nature geoscience

Article

https://doi.org/10.1038/s41561-022-01116-9

Asthenospheric low-velocity zone consistent with globally prevalent partial melting

In the format provided by the authors and unedited

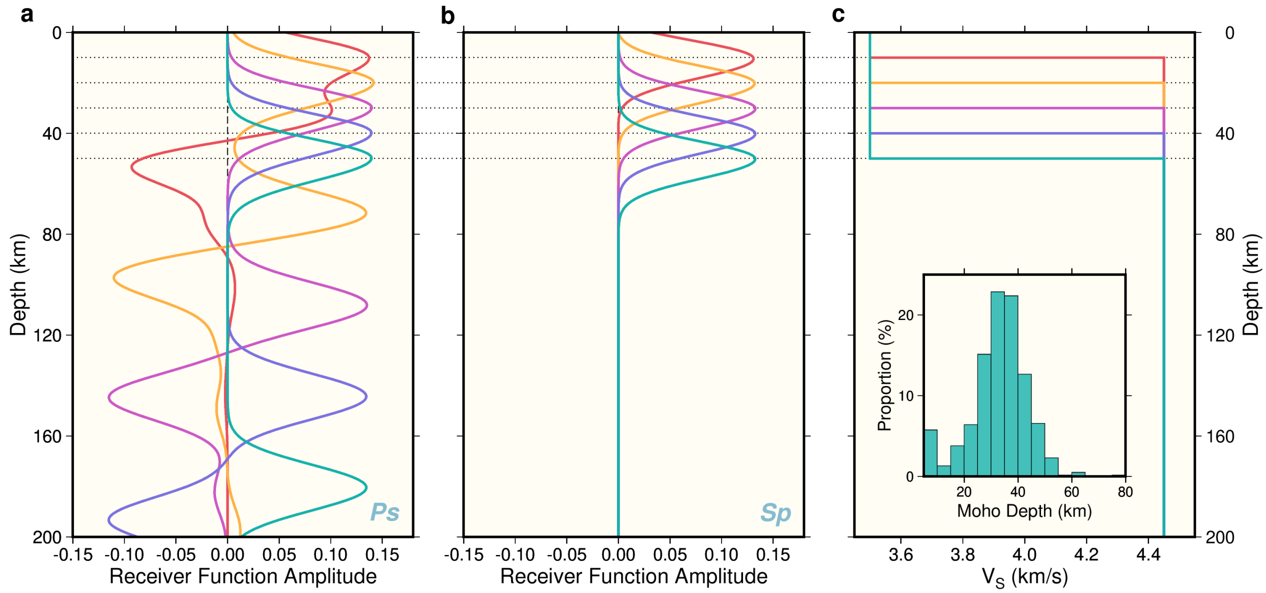
Table of contents:

Supplementary Method 1: Pre-processing and quality-control of the Sp receiver functions

To obtain an accurate velocity model for migrating Sp receiver functions, we first constrained crustal structure for each station by H-κ stacking. We followed the algorithm in ref.[45], but instead of determining crustal V_P and V_S simultaneously, we fixed crustal V_P at values from CRUST1.0[65], and more conservatively only solved for crustal thickness and V_P/V_S based on H-κ stacks of both Ps and Sp receiver functions. The receiver functions in this step were calculated with time-domain deconvolution[44], using a 2-100 s bandpass filter before deconvolution, and a post-deconvolution Gaussian shaping filter whose half-width (standard deviation) is 0.2s. During H-κ stacking, the six weighting factors w_1-w_6 in ref.[45] for Ps, PpPs, PpSs+PsPs, Sp, SsPp and SsSp phases were set to 0.275, 0.125, 0.275, 0.175, 0.05 and 0.1. In contrast to ref.[45], we estimated the standard deviation for the H-κ stacks based on the expressions in ref.[19]. We first picked the optimal crustal thickness and V_P/V_S at the maximum stacked value (S_{max}), and its corresponding standard deviation (S_{std}) was also estimated. Then we determined the area in the H-κ space around the optimal solution that has a stacked value higher than S_{max}- S_{std}. to represent the uncertainty. If the V_P/V_S and crustal thickness ranges for that area was less than 0.3 and 8 km, or the number of receiver functions was less than 30, we used the V_P/V_S and crustal thickness values estimated from the stack. Otherwise parameters from CRUST1.0 were assumed. The mantle portion of the migration velocity model was taken from GLAD-M25[22].

We then calculated and stacked Sp receiver functions. In this step, a 10-100 s filter (or 6-100 s filter in a few cases where explicitly stated) was applied to the waveforms before deconvolution, and a Gaussian shaping filter with a half-width of 1s was applied post-convolution, together with a 10 s (or 6 s) low pass filter. The receiver functions were migrated to depth, and careful quality control criteria were applied. S phases prior to deconvolution were required to have a signal to noise ratio higher than 2. We also required the maximum receiver function amplitude to be between 0.02 and 0.25 within 10 km of Moho depth from H-κ stacking, to eliminate receiver functions lacking a potential Moho phase and receiver functions with a very large Moho amplitude. In addition, we eliminated receiver functions with physically implausible amplitudes at depths between 200 and 350 km (root mean square amplitude higher than 0.1, and more than 30% of depth range with amplitudes higher than 0.1). The remaining 184,340 receiver functions were stacked by station (e.g. Fig. 1b-c), and the covariance matrices for the stacked receiver functions were estimated by bootstrapping the stacking process 20,000 times. The Sp receiver function amplitudes were flipped to match the Ps receiver function convention.

