nature geoscience

Analysis

https://doi.org/10.1038/s41561-024-01558-3

Large-scale mantle heterogeneity as a legacy of plate tectonic supercycles

Received: 15 October 2023

Accepted: 11 September 2024

Published online: 18 October 2024

Check for updates

Luc S. Doucet [1,2] & Zheng-Xiang Li [1,2]

The Earth's mantle is divided by the circum-Pacific subduction girdle into the African and Pacific domains, each featuring a large low-shear-wave-velocity province (LLSVP) in the lower mantle. However, how this hemispherical-scale mantle structure links to Earth's plate tectonic evolution remains unclear. Previous geochemical work has suggested the presence of a north–south hemispheric subdivision, with large-scale mantle heterogeneities in the Southern Hemisphere, termed the DUPAL (Dupré and Allegre) anomaly. Here we compile elemental and isotopic data of both shallow-mantle-derived oceanic igneous rocks from mid-ocean ridges and deeper-mantle plume-related samples (ocean islands and oceanic plateaus) and analyse these using supervised machine learning classification methods. Data from both shallow- and deeper-mantle-sourced samples illustrate a consistent chemical dichotomy. Our results indicate that heterogeneities in the present-day shallow and deep mantle are not exclusively controlled by the north–south hemispheric DUPAL anomaly. Instead, they are consistent with a chemical dichotomy between the African and Pacific mantle domains and their associated LLSVPs. These observations can best be explained by tectonic supercycles over the past one billion years involving two supercontinents and two superoceans.

The present-day Earth's mantle structure is dominated by a degree-two spherical harmonic, featuring two equatorial and antipodal mantle domains bisected by a time-integrated girdle of subducted oceanic slabs in the mantle but represented on the surface by the circum-Pacific subduction system[1,2]. Each of the two antipodal African and Pacific mantle domains features a large low-shear-wave-velocity province (LLSVP) in the lower mantle (Fig. 1), an expression of large physical heterogeneity. Whether there is a dynamic connection between the configuration of such mantle domains and global tectonic and geodynamic evolution is a long-standing debate. Clues to the answer may lie in the chemical heterogeneity of the mantle.

Dupré and Allegre[3] were the first to document large-scale mantle heterogeneity in their seminal 1983 paper. They used long-lived radiogenic isotope systems Pb and Sr from ocean island basalts (OIBs), oceanic large igneous provinces (O-LIPs) and mid-ocean ridge basalts (MORBs) to fingerprint the composition of their respective mantle sources at the times and locations of their formation. Isotopic differences were observed between Indian Ocean basalts and North Atlantic and Pacific basalts. The authors hypothesized that such large-scale heterogeneity was caused by reinjection of sediments (continental crust materials) into the sources of Indian OIBs and O-LIPs, which led to higher $^{87}Sr/^{86}Sr$, elevated $^{207}Pb/^{204}Pb$ and $^{208}Pb/^{204}Pb$ ratios for a given $^{206}Pb/^{204}Pb$ ratio and a more radiogenic helium signature. They also showed that Indian Ocean MORBs were more radiogenic, interpreted as evidence for mixing of an originally homogeneous upper mantle with blobs of sediment-enriched material from the lower mantle (the so-called pollution model). In 1984, Hart[4] popularized the idea of a large-scale isotope anomaly in the mantle by coining the now well-known term DUPAL (Dupré and Allegre) anomaly. On the basis of the observation that the identified enriched basalts were located in the Southern Hemisphere, a more comprehensive isotopic dataset of MORBs and OIBs from Hawaii, Iceland, the Azores, the Canaries

[1]Earth Dynamics Research Group (EDRG), School of Earth and Planetary Sciences, Curtin University, Perth, Western Australia, Australia. [2]Earth Evolution and Dynamics Research Centre (EDRC), Laoshan Laboratory, Qingdao, China. e-mail: luc-serge.doucet@curtin.edu.au; z.li@curtin.edu.au

https://doi.org/10.1038/s41561-024-01558-3

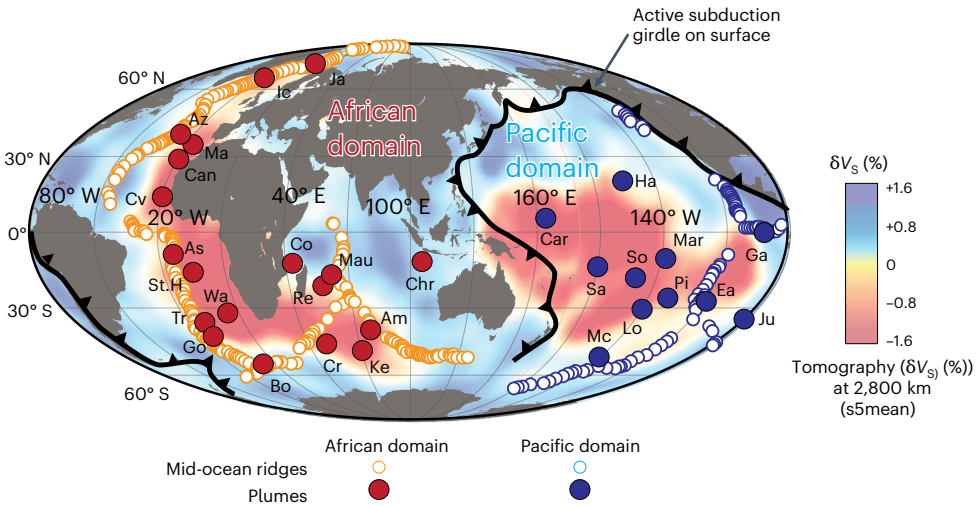


Fig. 1 | Global mantle structure highlighting the two LLSVPs. Seismic shear-wave-velocity anomalies (δV_s) are those of the mean S-wave tomography model s5mean[24]. The map shows the positions of the two mantle domains, their respective LLSVPs and the circum-Pacific subduction girdle (thick black line). The locations of previously studied MORBs (Extended Data Fig. 1 and Supplementary Table 1), OIBs and O-LIP[6] samples used in this study are also shown. The map was produced using the open-source programming language Python 3.0 and the open-source library pygmt. Am, Amsterdam; As, Ascension; Az, Azores; B, Bouvet; Can, Canary; Car, Caroline; Chr, Christmas; Co, Comores; Cr, Crozet; Cv, Cape Verde; Ea, Easter; Ga, Galapagos; Go, Gough; Ha, Hawaii; Ic, Iceland; Ja, Jan Mayen; Ju, Juan Fernandez; Ke, Kerguelen; Lo, Louisville; Ma, Madeira; Mar, Marquesas; Mau, Mauritius; Mc: Mcdonald; Pi, Pitcairn; Re, Reunion; Sa, Samoa; So, Societies; St.H, Saint Helenas; Tr, Tristan; Wa, Walvis.

and Cape Verde was further used to define an 'arbitrary but convenient' Northern Hemisphere reference line as a counterpart to the enriched DUPAL mantle province[4]. Hart[4] suggested that the DUPAL anomaly is a Southern Hemisphere feature centred around latitude 30° S, a definition still widely used today by much of the mantle geochemistry community. In addition, Hart[4] demonstrated that the high time-integrated Rb/Sr, U/Pb and Th/U of the DUPAL anomaly imply ancient (>500 million years ago (Ma)) re-enrichments, probably over billions of years.

Castillo[5] was one of the first to link the location of hotspots and isotopic anomalies with mantle structures identified by mantle tomography. Castillo[5] showed that the DUPAL anomaly has two maxima, one close to the southern tip of Africa and one in the Pacific, both above the "two large-scale regions of low velocity (LVRs) in the lower mantle" and that such LVRs (now known as LLSVPs) are surrounded by present-day and recent subduction zones, thus pointing to possible links between the DUPAL anomaly, LLSVPs, upwelling lower mantle and subduction. Over recent decades, numerous attempts have been made to further map out the DUPAL anomaly, but its possible links with global tectonic and geodynamic evolution remain unclear.

Isotopic fingerprinting of deep-mantle sources

Geochemical and isotopic data for oceanic rocks showed that the deep mantle of the African and Pacific mantle domains have distinct geochemical features, as demonstrated by long-lived radiogenic isotope systems (Pb, Sr, Nd) in both OIBs and O-LIPs[6–8] (Figs. 2 and 3). Compared with the deep mantle of the Pacific domain, the deep mantle of the African domain generally has lower $^{143}Nd/^{144}Nd$ and $^{176}Hf/^{177}Hf$, higher $^{87}Sr/^{86}Sr$ and elevated $^{207}Pb/^{204}Pb$ and $^{208}Pb/^{204}Pb$ ratios for a given $^{206}Pb/^{204}Pb$ ratio, all characterizing distinctive isotopic enrichment. These observations suggest that heterogeneities in the present-day deep mantle are not controlled by a north–south hemispheric subdivision, as previously ascribed for the DUPAL anomaly (Extended Data Fig. 1a–d). Instead, they are dominantly associated with the two antipodal mantle domains and their respective LLSVPs, both centred close to the Equator. Subduction associated with the formation of the supercontinent Pangaea, which created a slab graveyard through prolonged and multiple ocean closures, can best explain why the African mantle domain is more enriched with continental crustal materials[6,9–11], whereas the Pacific mantle domain has largely been protected from such 'contamination' for much of the past 600 million years[6,9–11]. The deep-mantle geochemical dichotomy is consistent with a dynamic nature of mantle structure, coupled with plate tectonics and supercontinent cycles[6,9,11–13]. If this model is correct, the geochemical and isotopic composition of the upper mantle should also have been affected by such global plate and mantle re-organizations.

