyst.* **20**, 4130–4145 (2019). doi:10.1029/2019GC008437

33. J. K. Wicks, J. M. Jackson, W. Sturhahn, Very low sound velocities in iron-rich (Mg,Fe) O: Implications for the core-mantle boundary region. *Geophys. Res. Lett.* **37**, L15304 (2010). doi:10.1029/2010GL043689

34. A. Mundl, M. Touboul, M. G. Jackson, J. M. Day, M. D. Kurz, V. Lekic, R. T. Helz, R. J. Walker, Tungsten-182 heterogeneity in modern ocean island basalts. *Science* **356**, 66–69 (2017). doi:10.1126/science.aal4179 Medline

35. B. Steinberger, Plumes in a convecting mantle: Models and observations for individual hotspots. *J. Geophys. Res. Solid Earth* **105**, 11127–11152 (2000). doi:10.1029/1999JB900398

36. M. van Driel, L. Krischer, S. C. Stähler, K. Hosseini, T. Nissen-Meyer, Instaseis: Instant global seismograms based on a broadband waveform database. *Solid Earth* **6**, 701–717 (2015). doi:10.5194/se-6-701-2015

37. J. Ritsema, E. Garnero, T. Lay, A strongly negative shear velocity gradient and lateral variability in the lowermost mantle beneath the Pacific. *J. Geophys. Res. Solid Earth* **102**, 20395–20411 (1997). doi:10.1029/97JB01507

38. L. V. D. Maaten, G. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

39. Y. Rubner, C. Tomasi, L. J. Guibas, A metric for distributions with applications to image databases. *Proc. IEEE Int. Conf. Comput. Vis.* **6**, 59–66 (1998). doi:10.1109/ICCV.1998.710701

40. J. M. Kolb, V. Lekić, Receiver function deconvolution using transdimensional hierarchical Bayesian inference. *Geophys. J. Int.* **197**, 1719–1735 (2014). doi:10.1093/gji/ggu079

41. J. Tromp, D. Komatitsch, Q. Liu, Spectral-element and adjoint methods in seismology. *Commun. Comput. Phys.* **3**, 1–32 (2008).

42. A. M. Dziewonski, T.-A. Chou, J. H. Woodhouse, Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res.* **86**, 2825–2852 (1981). doi:10.1029/JB086iB04p02825

43. J. W. Hernlund, C. Houser, On the statistical distribution of seismic velocities in Earth's deep mantle. *Earth Planet. Sci. Lett.* **265**, 423–437 (2008). doi:10.1016/j.epsl.2007.10.042

44. F. A. Dahlen, S. H. Hung, G. Nolet, Fréchet kernels for finite-frequency traveltimes – I. Theory. *Geophys. J. Int.* **141**, 157–174 (2000). doi:10.1046/j.1365-246X.2000.00070.x