In principle, Ps receiver functions also contain information about mantle velocity gradients. However, for the stations used in this study, Moho depths are concentrated at 20-50 km, which generates strong crustal reverberation phases that would be migrated to the depth range of the PVG-150 (Supplementary Fig. 1). Hence, we focus on Sp receiver functions in this study.



Supplementary Fig. 1 | Crustal effects on Ps and Sp receiver functions. Synthetic Ps (**a**) and Sp (**b**) receiver functions for the velocity models in **c** were calculated[47] for epicentral distances of 50° (**a**) and 60° (**b**); post-deconvolution filters were only not applied here to better preserve the shape of crustal multiples. Colors correspond to structures with different Moho depths from 10 to 50 km (as shown in **c**). Horizontal dotted lines show the center of V_S discontinuities. The inset in **c** shows the Moho depth distribution for stations used in this study based on CRUST1.0[65]. Vertical dashed lines mark zero amplitude for receiver functions. This analysis shows that significant crustal reverberations in Ps receiver functions would be migrated to the depth range of PVG-150 (100-200 km) when the Moho is deeper than 20 km depth (**a**), which is not the case for Sp receiver functions (**b**).

Supplementary Method 2: Parameterization of mantle adiabats for thermal models

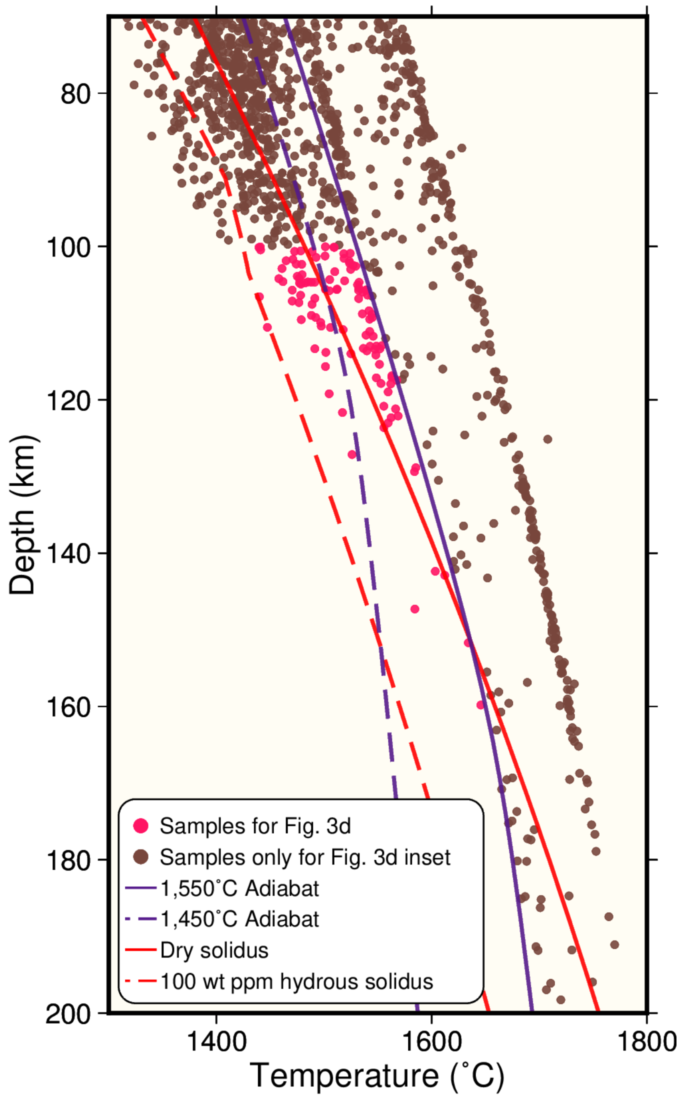
When calculating adiabatic temperature gradients, we used thermal expansion coefficients, heat capacities and densities for the solid mantle that were calculated by Perple_X[51] with the thermodynamic data and solution model of ref.[70] assuming a Harzburgite composition[14]. Perple_X-defined parameters were used here to be consistent with elastic seismic velocity estimates also based on Perple_X. The entropy difference between solid and melt were from ref.[54]. The solidus was based on ref.[55], and the degree of melting was estimated based on ref.[37] for consistency. The ref.[55] hydrous solidus was used as it has a relatively complete consideration of compositional effects. Although the degree of melting estimates from ref.[37] are primarily for small values, the predicted trend should be applicable for larger degrees of melting, and in any case, an accurate estimation of the degree of melting is not the target of this paper while the maximum allowed melt fraction in the mantle is less than 1%. We assumed batch melting to calculate the water partition coefficient when melt is present, so that we avoided complexity associated with fractional melting. With all of these choices and assumptions, the adiabatic temperature profiles were estimated based on ref.[71]. In the lithosphere, we assumed a conductive geotherm and, for simplicity, no heat generation. Mantle and crust heat conductivities[72] were 4 $Wm^{-1}K^{-1}$ and 2.5 $Wm^{-1}K^{-1}$, respectively, and the surface temperature was fixed at 10°C. When the adiabatic temperature at the LAB is less than the dry solidus temperature[55], the LAB conductive geotherm temperature is set to be equal to the LAB adiabat temperature. When the LAB adiabatic temperature exceeds the dry solidus temperature, the LAB temperature for the conductive geotherm is set to be the dry solidus temperature, and within the asthenosphere, the geotherm is set to the lower temperature between the adiabat and the extension of the conductive geotherm to asthenospheric depths[73].