Isotopic composition of MORB mantle sources

Here we compare the geochemical and radiogenic isotopes of basaltic lava flows from present-day mid-ocean ridges (MORBs) with data from OIBs and O-LIPs (Fig. 1) to examine the upper mantle response to the hemispherical chemical dichotomy identified from deep-mantle-derived mafic rocks. Basalts along the global mid-ocean ridge chains on present-day Earth (Fig. 1), with a total length of ~65,000 km, represent the most recent decompression melts from the upper mantle. Such MORBs indirectly record the composition of their underlying upper mantle sources[14]. At first order, MORBs are geochemically and isotopically more uniform relative to OIBs, for example, they show a relatively homogeneously depleted signature (known as the depleted MORB mantle)[15]. Nonetheless, heterogeneities have been identified in MORBs[14]. On the basis of their trace element compositions, MORBs can be classified as normal-MORB (N-MORB), depleted-MORB (D-MORB) and enriched-MORB (E-MORB)[16]. Global MORBs, regardless of their classification, also show heterogeneity in their radiogenic isotopes (for example, Pb, Sr)[3,17,18], lithophile elements (for example, Ba, K)[8] and volatiles (for example, C)[19].

We compiled geochemical and isotopic data from a total of 3,983 global MORB samples (Fig. 1), including whole-rock, glass and melt inclusions, with 1,661 (42%) from the African domain and 2,322 (58%) from the Pacific domain (Methods and Supplementary Table 1). To examine variability in the MORB dataset, we first classified MORBs following the approach in ref. 16, using whole-rock La and Sm contents (Methods). We classified the MORB samples into N-MORB, D-MORB, E-MORB and 'unknown MORB' when no trace element data were

https://doi.org/10.1038/s41561-024-01558-3

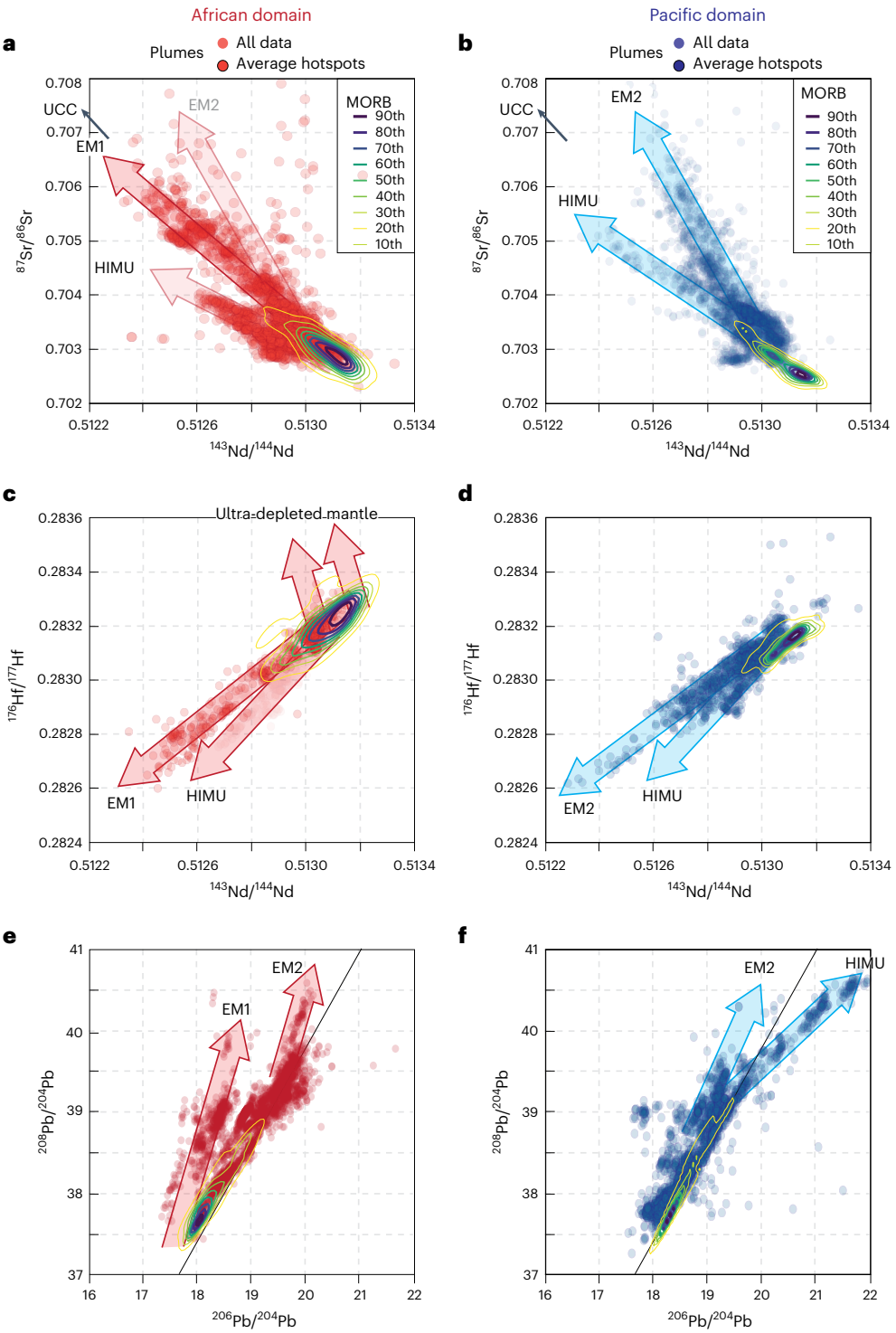


Fig. 2 | Geochemical and isotopic data of MORB, OIB and oceanic-LIP samples from the African and Pacific domains. a,**b**, $^{87}Sr/^{86}Sr$ versus $^{143}Nd/^{144}Nd$ for the African (**a**) and Pacific (**b**) domains. **c**,**d**, $^{177}Hf/^{176}Hf$ versus $^{143}Nd/^{144}Nd$ for the African (**c**) and Pacific (**d**) domains. **e**,**f**, $^{208}Pb/^{204}Pb$ versus $^{206}Pb/^{204}Pb$ for the African (**e**) and Pacific (**f**) domains. The contour lines represent the percentile relative density of the kernel density estimation for MORB data (corrected for geographical bias; Methods). The average values of individual OIB and O-LIP, EM1, EM2, HIMU and the UCC isotopic endmembers[15] are also shown. Geochemical and isotopic data used in this study were compiled from ref. 6 and PetDB through the EarthChem Portal.

available in the dataset (Methods and Extended Data Fig. 2). Overall, 54% of MORB samples in online databases do not have trace element compositions and thus cannot be classified. However, the unclassified MORBs have isotopic compositions that plot within the distribution fields of N-, D- and E-MORB samples (Extended Data Fig. 3), indicating that these MORBs do not require new MORB classifications.

While samples near known mantle plumes or plume-influenced rocks do not necessarily give E-MORB signatures[16], the proximity of a mantle plume could potentially cause 'contamination' of the nearby MORB mantle source. We therefore used the distance between a ridge segment and the nearest hotspot to examine potential plume influence on geochemical (Ba/La and La/Sm) and Sr, Nd, Pb and Hf isotopic