Supplementary Method 3: Evidence from basaltic magma samples

Basaltic magma samples that erupted since the Pliocene from the GEOROC database (http://georoc.mpch-mainz.gwdg.de/Georoc) were analyzed globally. Most of the basalt samples are from intraplate volcanoes (Fig. 1a). The FeO and Fe_2O_3 weight percentages from the samples were first converted to FeO* to represent the total iron weight percentage. Then we calculated the sum of the weight percentages for major elements, and only samples with summed weight percentage between 98% and 102% were used. The major element weight percentages were normalized to make the sum equal to 100. We only used samples in an olivine control trend by observation of major element (FeO, TiO_2, CaO, Al_2O_3, CaO/Al_2O_3, Na_2O, etc.) variations with MgO following the recommendation discussed in ref.[36] and ref.[74], limiting samples to MgO higher than 9 wt %, and SiO_2 higher than 45 wt %.

We calculated the primary magma equilibration conditions based on ref.[36] (Supplementary Fig. 2). We assumed water in the primary melt to be 1 wt %, a typical value for intraplate basalts[73,75]. Changes in this parameter of 1 wt % could introduce a ~30°C difference in the determined mantle temperature. The source mantle forsterite content was set to 0.9, and the oxidation state parameter Fe^{3+}/Fe^{Total} was fixed at 0.14[76]. The equilibration pressures were converted to depth by assuming a 30 km thick crust with a density of 2,800 kg/m^3, and a mantle with a density of 3,300 kg/m^3. We only considered the 1100 samples that equilibrated between 70 and 250 km depths to represent asthenospheric melting. The values of V_{SV} at 110-130 km depth[22] at the geographic locations of these samples are shown in the inset to Fig. 3d. For estimates of primary magma equilibration conditions, we focused on samples which equilibrated beneath 100 km depth, to avoid samples affected by lithospheric cooling. When analyzing the depth distribution of the samples equilibrated in the asthenosphere, we only used 113 samples whose corresponding mantle potential temperatures are less than 1,550°C to eliminate very deep samples from hot regions (Supplementary Fig. 2). Mantle potential temperatures were determined by finding the best-fit asthenospheric adiabat for a hydrous mantle with 100 wt ppm of water with the solidus specified in ref.[55]. The degree of melting was estimated based on ref.[37]. When CO_2 was considered, its freezing point depression effect was based on ref.[38], but the corresponding change in temperature with the change in the degree of melting was taken from ref.[37]. Thermodynamic parameters to calculate asthenospheric adiabats in this section for both solid and melt are based on ref.[54].

As shown in Supplementary Fig. 2, most of the basalt samples were likely generated at higher-than-normal mantle temperatures. On a global scale, these samples are valuable to probe the high temperature mantle regions where observations of the PVG-150 phase often occur.



Supplementary Fig. 2 | Primary magma record of equilibration temperature and depth conditions for basalt samples. Dots show the equilibration temperature and depth values for the samples. Red dots are samples used for the degree of melting analysis in Fig. 3d, and brown dots are other post-Pliocene samples that are only used for the V_{SV} distribution in the inset of Fig. 3d, and not for the degree of melting analysis. Locations of these samples are shown in Fig. 1a with the same colors. The dry solidus and 100 wt ppm hydrous solidus are shown by red solid and dashed lines. The mantle adiabat for a mantle potential temperature of 1,550°C is shown by the purple solid line, and it is the upper limit for samples to be employed in the degree of melting analysis. The adiabat for a mantle potential temperature of 1,450°C is shown by the purple dashed line, and the corresponding degree of melting is shown in Fig. 3d.

Supplementary Method 4: Solving for the 1D horizontal mantle flow velocity

In the flow models, both stress-dependent dislocation creep and diffusion creep were considered. For the lithosphere, dry creep laws were assumed, which resulted in very high viscosities, and for the asthenosphere, wet dislocation and diffusion creep flow laws were employed[34]. For both Couette and Poiseuille flow, their governing equations are expressed as

$$\frac{\mathrm{d}\tau(z)}{\mathrm{d}z} = \frac{\mathrm{d}P}{\mathrm{d}x}(z),$$

where τ is the shear stress and dP/dx is the horizontal pressure imposed on the mantle. The dislocation and diffusion creep flow laws require

$$\frac{\mathrm{d}v}{\mathrm{d}z} = 3^{\frac{n_{dis}+1}{2}} A_{dis}(z)\tau|\tau|^{n_{dis}-1} + 3^{\frac{n_{dif}+1}{2}} A_{dif}(z)\tau|\tau|^{n_{dif}-1},$$

where v is the horizontal mantle flow velocity, and n_{dis} and n_{dif} are the stress dependence of dislocation creep and diffusion creep from ref.[34] which are 3.5 and 1, respectively. A_{dis} and A_{dif} are prefactors from ref.[34], and both have the form $AC_{OH}^{r}d^{-p}\exp(-\frac{E+PV}{RT})$, where C_{OH} represents water content (~380 H/10^{6}Si in olivine for 100 wt ppm bulk water content), d is grain size (10 mm), and pressure (P) and temperature (T) are from the assumed geotherm. Factors A, r, p, the activation energy (E) and the activation volume (V) are from ref.[34], but A for wet creep laws were reduced to 1/3 of their values in ref.[34] to account for calibrations of FTIR data in olivine[77]. The terms before the prefactors in the flow law are scaling factors to convert the laboratory defined flow law based on uniaxial compression to simple shear[78]. When variations in melt fraction are considered, A_{dis} and A_{dif} are multiplied with an $\exp(\alpha\phi)$ term[34], where ϕ is the melt fraction, and α is set to 30. When a sharp viscosity change is considered at 150 km depth due to partial melting, in addition to the $\exp(\alpha\phi)$ term above 150 km depth, depths below 150 km depth had their A_{dis} and A_{dif} values divided by 5. A factor of 5 is the suggested viscosity effect from a very small amount of melt for diffusion creep[10], although the magnitude of the decrease has also been inferred to be much smaller[79]. Moreover, results for diffusion creep are not directly applicable to the dislocation creep models considered here. However, since we are simply choosing a factor that significantly reduces viscosity for a trial model, a factor of 5 is reasonable. We chose to increase the viscosity beneath 150 km depth, instead of decreasing the viscosity above this depth, because if the small melt fraction influence is real, the samples used to define the empirical diffusion creep flow laws with the $\exp(\alpha\phi)$ term[34] are likely already slightly partially molten during experiments. Therefore, at melt-free depths, the viscosities would be increased 5 times relative to the experimental relationships. For Couette and Poiseuille flows, the combined effects of these terms are reflected by their equivalent viscosity ($\tau \bullet (\mathrm{d}v/\mathrm{d}z)^{-1}$) in Extended Data Fig. 9.

Solutions for horizontal mantle flow with the assumed 1D properties were obtained through integration, and no numerical solvers were necessary. Based on the force balance, the shear stress can be expressed as $\tau(z) = \int_0^z \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)\,\mathrm{d}z_1 + C_1$, where C_1 is a constant. After substituting this τ into the flow law and integrating the right side of that equation with depth, the flow velocity can be expressed as

$$\begin{aligned} v(z) = {} & 3^{\frac{n_{dis}+1}{2}} \int_0^z A_{dis}(z_2)\left(\int_0^{z_2} \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)\,\mathrm{d}z_1 + C_1\right)\left|\int_0^{z_2} \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)\,\mathrm{d}z_1 + C_1\right|^{n_{dis}-1} \mathrm{d}z_2 + \\ & 3^{\frac{n_{dif}+1}{2}} \int_0^z A_{dif}(z_2)\left(\int_0^{z_2} \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)\,\mathrm{d}z_1 + C_1\right)\left|\int_0^{z_2} \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)\,\mathrm{d}z_1 + C_1\right|^{n_{dif}-1} \mathrm{d}z_2 + C_2 \end{aligned},$$

where C_2 is another constant. To meet the boundary condition at the surface, C_2 is equal to the surface motion velocity (3 cm/yr for Couette flow and 0 for Poiseuille flow). C_1 was obtained based on the boundary condition that the velocity at 400 km depth is 0 for both flow types. For Couette flow, since dP/dx is zero, to meet the boundary condition C_1 would be higher than $-\left(3^{\frac{2}{n_{dis}+1}} C_2 \Big/ \int_0^{400} A_{dis}(z)\,\mathrm{d}z\right)^{\frac{1}{n_{dis}}}$ and less than zero. For Poiseuille flow, since C_2 is assumed to be zero, to make the velocity at 400 equal zero,

C_1 has to be in the range of $\pm\max\left|\int_0^z \frac{\mathrm{d}P}{\mathrm{d}x}(z_1)dz_1\right|$ for all z in the range from 0 to 400 km. With the range of potential C_1 thus defined, a grid search within this range was conducted to obtain the C_1 that satisfies the force balance, the flow law and boundary conditions. Then the flow velocity *v(z)* at different depths are determined. During the modeling, for Couette flow, the surface velocity was set to 3 cm/yr, the velocity at 400 km depth was set to zero, and d*P*/d*x* is zero at all depths. For Poiseuille flow, the velocities for both the surface and 400 km depth were set to zero, but a d*P*/d*x* of -3 Pa/m was imposed on the asthenosphere between 88 and 150 km depths.