https://doi.org/10.1038/s41561-024-01558-3

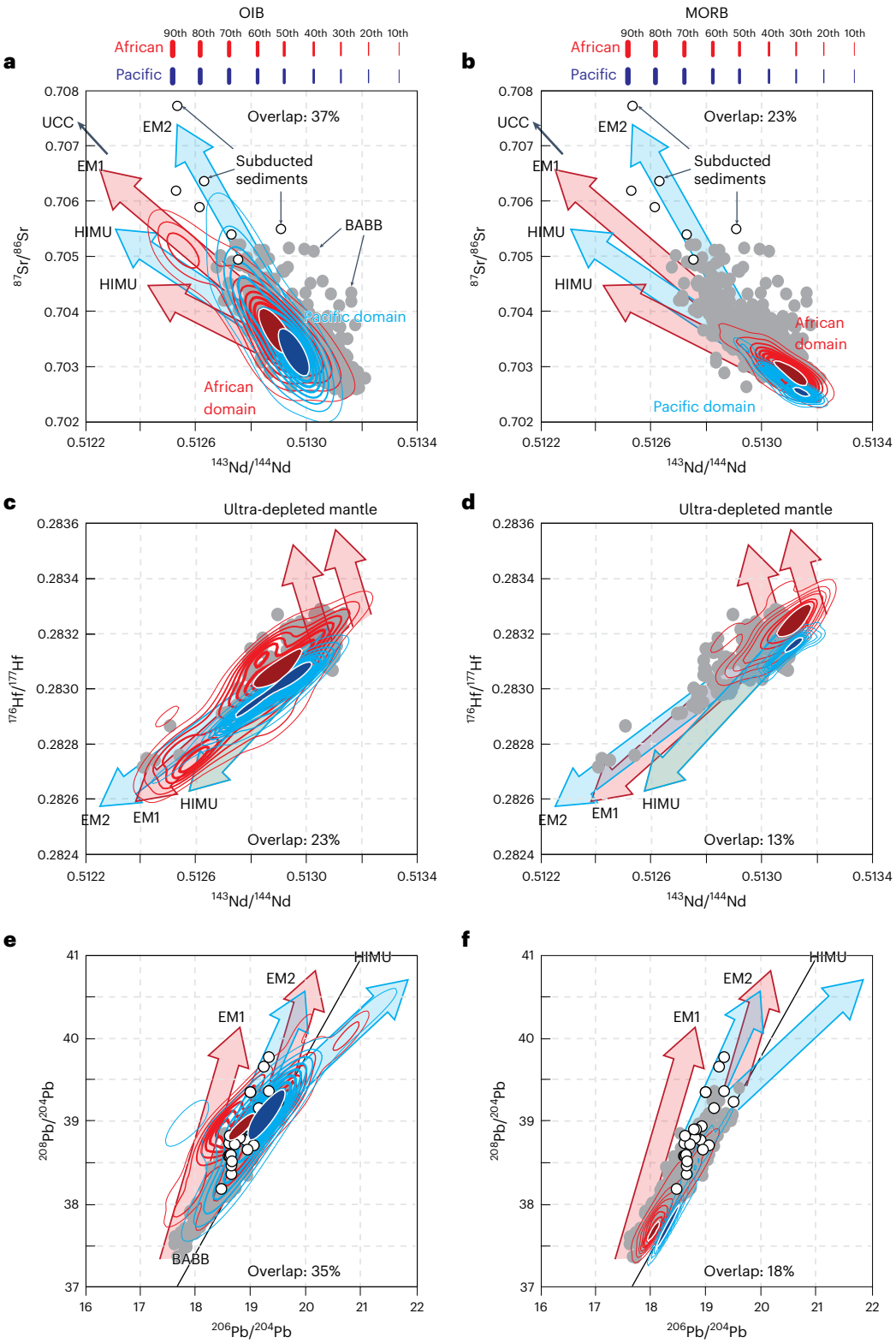


Fig. 3 | Geochemical dichotomy observed in MORB, OIB and oceanic-LIP samples from the African and Pacific domains. a,**b**, $^{87}Sr/^{86}Sr$ versus $^{143}Nd/^{144}Nd$ for the African and Pacific domain OIB (**a**) and MORB (**b**) samples. **c**,**d**, $^{177}Hf/^{176}Hf$ versus $^{143}Nd/^{144}Nd$ for the African and Pacific domain OIB (**c**) and MORB (**d**) samples. **e**,**f**, $^{208}Pb/^{204}Pb$ versus $^{206}Pb/^{204}Pb$ for the African and Pacific domain OIB (**e**) and MORB (**f**) samples. **a**,**c**,**e**, The contour lines represent percentiles of kernel density estimation of isotopic data for all OIB and oceanic-LIP (O-LIP) data for the African (red) and Pacific (blue) mantle domains corrected for geographical bias (Methods). **b**,**d**,**f**, The contour lines represent the percentile relative density of the kernel density estimation of isotopic data for MORBs from the African (red) and Pacific (blue) mantle domains. The average values of individual OIB and O-LIP, EM1, EM2, HIMU and the UCC isotopic endmembers[15], subducted sediments[25] and back-arc basalts (BABB, as grey dots)[8] are also shown. Geochemical and isotopic data used in this study were compiled from ref. 6 and PetDB through the EarthChem Portal, with MORB data given in Supplementary Table 1.

https://doi.org/10.1038/s41561-024-01558-3

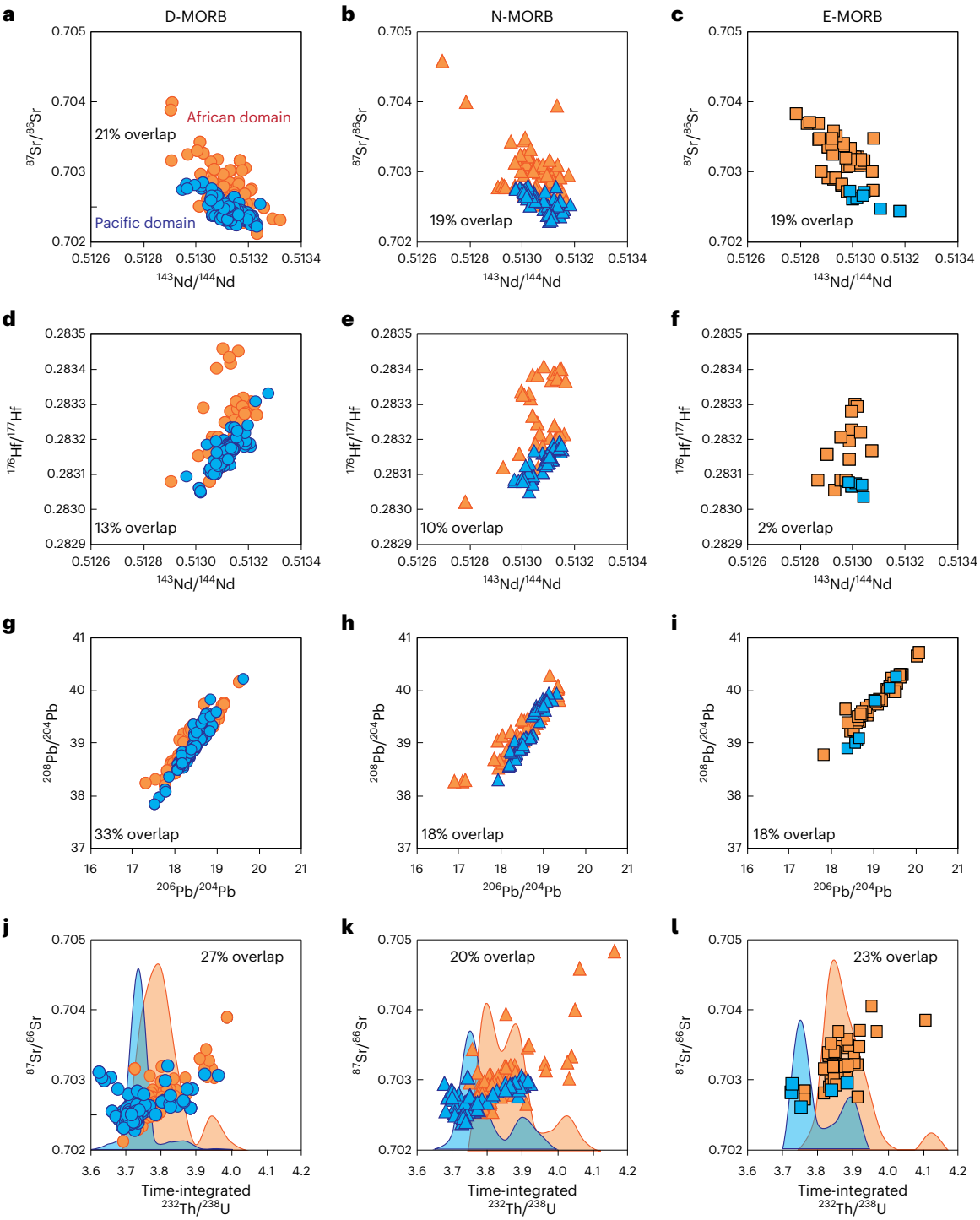


Fig. 4 | Geochemical dichotomy observed in D-, N- and E-MORB results from the African and Pacific domains. a–**c**, $^{143}Nd/^{144}Nd$ versus $^{87}Sr/^{86}Sr$ for D-MORB (**a**), N-MORB (**b**) and E-MORB (**c**). **d**–**f**, $^{143}Nd/^{144}Nd$ versus $^{177}Hf/^{176}Hf$ for D-MORB (**d**), N-MORB (**e**) and E-MORB (**f**). **g**–**i**, $^{206}Pb/^{204}Pb$ versus $^{208}Pb/^{204}Pb$ for D-MORB (**g**), N-MORB (**h**) and E-MORB (**i**). **j**–**l**, $^{87}Sr/^{86}Sr$ versus $^{232}Th/^{238}U$ for D-MORB (**j**), N-MORB (**k**) and E-MORB (**l**). Distribution of the $^{232}Th/^{238}U$ field is also shown. Geochemical and isotopic data used in this study were compiled from PetDB through the EarthChem Portal, and are listed in Supplementary Table 1.

compositions of MORBs (Extended Data Fig. 4). While geochemical and isotopic variabilities tend to decrease with greater distance from hotspots (~1,000–1,500 km)[19], there is no clear correlation between distance to a plume and the geochemical and isotopic compositions of MORBs (Extended Data Fig. 4). We also tested the geochemical and isotopic variabilities relative to the degree of partial melting, using spreading rate as a proxy[19], and found no clear correlation between the spreading rate and the geochemical and isotopic compositions of MORBs (Extended Data Fig. 4). In addition, the isotopic composition of E-MORBs overlaps with that of N- and D-MORBs (Extended Data Fig. 3), indicating that regardless of their classification, MORB types do not show a clear correlation[19]. By contrast, D-MORBs, N-MORBs and E-MORBs from the African and Pacific domains all show distinct domain-related geochemical and isotopic compositions, with African MORBs having higher $^{87}Sr/^{86}Sr$, $^{177}Hf/^{176}Hf$, time-integrated $^{232}U/^{238}U$ and lower $^{143}Nd/^{144}Nd$ than Pacific E-MORBs (Fig. 4).