Grain size will affect the partitioning between dislocation and diffusion creep, and with it absolute viscosity profiles and where crystallographic preferred orientation is being generated if that happens by dislocation creep only[49,57,80]. Our reference models with 10 mm grain size are consistent with evolution models[49,50]. However, there is some uncertainty about grain size evolution[81], and different grain size choices lead to different reference structures and anisotropy formation in our 1D models (e.g. the cases with 1 mm grain size in Extended Data Fig. 9). Importantly, regardless of the background deformation and anisotropy formation, whenever seismic anisotropy is developed, the cases with the sharp viscosity reduction show significantly larger anisotropy and modified depth dependence for the reduced grain size. Thus, while radial anisotropy can be influenced by grain size, the differences between cases with and without the viscosity reduction are significant, and our conclusions regarding whether the PVG-150 is a viscosity boundary are not affected.

Supplementary Discussion 1: Test for artificial side-lobe effects and the sharpness of the PVG-150

To verify the PVG-150 phase is not an artificial side-lobe of the stronger NVG from the top of the LVZ (the LAB), we stacked receiver functions obtained after applying a 6-100 s bandpass filter to the waveform data, in contrast to the 10-100 s bandpass used elsewhere in this study. The stacks that retain the shorter periods yield a similar variation of the PVG-150 with mantle velocity, while the PVG-150 is more clearly separated in depth from the LAB, indicating that it is not an artifact (Extended Data Fig. 3f). We also analyzed the relative amplitudes between PVG-150 and LAB phases at different V_{SV}, as well as between LAB and Moho phases (Extended Data Fig. 4a). These ratios vary significantly with V_{SV}, indicating that the phases are not simply side-lobes; if they were, the amplitude of the PVG-150 would correlate with LAB phase amplitude, and LAB phase amplitude would be strongly controlled by Moho phase amplitude.

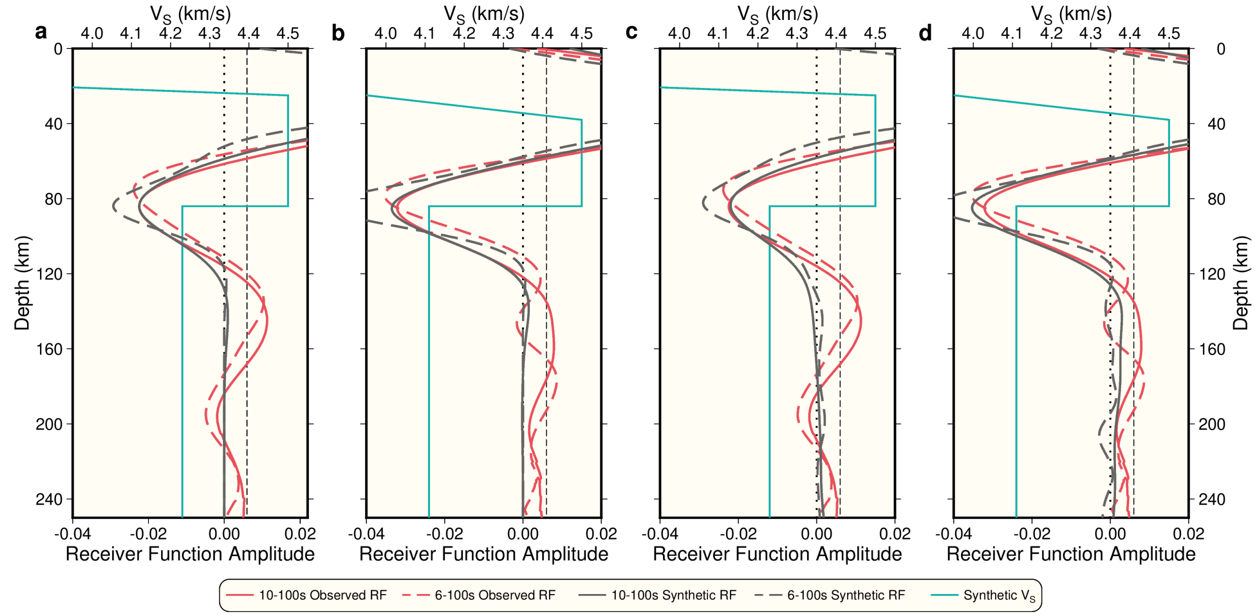
To show that the observed phases are not side-lobes or other types of artifacts, we compared synthetic receiver functions with two observed stacked receiver functions, one with LAB and PVG-150 phase amplitudes close to the median of all stations (a typical case) and the other with a very strong LAB phase and one of the weakest PVG-150 phase (Extended Data Fig. 4) (an extreme case that is most vulnerable to side-lobe effects) (Supplementary Fig. 3). For candidate velocity models, synthetic seismograms were generated for the ray parameters represented in the observed seismograms and processed in the same manner as the real data. Predicted receiver functions were migrated and stacked to compare with the observed receiver function stacks. In the velocity models, we adjusted Moho and LAB velocity gradients to approximately fit synthetic receiver functions to the observed 10-100 s stacks for these stations. We did not include a PVG-150 velocity gradient, so any apparent receiver function amplitudes at depths below the LAB would be artifacts. Based on the synthetic tests, the data processing we used would not generate strong side-lobes that are comparable to the observed PVG-150 phases (Supplementary Fig. 3). With noise-free synthetic seismograms (Supplementary Figs. 3a and 3b), the predicted receiver function stacks contain only very small side-lobes at depths directly below the LAB phases. To test the effects of observed waveform noise, we added real noise sampled from before P-wave arrivals from different events to the synthetics, and adjusted synthetic seismogram amplitudes to make signal to noise ratios equal for synthetic and observed waveforms. In the synthetic receiver function stacks with noise added (Supplementary Figs. 3c and 3d), some additional artificial signals below the LAB are present, but they are still much smaller than the observed PVG-150 phases, and they are too low amplitude to meet the threshold for picking the PVG-150 (vertical dashed lines in (Supplementary Fig. 3). These tests support the conclusion that the observed PVG-150 phases represent real structure and are not generated by data-processing or noise.