https://doi.org/10.1038/s41561-024-01558-3

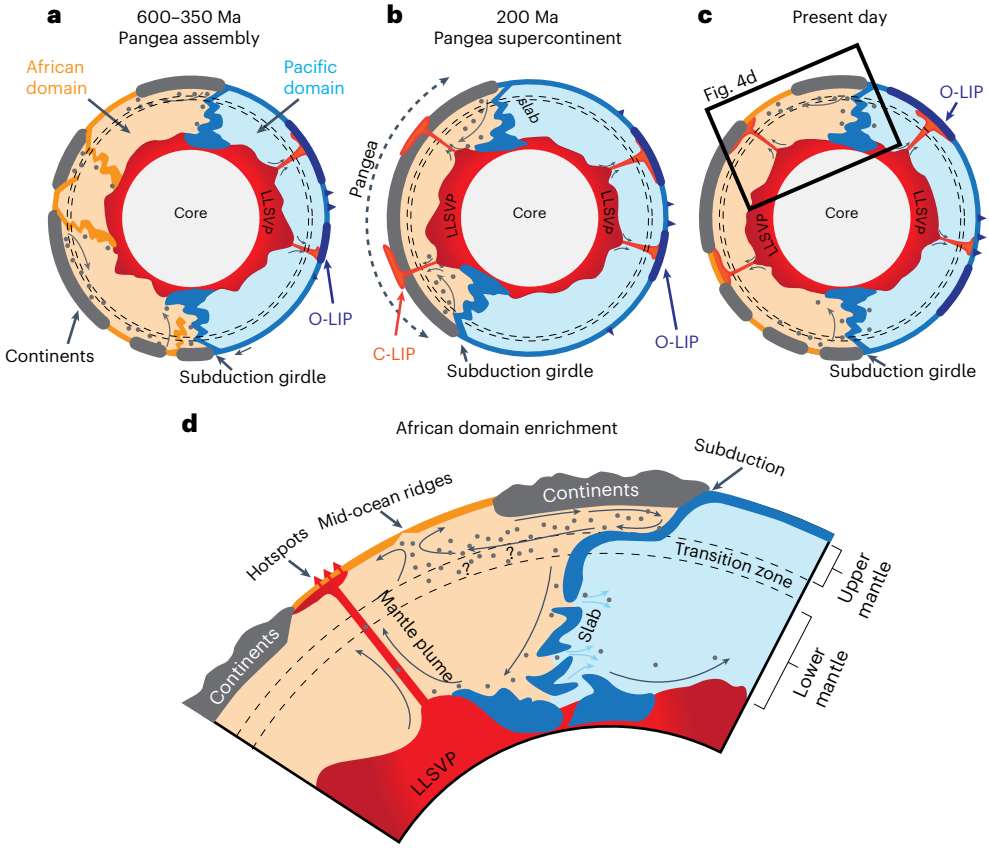


Fig. 5 | Schematic showing a dynamic and coupled supercontinent-plume-mantle domain model. a, Between 600 Ma and 350 Ma, Gondwana and Pangea started to assemble after the break-up of the supercontinent Rodinia by ~700 Ma. **b**, Around 200 Ma, formation of antipodal LLSVPs after the final assembly of the supercontinent Pangea at ~320 Ma. **c**,**d**, Present-day configuration of mantle domains with residual equatorial African and Pacific LLSVPs after Pangea break-up, but before the assembly of the future supercontinent Amasia. **d** shows an enlarged area of **c**. C-LIP, continental large igneous provinces. During the Rodinia-to-Pangea supercontinent cycle (**a**–**c**), the subducted slabs (dark blue from Pacific domain ocean(s) and orange from the African domain oceans) all dominantly rested in the African domain, bringing continental material primarily to the African domain.

To further test the MORB mantle dichotomy statistically, we applied supervised machine learning classification methods using convolutional neural network (CNN), random forest and quadratic discriminant analysis (QDA) (Methods and Extended Data Fig. 5). During the training phase, the models were trained to recognize distinguishing features that separate one group from the others. To test the accuracy of the methods, the models were then used to generate predicted datasets, which were compared with each data group. The models' ability to correctly identify the group affiliation of each dataset was demonstrated by an accuracy value of 0.99 for all models (that is, 99%; Extended Data Fig. 5). While the models showed no clear distinction between the N- and D-MORB data groups, they were in general able to distinguish data from the two mantle domains, indicating distinct mantle characteristics between the two domains.

Overall, both the geochemical and isotopic data show a notable compositional difference between MORBs from the African and Pacific mantle domains (Figs. 3b,d,f and 4, and Supplementary Figs. 1–4), similar to what is observed from the OIB and O-LIP datasets[6] (Figs. 2 and 3a,c,e, and Supplementary Fig. 5). MORBs from the Pacific domain's upper mantle have relatively homogeneous geochemical and isotopic compositions (Figs. 3 and 4). By contrast, MORBs from the African domain's upper mantle show a broader spread in their geochemical and isotopic compositions, similar to African domain OIBs and O-LIPs (Figs. 2–4). The African domain MORBs show enrichments relative to the Pacific domain, with higher $^{87}Sr/^{86}Sr$ and $^{177}Hf/^{176}Hf$ and time-integrated $^{232}Th/^{238}U$ (Figs. 3 and 4). The observed geochemical and isotopic enrichment in the African domain is relatively evenly distributed along the Atlantic and Indian ridges (Extended Data Fig. 6). Nonetheless, the Southern Hemisphere sections of the Atlantic and Indian ridges in general have somewhat higher Sr and Pb isotopes and lower Hf isotopes than their Northern Hemisphere counterparts (Extended Data Figs. 1 and 6), in line with previous observations[3,4,20].

East–west mantle dichotomy caused by sypercycle plate tectonics

Our results show that heterogeneities in the present-day shallow and deep mantle are not exclusively controlled by the north–south hemispheric subdivision ascribed to the DUPAL anomaly. Instead, they are dominantly controlled by the two antipodal mantle domains, with associated LLSVPs centred on the Equator[6]. Such a global mantle geochemical and isotopic dichotomy between the entire African and Pacific mantle domains can best be explained by the supercontinent cycle and related superocean evolution[9], a mechanism previously invoked to explain just deep mantle signatures[6] (Figs. 2 and 3). In our model (Fig. 5), the composition of the whole mantle has evolved over time through the competing effects of (1) progressive depletion by the extraction of mafic magma (that is, through partial melting of mantle peridotite, most commonly along mid-ocean ridges and arcs) and (2) enrichment by recycling of lithosphere and sediments (rich in Rb, Sr, Th, Pb, Ba, K, Hf and so on) through subduction or crustal delamination[21]. In particular, subduction and collision in the African domain since about 700 Ma during the formation of the latest supercontinent Pangaea (Fig. 5a,b), as well as ongoing subduction towards the African domain since that time (Fig. 5c,d), have enriched both the upper and lower mantle of that

domain[6,9,11,22], where the subduction of continental-derived materials and modification by subduction fluids increased Rb/Sr, Th/Pb, U/Pb and Lu/Hf that would, over time, elevate such isotopic ratios.

The observed whole-mantle east–west geochemical dichotomy, therefore, shows that the two mantle domains (including their LLSVPs) may have been geodynamically and chemically coupled with supercontinent cycles and superocean evolution, thus suggesting a dynamic, instead of stationary, nature of the mantle structure and composition. We note in Fig. 2a the close alignment between the enriched mantle 1 (EM1) trend shown by the African domain data and that of the upper continental crust (UCC), which is lacking in the Pacific domain (Fig. 2b). By contrast, the enriched mantle 2 (EM2) trend not only is well developed in the Pacific domain (Fig. 2b,d,f) but also appears present in the African domain (Fig. 2a,e). We therefore speculate that EM1 may represent the signature of the enrichment of UCC materials during the Rodinia-to-Pangea supercontinent cycle since about 700 Ma (Fig. 5), whereas EM2 could represent the signature of older subduction-related enrichment associated with the Nuna-to-Rodinia supercontinent cycles[13]. Although this suggests secular changes in Earth's dynamic chemical cycling during its evolution, the precise process for such changes is yet to be investigated. High-μ (HIMU, where $\mu = {}^{238}U/{}^{204}Pb$) is present in both mantle domains but in a rather discrete way in St Helena and Cook Austral. There is increasing evidence that the high U/Pb ratio of HIMU mantle domains might be related to more ancient processes (potentially Archaean era)[23] during Earth's earlier history. Future heterogeneity mapping of the mantle, combined with more reliable global reconstructions, geodynamic analysis and four-dimensional numerical simulation, could potentially allow us to achieve a more complete understanding of the evolution of Earth's tectonic plates, its mantle and the underpinning geodynamic processes.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-024-01558-3.