To demonstrate how waveforms filtered at 10-100 s and 6-100 s resolve velocity gradients over different depth ranges (Extended Data Fig. 5), synthetic seismograms were calculated with the propagator matrix method[47] for a range of models. The assumed velocity models contain a crust with a V_S of 3.9 km/s, a mantle lithosphere with a V_S of 4.43 km/s, a LVZ with a V_S of 4.3 km/s, and an underlying mantle with a V_S of 4.39 km/s. The velocity increase between the LVZ and the underlying mantle was accommodated by velocity gradients whose thicknesses were assumed to be 10, 20, 30, 40 and 50 km (Extended Data Fig. 5c). The V_S values were chosen to make the peak amplitudes of the resulting Sp receiver functions for the 10-100 s case match the median amplitudes of the observed PVG and NVG phases (Extended Data Fig. 4) to provide approximate estimates of the velocity changes across layers.

As expected, in the synthetic waveforms filtered at 10-100 s (Extended Data Fig. 5b), the phase produced by the PVG at 150 km shows less amplitude reduction as the velocity gradient becomes more gradual, relative to the case filtered at 6-100 s (Extended Data Fig. 5a). In addition, when the velocity gradient centered at 150 km depth is distributed over 50 km, the PVG-150 phase predicted for the 6-100 s filter is as wide in depth as the PVG-150 phase predicted for the 10-100 s filter (Extended Data Fig. 5a versus Extended Data Fig. 5b). However, the predicted PVG-150 for 6-100 s becomes systematically narrower in depth as the velocity gradient becomes more localized in depth, but the predicted PVG-150 for 10-100 s does not. Because the widths of the observed PVG-150 phases are narrower for the 6-100 s data than the 10-100 s data (Extended Data Fig. 3f versus Fig. 2a; Extended Data Fig. 3g versus Extended Data Fig. 3c; Extended Data Fig. 3h versus Extended Data Fig. 3d), the observed PVG-150 is likely to be typically distributed over less than 50 km. For the 10-100 s receiver function stack, a velocity increase of 4.0% provides a good fit to the median PVG-150 amplitude (shown as the dotted-dashed line at an amplitude of 0.01 in Extended Data Fig. 5).

To further demonstrate the effects of filtering, for a station whose LAB and PVG-150 phase amplitudes are close to the median of all stations (station ANM in the AK network) we calculated the receiver function stack for a range of filters (Extended Data Fig. 5d). All filters result in the same phases (Moho PVG, LAB NVG and PVG-150) and similar depths for the mantle phases. However, the width of the PVG-150 phase narrows when shorter periods are included, consistent with the trends in the global data and the conclusion that the PVG-150 velocity gradient is distributed over less than 50 km, as described in the preceding paragraph. In addition, with the 2-100 s filter, the PVG-150 is more clearly separated in depth from the LAB phase, providing additional evidence that the PVG-150 is not a LAB side-lobe.

While these examples (Extended Data Figs. 3f-h and 5d) show that the PVG-150 phase is robustly observed at a broad range of different period bands, we carried out much of the analysis and modeling of the PVG-150 using the 10-100 s filter. Because the phases in the 10-100s stacks are broader in depth, the amplitudes of the stacked receiver functions are less sensitive to lateral variations in the depth of the PVG-150, and, as shown in Extended Data Fig. 5a-b, data with the 10-100 s filter are less influenced by variations in the depth range over which the velocity gradient is distributed. Thus data filtered at 10-100 s will yield a more accurate distribution of amplitudes when stacked over large areas which may contain second-order variations in PVG-150 properties.