References

1. Mitchell, R. N. et al. The supercontinent cycle. *Nat. Rev. Earth Environ.* **2**, 358–374 (2021).
2. Zhong, S., Zhang, N., Li, Z.-X. & Roberts, J. H. Supercontinent cycles, true polar wander, and very long-wavelength mantle convection. *Earth Planet. Sci. Lett.* **261**, 551–564 (2007).
3. Dupré, B. & Allègre, C. J. Pb–Sr isotope variations in Indian Ocean basalts and mixing phenomena. *Nature* **303**, 142–146 (1983).
4. Hart, S. R. A large-scale anomaly in the Southern Hemisphere mantle. *Nature* **309**, 753–757 (1984).
5. Castillo, P. The Dupal anomaly as a trace of the upwelling lower mantle. *Nature* **336**, 667–670 (1988).
6. Doucet, L. S. et al. Distinct formation history for deep-mantle domains reflected in geochemical differences. *Nat. Geosci.* **13**, 511–515 (2020).
7. Liu, X. et al. Long-lived low Th/U Pacific-type isotopic mantle domain: constraints from Nd and Pb isotopes of the Paleo-Asian Ocean mantle. *Earth Planet. Sci. Lett.* **567**, 117006 (2021).
8. Yang, A. Y. et al. A subduction influence on ocean ridge basalts outside the Pacific subduction shield. *Nat. Commun.* **12**, 4757 (2021).
9. Li, Z. X. et al. Decoding Earth's rhythms: modulation of supercontinent cycles by longer superocean episodes. *Precambrian Res.* **323**, 1–5 (2019).
10. Flament, N., Bodur, Ö. F., Williams, S. E. & Merdith, A. S. Assembly of the basal mantle structure beneath Africa. *Nature* **603**, 846–851 (2022).
11. Yuan, Q. & Li, M. Instability of the African large low-shear-wave-velocity province due to its low intrinsic density. *Nat. Geosci.* **15**, 334–339 (2022).
12. Li, Z.-X. & Zhong, S. Supercontinent–superplume coupling, true polar wander and plume mobility: plate dominance in whole-mantle tectonics. *Phys. Earth Planet. Inter.* **176**, 143–156 (2009).
13. Li, Z.-X., Liu, Y. & Ernst, R. A dynamic 2000–540 Ma Earth history: from cratonic amalgamation to the age of supercontinent cycle. *Earth Sci. Rev.* **238**, 104336 (2023).
14. Stracke, A. A process-oriented approach to mantle geochemistry. *Chem. Geol.* **579**, 120350 (2021).
15. White, W. M. Isotopes, DUPAL, LLSVPs, and Anekantavada. *Chem. Geol.* **419**, 10–28 (2015).
16. Gale, A., Dalton, C. A., Langmuir, C. H., Su, Y. & Schilling, J. G. The mean composition of ocean ridge basalts. *Geochem. Geophys. Geosyst.* **14**, 489–518 (2013).
17. Pearce, J. A., Kempton, P. D. & Gill, J. B. Hf–Nd evidence for the origin and distribution of mantle domains in the SW Pacific. *Earth Planet. Sci. Lett.* **260**, 98–114 (2007).
18. Zhang, Y. & Gan, T. Depletion ages and factors of MORB mantle sources. *Earth Planet. Sci. Lett.* **530**, 115926 (2020).
19. Le Voyer, M. et al. Carbon fluxes and primary magma CO_2 contents along the global mid-ocean ridge system. *Geochem. Geophys. Geosyst.* **20**, 1387–1424 (2019).
20. Patchett, P. J. & Tatsumoto, M. Hafnium isotope variations in oceanic basalts. *Geophys. Res. Lett.* **7**, 1077–1080 (1980).
21. Willbold, M. & Stracke, A. Formation of enriched mantle components by recycling of upper and lower continental crust. *Chem. Geol.* **276**, 188–197 (2010).
22. Pearce, J. A., Kempton, P. D., Nowell, G. M. & Noble, S. R. Hf–Nd element and isotope perspective on the nature and provenance of mantle and subduction components in Western Pacific arc-basin system. *J. Petrol.* **40**, 1579–1611 (1999).
23. Weiss, Y., Class, C., Goldstein, S. L. & Hanyu, T. Key new pieces of the HIMU puzzle from olivines and diamond inclusions. *Nature* **537**, 666–670 (2016).
24. Doubrovine, P. V., Steinberger, B. & Torsvik, T. H. A failure to reject: testing the correlation between large igneous provinces and deep mantle structures with EDF statistics. *Geochem. Geophys. Geosyst.* **17**, 1130–1163 (2016).
25. Plank, T. *Treatise on Geochemistry* 2nd edn (Elsevier, 2014).

Methods

Geochemical and isotopic MORB data used in this study were compiled from ref. 6 and PetDB through the EarthChem Portal (https://search.earthchem.org/)[11,19], and are presented in Supplementary Table 1. The dataset contains data for 3,983 MORB samples (both whole-rock and mineral inclusion analyses) from mid-ocean ridges in the Arctic, Atlantic, Indian and Pacific oceans (Fig. 1). The MORB samples were classified as belonging to the African or Pacific domains based on their geological locations, with the circum-Pacific subduction system being the boundary between the two domains (Fig. 1). No further filtering was applied. The OIB and O-LIP data are from PetDB through the EarthChem Portal (https://search.earthchem.org/). For samples with available La and Sm data, we followed the approach in ref. 16 to classify the different types of MORBs using La_N/Sm_N ratios (La and Sm being normalized to primitive mantle estimate[26]). E-MORB has $La_N/Sm_N > 1.5$, N-MORB has La_N/Sm_N ranging from 0.8 to 1.5 and D-MORB has $La_N/Sm_N < 0.8$. Samples without La and Sm data could not be classified, and are thus labelled as 'unclassified MORB'. In total, 2,054 out of the 3,983 samples (~52%) are 'unclassified', with 570/3,983 (~14%) being N-MORB, 1,194/3,983 (30%) being D-MORB and 165/3,983 (4%) being E-MORB. Among the 1,929 samples that could be classified, N-MORB represents 30%, D-MORB 62% and E-MORB 9%. If we extrapolate these results to the entire dataset, we can speculate that E-MORB represents only a small portion of the data, while D-MORB represents the vast majority. It is important to note that we followed the N-MORB definition of "any ridge segment >500 km away from any hotspot" given in ref. 16. In detail, for the Pacific domain, 9% of MORB samples are N-MORB, 30% are D-MORB, 2% are E-MORB and close to 60% are unclassified. For the African domain MORB samples, 22% are N-MORB, 31% are D-MORB, 8% are E-MORB and ~40% are unclassified.

Kernel density plots, corrected for geographical bias (see below), for both MORB and OIB data from the African and Pacific mantle domains are shown in Figs. 2 and 3, and Supplementary Figs. 1–4 (for MORB) and Supplementary Fig. 5 (for OIB and oceanic-LIP). A location-based weighting approach was employed to ensure equal representation of samples from each domain. The data were grouped by location, and a weighting scheme was applied where the weight assigned to each sample was inversely proportional to the number of samples from that location. This approach ensured that variations in sampling intensity had no influence on the density estimates.

Density overlap for the two domains was quantified by computing the overlap area between their kernel density estimation distributions for both MORB and OIB. Specifically, we summed the minimum density values at each overlapping grid point and expressed this as a percentage of the total density. This overlap percentage measures the similarity between the isotopic compositions of the Atlantic and Pacific domains.

Machine learning is applied in this paper to analyse MORBs. The code 'morb_ml_model.ipyb' was written using the open-source Python 3.10 programming language and open-source libraries and can be run in Jupyter Notebook (Supplementary Code 1). Overall, the code pre-processes the data, performs iterative imputation, normalizes the data and finally trains the machine using CNN, random forest and QDA methods for the classification of MORBs based on selected geochemical features. In detail, the code begins with importing necessary libraries: Pandas, NumPy, Seaborn, Matplotlib.pyplot, Sklearn and Tensorflow.keras. The data, stored in a Parquet file (Excel version also available), are read into a Pandas Data Frame using the 'read_parquet' function. A list of MORB types (E-, N- and D-MORB for each domain) is defined alongside specific features: La/Sm, Ba/La, $^{87}Sr/^{86}Sr$, $^{143}Nd/^{144}Nd$, $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$, $^{208}Pb/^{204}Pb$ and $^{176}Hf/^{177}Hf$ for the analysis. Next, separate Pandas Data Frames are created for African and Pacific MORBs with different MORB types and selected features. As the datasets are sparse, data imputation is then performed using IterativeImputer from the sklearn.impute library with a maximum of 500 iterations and the random state set to 0. IterativeImputer imputes missing data by modelling each feature with missing values as a function of other features (correlation or absence of correlation) and using that model to impute the missing values. Unlike traditional imputation methods that rely on simple column-wise statistics such as mean, median or mode, the 'IterativeImputer' method employs a comprehensive multivariate approach, leveraging the entire set of available features to impute missing values based on complex inter-feature relationships. This technique not only preserves the inherent data structure but also enhances the quality of predictive models by iteratively refining imputations, ensuring convergence. The imputed data are then concatenated and normalized using Z-score normalization, a technique that rescales the input data to have a mean of 0 and a standard deviation of 1 in preparation for machine learning training. Z-score normalization follows the equation:

$$Z = \frac{X - \mu}{\sigma}$$

where X is the original value, μ here is the mean of the data and σ is the standard deviation of the data. Normalization is a crucial step to mitigate the effects of differences in the distribution of input data and improve the convergence rate of the optimization algorithm used to train the CNN. The data are then split into a training dataset (randomly selecting 33% of the original dataset for each MORB group) and a testing dataset using the train_test_split function. For the CNN, the data are then reshaped into a three-dimensional tensor to fit the input shape of the CNN. The target variable is one-hot encoded using the to_categorical function. A CNN is created using the Sequential class in the tensorflow.keras library. As the objective of using CNN is to test whether the observations and classifications are consistent, we use rather conventional hyperparameter settings. The CNN comprises a Conv1D layer followed by MaxPooling1D, Flatten and Dense layers. The CNN is trained using the fit function with a batch size of 32, 1,000 epochs and a validation split of 0.2. The initial configuration of 32 filters and a 3×3 kernel size is a popular choice for balancing feature detection with computational efficiency. This set-up is enhanced by the rectified linear unit (ReLU) activation function, which mitigates the vanishing gradient problem, thereby accelerating the learning process. A MaxPooling layer with a pool size of 2 reduces spatial dimensions and computational load, while dense layers with numerous neurons, such as 128 with rectified linear unit activation, interpret high-level features. The Adam optimizer, known for handling sparse gradients and adaptive learning rates, is a standard choice. For multi-class classification, the categorical_crossentropy loss function is typically used. Choices of batch size and epoch count, such as 32 and 1,000, are adjusted based on dataset complexity and resource availability, often necessary for convergence in complex scenarios. Finally, for each machine learning model, the confusion matrices are plotted using the confusion_matrix function from the sklearn.metrics library (Extended Data Fig. 5). The implemented models performed well in classifying MORBs based on selected geochemical features. The confusion matrices showed that the CNN, random forest and QDA methods were all able to classify MORBs into their respective types with high accuracy, with an overall accuracy > 0.9 (Extended Data Fig. 5). The overall results indicate that the two mantle domains show distinct E-, N- and D-MORB compositions.

Data sources

All data supporting the findings of this study are available from the online database PetDB through the EarthChem Portal (https://search.earthchem.org/).

Data availability

All data used in this study are compiled from ref. 6 and PetDB through the EarthChem Portal. A full compilation can be found in Supplementary Table 1. Source data are provided with this paper.

Code availability

All the codes are available upon request to the corresponding authors. The code used in this study is provided in Supplementary Code 1.

References

26. McDonough, W. F. & Sun, S.-S. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).

Acknowledgements

L.S.D. and Z.-X.L. thank the team members of the Earth Evolution and Dynamics Research Centre, including members of the Earth Dynamics Research Group in Perth. This work was supported by Australian Research Council Future Fellowship grant to L.S.D. (FT230100348), Australian Research Council Laureate Fellowship grant to Z.-X.L. (FL150100133) and the National Natural Science Foundation of China grant to Z.-X.L. (42350710792). This is a contribution to IGCP 648: Supercontinent Cycles and Global Geodynamics.

Author contributions

L.S.D. and Z.-X.L. contributed equally to the work.

Competing interests

The authors declare no competing interests.

Additional information

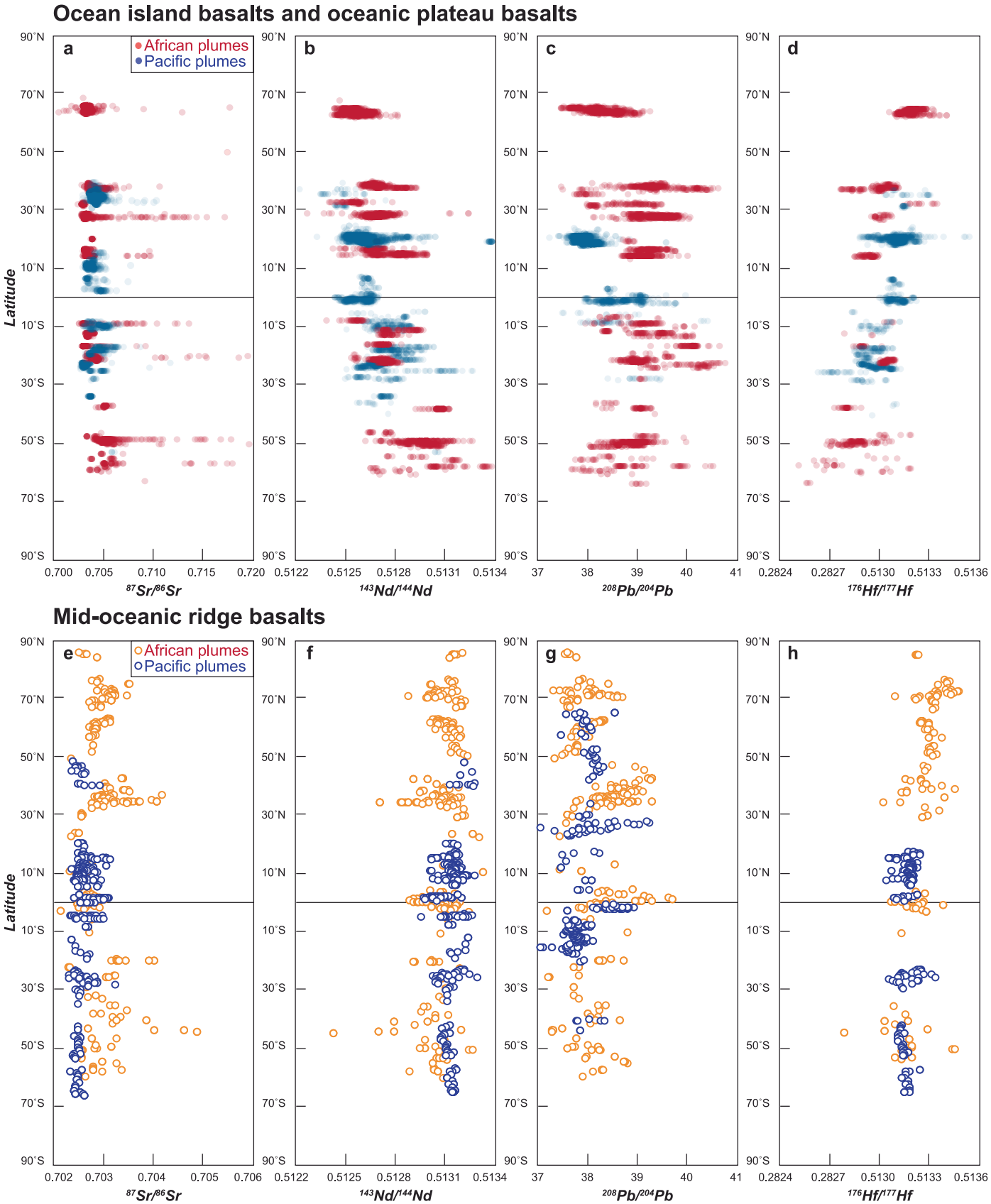
Extended data is available for this paper at https://doi.org/10.1038/s41561-024-01558-3.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-024-01558-3.

Correspondence and requests for materials should be addressed to Luc S. Doucet or Zheng-Xiang Li.

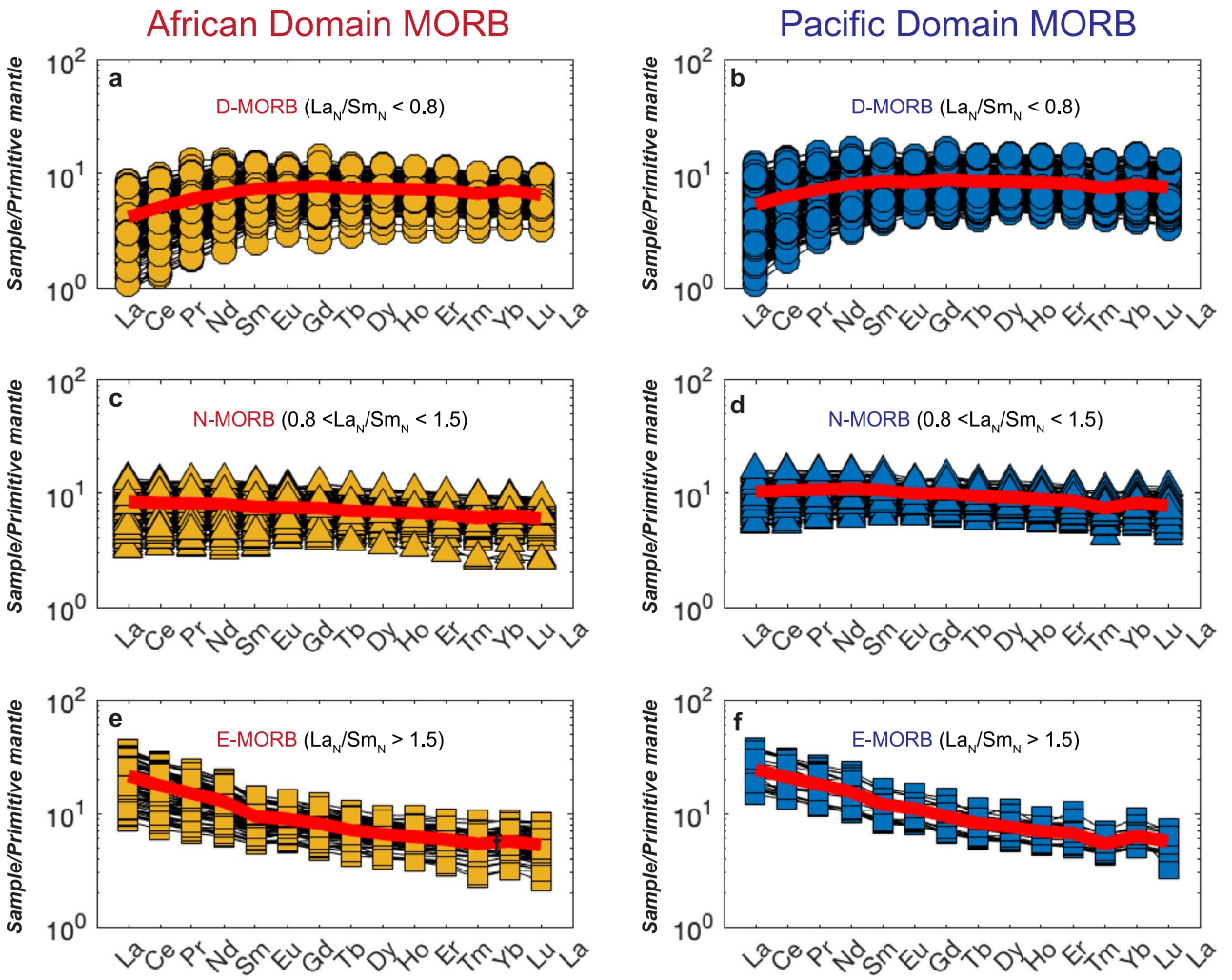
Peer review information *Nature Geoscience* thanks Guoxiong Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Alison Hunt, in collaboration with the *Nature Geoscience* team.