Supplementary Fig. 3 | Testing for artifacts in the receiver function stacks. Synthetic Sp receiver functions (filtered from 10-100 s) were fit to observed Sp receiver functions from station ANM in network AK (**a** & **c**) and station DLV in network RM (**b** & **d**) by adjusting positive velocity gradients at the Moho and negative velocity gradients at the LAB while not allowing any positive velocity gradients for the PVG-150. Red lines are observed receiver functions; black lines are synthetic receiver functions; and green lines are the shear velocities used to produce the synthetic seismograms. Solid lines are receiver functions filtered from 10-100 s, and dashed lines from 6-100 s. Vertical dotted lines mark zero amplitude for receiver functions. Vertical dashed lines show the threshold of 0.006 to pick the PVG-150 phase. For **a**, the observed LAB and PVG-150 phase amplitudes are close to the median of all stations (Extended Data Fig. 4), making it a typical case. In **b**, one of the strongest LAB phases and one of the weakest PVG-150 phases (Extended Data Fig. 4) occurs, making it most vulnerable to side-lobe effects. For **c**, noise from before P-wave arrivals at station ANM were added to synthetic seismograms in **a** before deconvolution to test the effects of noise. **d** is similar to **c**, but is for station DLV. While weak side-lobes and noise-related artifacts are present in the synthetics in the 100-200 km depth range, they are significantly lower than the observed PVG-150 phases and would not meet the criteria for PVG-150 picking.

Supplementary Discussion 2: Test for the effects of radial anisotropy on the PVG-150

We used synthetic Sp receiver functions to test whether layer with radial anisotropy (a transversely isotropic layer) could result in the observed PVG-150 (Extended Data Fig. 6). With RAYSUM[82] we generated synthetic seismograms for models containing a 30 km thick crust with a V_S of 3.9 km/s, a 58 km thick mantle lithosphere with a V_S of 4.5 km/s, a LVZ, and an underlying mantle with a V_S of 4.3 km/s. In isotropic layers, the mantle V_P/V_S was fixed at 1.8, and the mantle density was set to 3,300 kg/m^3. When the LVZ was assumed to be isotropic, its V_S was 4.23 km/s. As expected, this type of isotropic structure produces a clear PVG arrival at all ray parameters (Extended Data Fig. 6b), as seen in the PVG-150 observations (Extended Data Fig. 6a).

In models with S-wave radial anisotropy, we fixed the V_{Voigt} of the LVZ to be the same as the underlying mantle, but required ξ to be 1.1, so that the V_{SV} within the LVZ is the same as the LVZ V_S in the isotropic case (4.23 km/s). The fifth parameter needed to determine the stiffness tensor, η, was set to 0.9, which is approximately the η value in the PREM model[27] for a ξ value of 1.1. The model with S-wave radial anisotropy (Extended Data Fig. 6c) shows a decrease in the amplitude of the PVG-150 phase with increasing ray parameter that is not seen in the observed receiver functions (Extended Data Fig. 6a). Rather, the observed receiver functions are more consistent with an isotropic low velocity zone (Extended Data Fig. 6b). If η is increased to unity, the behavior of the PVG-150 amplitude with ray parameter becomes more similar to the isotropic case and to the data. However, η is more likely to be ~0.8[57] for a ξ of 1.1 for olivine, which would produce an even stronger reduction of PVG-150 amplitude with ray parameter than what is shown in Extended Data Fig. 6c, supporting our conclusion that uniform radial anisotropy alone is not sufficient for the observed PVG-150.

We also tested the potential influence of P-wave radial anisotropy, which is capable of generating receiver functions phases[26]. Here we fixed the Voigt average of horizontally polarized compressional velocity (V_{PH}) and vertically polarized compressional velocity (V_{PV}) ($\sqrt{\left(V_{PV}^2+4V_{PH}^2\right)/5}$) to be the same as the underlying mantle (7.74 km/s), while setting P-wave radial anisotropy φ^{-1} (V_{PH}^2 / V_{PV}^2) to be 0.96, and η at 0.9. The resulting synthetic receiver functions show that P-wave anisotropy could also generate a PVG phase (Extended Data Fig. 6d), but, similar to the case for S-wave radial anisotropy (Extended Data Fig. 6c), the predicted ray parameter dependence differs from the observations (Extended Data Fig. 6a). In addition, to produce a PVG phase, V_{PH} must be smaller than V_{PV}, which is generally not favored when mantle is not upwelling (Extended Data Fig. 8d), since φ^{-1} correlates positively with ξ[57].

Lastly, we tested the combined effects of S-wave and P-wave anisotropies. Based on PREM[27], when ξ is 1.1, φ^{-1} and η would be 1.04 and 0.9, and we set the Voigt average for both shear and compressional velocities to be the same as the underlying mantle. The result shows that the combined radial anisotropy does not generate a PVG (Extended Data Fig. 6e). Overall these modeling results indicate that the role of radial anisotropy in creating the PVG-150 phase is likely to be limited.

The set of parameters from PREM likely primarily represent anisotropy due to olivine crystallographic preferred orientations, and do not reflect additional anisotropy caused by the distribution and geometry of partial melt[83]. However, even if future modeling were to show that including the effects of partial melt on velocity anisotropy can produce a phase with the ray parameter dependence of the PVG-150 observations, this finding would not change the conclusion that partial melt exists in the LVZ since melt would be required to produce that anisotropy.

In addition, the representative radial anisotropy profiles for groups with or without the PVG-150 are nearly identical (Extended Data Fig. 8b), which would not be expected if radial anisotropy was the source of the observed PVG-150 phase.