https://doi.org/10.1038/s41561-024-01558-3



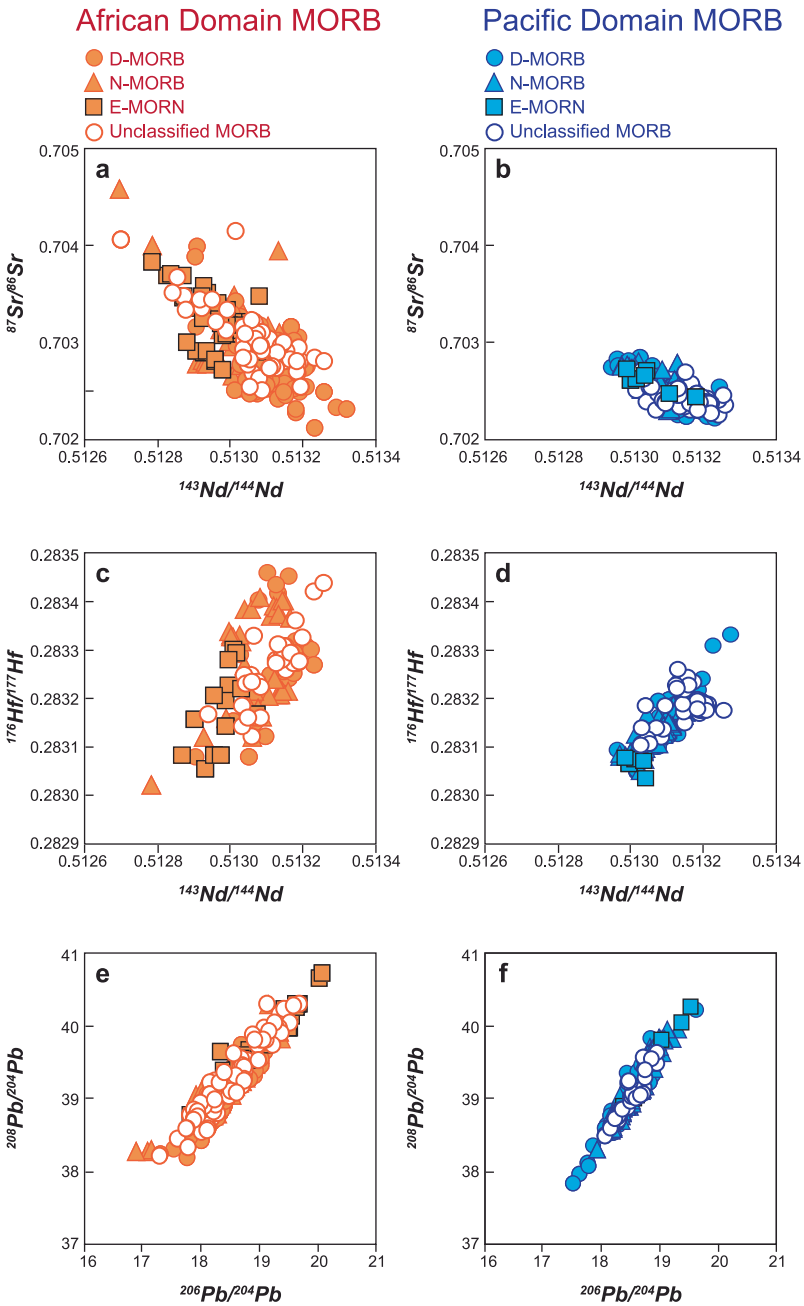
Extended Data Fig. 1 | Isotopic composition of OIB, oceanic-LIP and MORB, plotted against latitude for the African and Pacific domains. Isotopic composition of OIB and oceanic-LIP (**a**, **b**, **c**, and **d**), as well as MORB (**e**, **f**, **g**, **h**), plotted against latitude for the African and Pacific domains.

https://doi.org/10.1038/s41561-024-01558-3



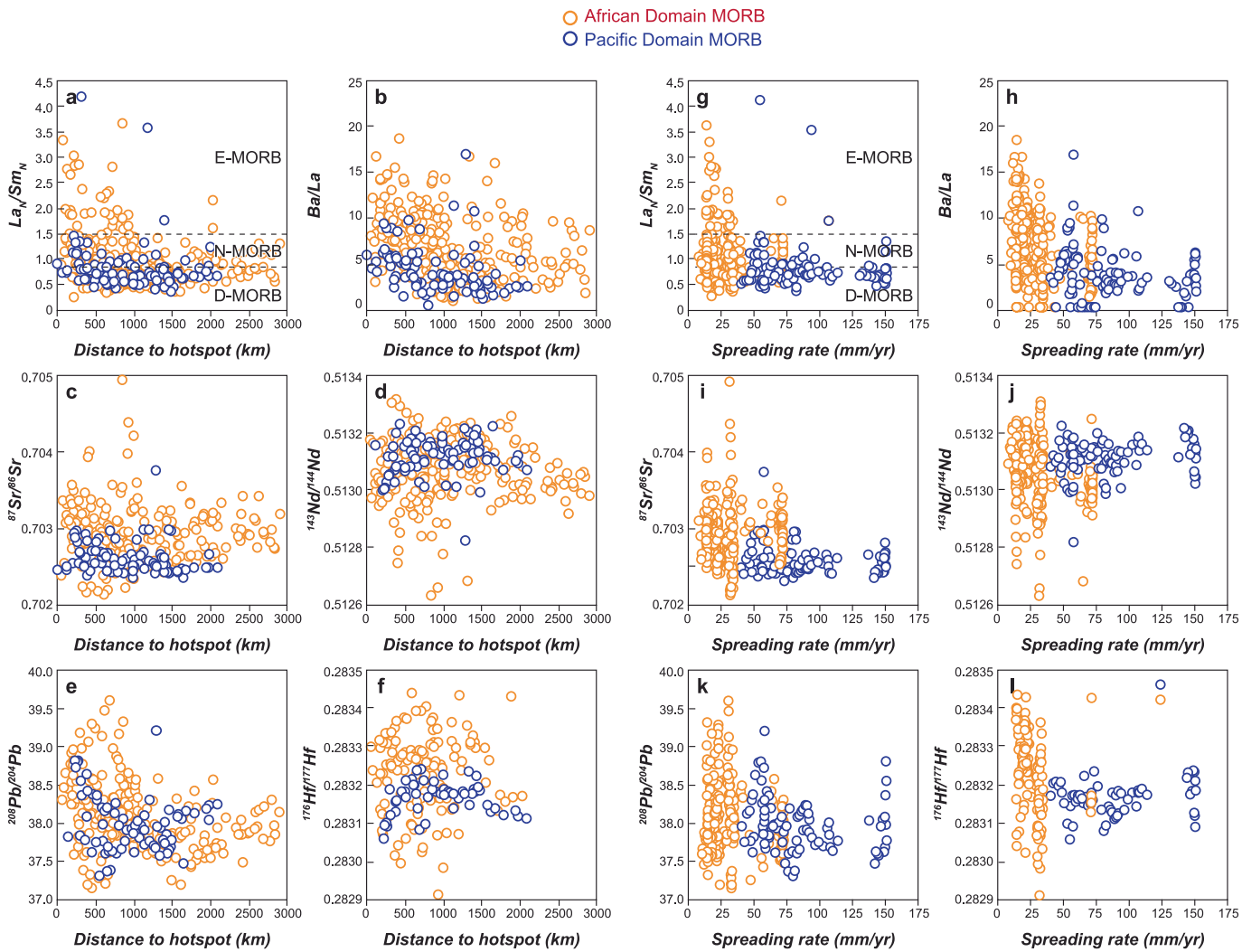
Extended Data Fig. 2 | Rare earth element patterns forMORB from the African and Pacific domains. Mantle-normalized rare earth element patterns of D- (**a**, **b**), N- (**c**, **d**), and E-MORB (**e**, **f**) for the African and Pacific domains, following the Gale et al.[16] classification.

https://doi.org/10.1038/s41561-024-01558-3



Extended Data Fig. 3 | Isotope composition for unclassified-, E-, N-, and D-MORB for the African and Pacific domain. The unclassified MORB data generally plot within the distribution fields of N-, D- and E-MORB samples.

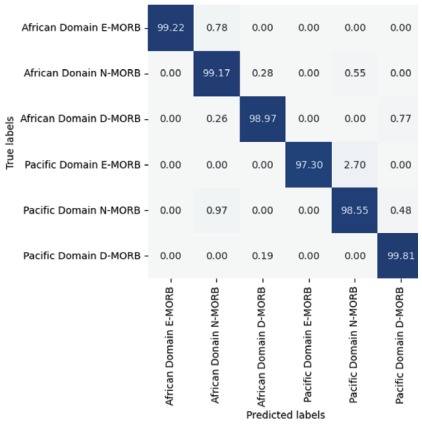
https://doi.org/10.1038/s41561-024-01558-3



Extended Data Fig. 4 | Isotopic composition of MORB as function of distance to nearest hotpot and degree of partial melting. Isotopic and geochemical composition of E-, N-, and D-MORB for the African and Pacific domains vs. the distance to the nearest hotspot (**a**, **b**, **c**, **d**, **e** and **f**) and spreading rate (**g**, **h**, **i**, **j**, **k**, and **l**) used as proxy for degree of partial melting.

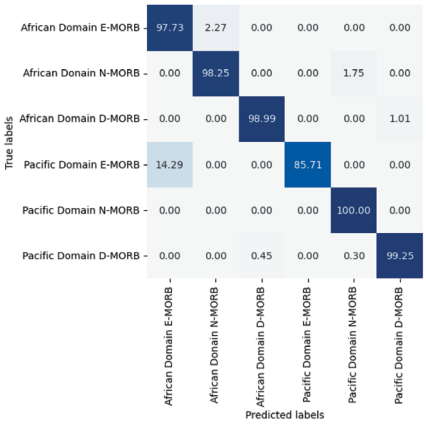
https://doi.org/10.1038/s41561-024-01558-3

a) Convolutional Neural Network



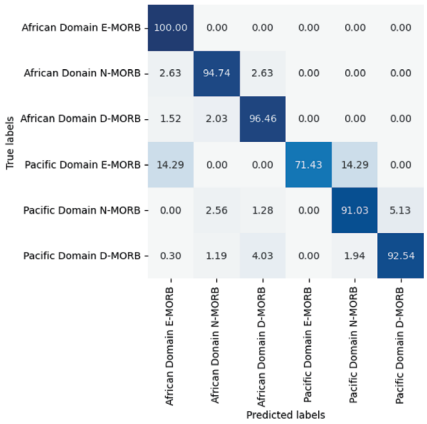
	precision	recall	f1-score	support
African Domain E-MORB	1.00	0.99	1.00	128
African Donain N-MORB	0.98	0.99	0.99	363
African Domain D-MORB	1.00	0.99	0.99	1170
Pacific Domain E-MORB	1.00	0.97	0.99	37
Pacific Domain N-MORB	0.99	0.99	0.99	207
Pacific Domain D-MORB	1.00	1.00	1.00	2078
accuracy			0.99	3983
macro avg	0.99	0.99	0.99	3983
weighted avg	0.99	0.99	0.99	3983

b) Random Forrest



	precision	recall	f1-score	support
	1.00	1.00	1.00	128
	0.99	1.00	0.99	363
	1.00	0.99	0.99	1170
	1.00	0.97	0.99	37
	1.00	0.99	0.99	207
	0.99	1.00	1.00	2078
			0.99	3983
	1.00	0.99	0.99	3983
	0.99	0.99	0.99	3983

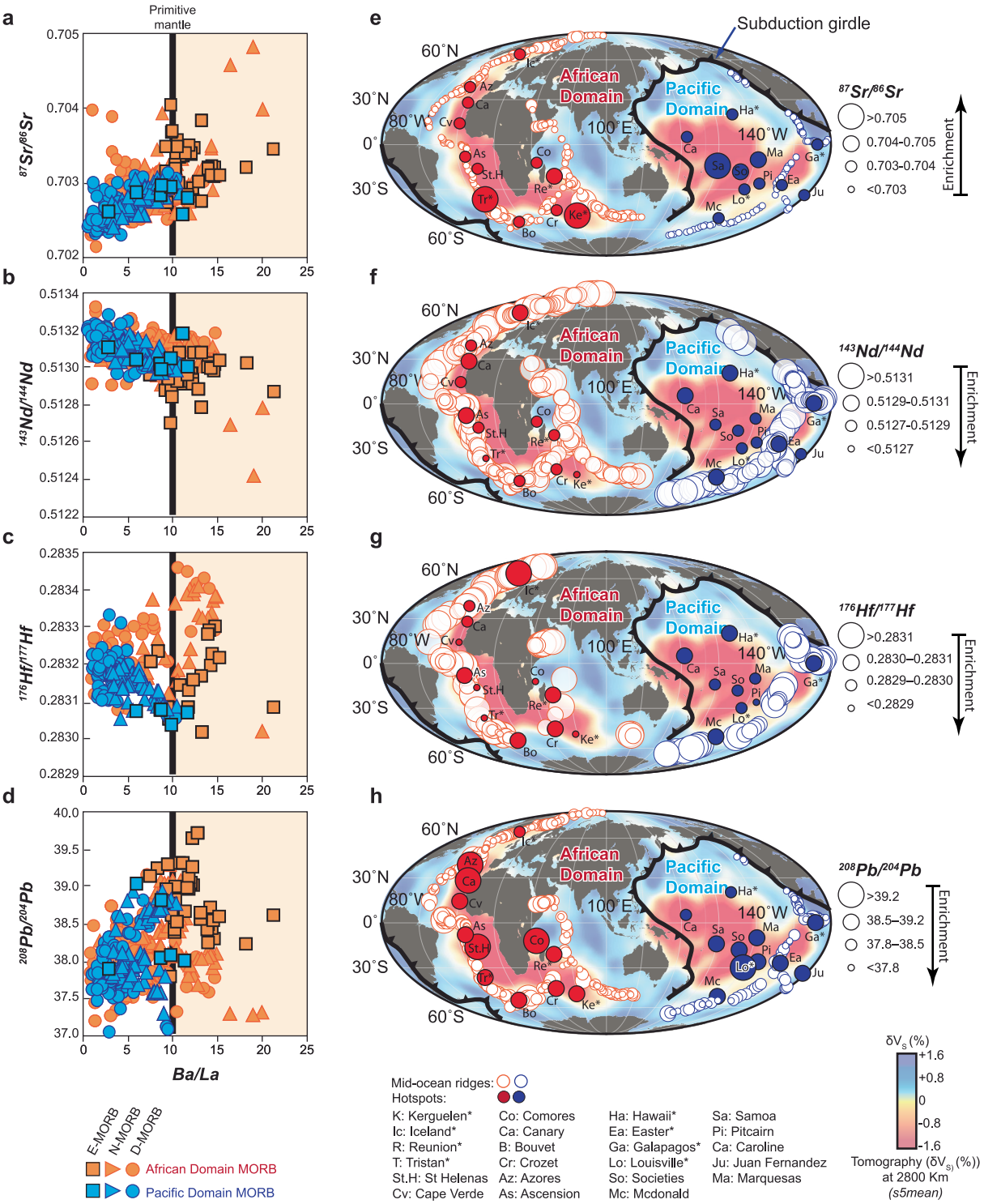
c) Quadratic discriminant analysis



	precision	recall	f1-score	support
	1.00	1.00	1.00	128
	0.99	1.00	0.99	363
	1.00	0.99	0.99	1170
	1.00	0.97	0.99	37
	1.00	0.99	0.99	207
	0.99	1.00	1.00	2078
			0.99	3983
	1.00	0.99	0.99	3983
	0.99	0.99	0.99	3983

Extended Data Fig. 5 | Confusion matrix and report for the three machine learning models used in this study. Machine learning methods applied include convolutional neural network (CNN) (**a**), random forest (**b**), and quadratic discriminant analysis (QDA) (**c**). Confusion matrix are calculated for k = 100 k-fold cross-validation tests using random train/test split of 67/33. The predictions demonstrate an overall accuracy of 99%. Precision = TP / (TP + FP). Recall = TP / (TP + FN). F1 Score=2 × 1/Precision+1/Recall. Accuracy = (TP + TN) / (TP + TN + FP + FN).Where TP, the number of true positives, TN, the number of true negatives, FP, the number of false positives, and FN, the number of false negatives.

https://doi.org/10.1038/s41561-024-01558-3



Extended Data Fig. 6 | Geographic distribution of isotopic heterogeneity for MORB, OIB and oceanic-LIP in the African and Pacific domains. Comparative isotopic plots of data from the two mantle domains are given in **a**–**d**. The level of enrichment in each isotopic data point is represented by the size of the circle (**e**–**h**). Seismic shear-wave velocity anomalies (δVs) are those of the mean S-wave tomography model s5mean of Doubrovine et al.[24]. The map shows the positions of the two mantle domains and their respective LLSVPs, and the circum-Pacific subduction girdle (thick black line) (**e**–**h**). The map was produced using open-source programming language Python 3.0 and open-source library pygmt.