Supplementary Discussion 3: Additional comparisons between predicted and observed PVG-150 phase patterns

Here we provide additional comparisons of the predicted receiver functions patterns with the observed receiver functions binned as a function of average V_{Voigt} at depths of 110-130 km (Fig. 3). The predicted pattern contains a significant Sp phase (Fig. 3b) that is generated at the base of the low velocity zone that contains partial melt (Fig. 3a). This predicted Sp phase appears at reference V_S values of less than 4.39 km/s, matching the Voigt velocity range where the Sp arrival corresponding to the PVG-150 appears in the observed receiver functions (Fig. 3c). The depths of the predicted and observed Sp phases are similar. The observed PVG-150 deepens slightly with increasing reference velocity, likely due to larger variations in lithosphere thickness and to an uneven sampling of mantle conditions (i.e. water content, mantle temperature, and melt fraction) in the real Earth. This effect is not seen as strongly in the predicted receiver functions, as expected, because the assumed lithospheric thickness is fully dependent on mantle potential temperature without the effects of other factors and the ranges of different mantle parameters are sampled evenly. Nonetheless, the fit between the predicted (Fig. 3b) and observed (Fig. 3c) PVG-150 phases is still good. Moreover, the apparent increase of observed PVG-150 depth with increasing reference velocity depends on the reference velocity model, and is not necessarily a robust feature of the data. For example, a slight shallowing of the PVG-150 with increasing V_{SV} is observed with SEMUCB_WM1[21] as the reference model (Extended Data Fig. 3a). Another consideration is that the velocity models used in the predictions assume a constant crustal thickness, whereas average crustal thickness is greater in the continental regions that dominate the bins at greater V_s, which explains why the Moho phases in the observed receiver function stacks are deeper at higher V_s values, relative to those in the predicted stacks. Hence, the discrepancies between the predicted and observed receiver function stacks are relatively minor compared to overall good fit of the PVG-150 phase depths at a given V_s.

We also tested whether poroelasticity is required to predict the PVG-150 pattern in detail. In the case where we directly removed the poroelasticity (Extended Data Fig. 7a), no significant Sp phase at depths near 150 km is observed, indicating that the anelastic effect from the variation of melt fraction alone is not sufficient to produce the observed PVG-150 (Extended Data Fig. 7a). Then, we also tested a different case that includes a sharp diffusion creep viscosity change where melting just starts to occur[10] (Extended Data Fig. 7b). This mechanism predicts an Sp phase at a different reference V_S range with a larger depth variation than the observed PVG-150, and thus is less likely than the poroelastic case[33].

In these predictions, grain size is fixed at 10 mm, but we also tested a case with a grain size of 1 mm. The resulting pattern of PVG-150 predictions (Extended Data Fig. 7c) is similar to the 10 mm case (Fig. 3b), although the PVG-150 phase emerges at slightly lower V_S. Because grain size and water content can affect seismic velocity by changing anelasticity (through their influence on diffusion creep viscosities), the same effect on V_S can be achieved by varying either of the parameters. We decided to keep grain size fixed at 10 mm in this study while allowing water content to vary.

We also tested the possible influence of a pre-melting effect on anelasticity[67] (Extended Data Fig. 7d). This effect generates a significant PVG-150 phase without poroelasticity, but it appears in a reference V_S range that differs from the observed PVG-150 distribution (Fig. 3c), and the amplitude of the PVG-150 phase is smaller at lower V_S as there the positive velocity gradient due to pre-melting is more gradual. By applying the model for the anelasticity relationship in ref.[67], we directly used its given diffusion creep relationship without specifying a grain size and water content, and water would only affect anelasticity through the pre-melting effect by modifying the hydrous solidus[55].

In summary, our analysis indicates that we cannot match the observed PVG-150 pattern successfully without considering the poroelasticity effect due to the presence of melt. To further demonstrate anelasticity alone is not the cause of the PVG-150, we also compared observed attenuation profiles for stations with or without the PVG-150 phase (Extended Data Fig. 8c). These patterns are approximately identical in the PVG-150 depth range for the two groups, suggesting no obvious anelasticity differences that are related to the existence of PVG-150.

Supplementary Discussion 4: Comparison of radial mantle flow to Sp receiver functions

To assess the role of mantle upwelling in creating the PVG-150, we compared the distribution of the PVG-150 to radial mantle flow velocity in two global mantle flow models, and found no clear relationship. The flow models incorporate plate motions and mantle density heterogeneity. One is based on the method in ref.[43] but using an updated tomographic model[68], and the other is from ref.[69]. Both models show that regions without the PVG-150 on average have stronger downwelling (Extended Data Fig. 8d). Based on ref.[43] the average mantle flow direction for regions with the PVG-150 is upward, although the rates are very small and uncertainties overlap rates of zero. Based on ref.[69], regions with the PVG-150 have small downward flow below depths of ~200 km. These comparisons are consistent with the observed correlation of the PVG-150 and low velocities in the upper mantle above 200-300 km (Extended Data Fig. 8a), since the buoyancy driven radial mantle flow velocity is dependent on the scaled density structure from tomographic V_S. They also indicate that while the absence of the PVG-150 tends to occur in regions with strong downwelling, strong upwelling is not required to produce the PVG-150.