Seismic evidence for a 920-km discontinuity in the mantle

Hitoshi Kawakatsu & Fenglin Niu

Earthquake Research Institute, University of Tokyo, 1-1-1 Yayoi Bunkyo-ku, Tokyo 113, Japan

Analysis of hundreds of seismograms from a seismic array in the Japanese islands reveals a seismic discontinuity at a depth of about 920 km in the mantle beneath the Tonga subduction zone. This discontinuity is also evident beneath subduction zones in the Japan and Flores seas. The discontinuity probably represents either a change in mantle chemical composition or the signature of the subducted slab's garnet layer.

ANSWERS to many key questions about mantle dynamics hinge on the nature of the mantle transition zone[1]. For example, the pattern of mantle convection depends strongly on whether the transition from the upper mantle to the lower mantle is related to a chemical or a phase change of the mantle material. In the first-generation models of the Earth's seismic velocity structure, Bullen[1] defined the mantle transition zone (region C) to be the region between the depths of 410 km and 980 km, where the seismic wave velocity changes drastically, though continuously, compared to that above and below.

Detailed seismological analyses in the past 30 years have yielded the current generation of Earth models[2,3] in which the mantle transition zone consists of two major discontinuities at depths of ~410 km and ~660 km. With the help of the results from high-pressure experiments on rock minerals, we now seem to have a rather universal (although still controversial) view of the nature of the transition zone: the 410-km and 660-km discontinuities are due to the phase transitions of olivine→spinel→perovskite + magnesiowüstite[4,5], and the lower mantle starts just below the 660-km discontinuity[6] with a rather smooth velocity gradient down to near the mantle-core boundary[2,3]. There exists however, appreciable evidence suggesting that the velocity gradient changes in the lower mantle. Summarizing earlier work[7 11], Johnson[12] concluded that most studies had identified change of velocity gradient in the depth ranges of 900–1,000, 1,200–1,300 and 1,900–2,000 km. Some of the recent studies have also identified reflected or converted seismic waves from those depth ranges, suggesting a possible presence of seismic discontinuities in the lower mantle[13 16].

In the past few years, signal enhancement techniques applied to a large number of seismograms have been successful in identifying many reflected or converted phases at seismic discontinuities in the upper mantle, helping us to obtain a better picture of the dynamic state of the Earth's deep interior[17-21]. In this paper, we use a large number of seismograms from Japanese stations to show that there is a discontinuity at a depth of ~920 km in the mantle. This discontinuity may actually represent the bottom of the transition zone and thus be the top of the lower mantle.

Signal enhancement with J-array

The J-array is a large-aperture short-period (1 Hz) seismic array across Japan which began recording in April 1991[22]. It combines several regional seismic networks developed under Japan's national programme for earthquake prediction. With an aperture of 3,000 km in length by 300–500 km in width, it covers the entirety of the Japanese islands (Fig. 1). With more than 300 high-quality short-period vertical-component seismographs, the J-array is ideally suited for analysing the weak signals of reflected or converted waves associated with mantle discontinuities.

In Fig. 2, we indicate the type of rays we wish to observe in the P-wave coda (the portion of records after the arrival of the

direct P wave) if discontinuities are present just below or above deep earthquakes. The notation p_dP (or S_dP) denotes a wave that propagates upward (downward) from the source as a P wave (S wave) and is subsequently reflected as (converted into) a P wave at a mantle discontinuity located at depth d km from the Earth's surface. In contrast to reflected waves, which have incident angles that are shallower than the direct P waves, waves converted from S to P (S_dP) are received at incident angles that are steeper than the direct P wave, as they have smaller values of slowness (the reciprocal of apparent velocity). We use S–P converted waves to constrain the depths of discontinuities. The travel time difference between S_dP and the direct P wave is a simple function of the vertical distance between the source and the mantle discontinuity for a laterally homogeneous structure. Therefore, the depth of the discontinuity can be constrained using estimates of the travel-time difference, assuming we have reasonable estimates of the source depth and that lateral inhomogeneity is relatively insignificant.

Data analysis. We use all earthquakes with depths greater than 500 km that occurred in the Tonga subduction zone and were
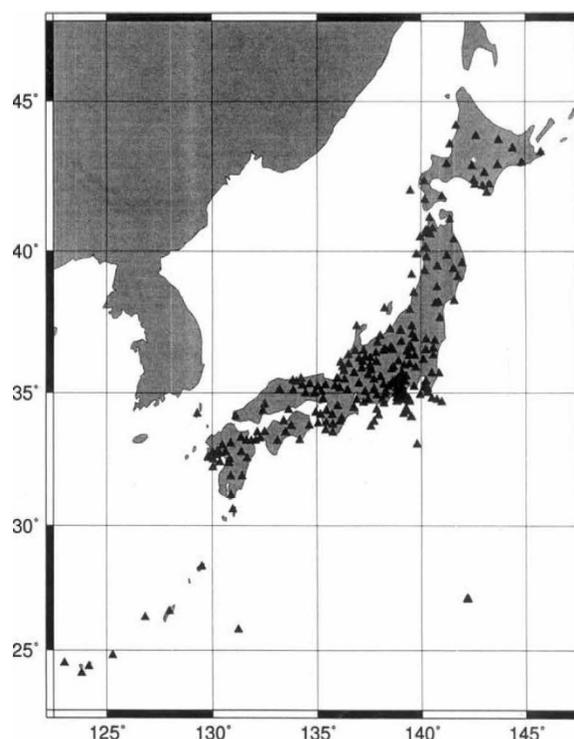
FIG. 1 Location of the J-array stations. The J-array's 'J' comes not only from 'Japan', but also from the shape of the Japanese islands.
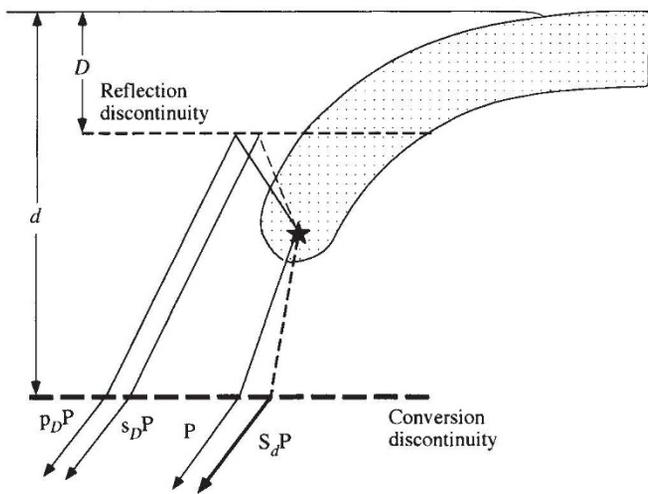
FIG. 2 Typical geometry of seismic rays used in this study to investigate transition zone discontinuities. Lower-case (p, s) and upper-case (P, S) letters are used for waves that initially propagate upward and downward from the source (star) respectively. Solid and broken lines indicate P and S waves respectively.

recorded by the J-array stations. There are a total of seven such events (for the events occurring in 1990 before the operation of the J-array, we use the data of the regional network of the Earthquake Research Institute of the University of Tokyo, which covers the central part of Japan). The epicentral distances are in the approximate range 65°–77°. We re-determine the hypocentral depths from handpicked pP phases using J-array data and globally distributed broadband data (Table 1). For the case of an unambiguous pP arrival, the source depth variance is as small as 3 km. In general, an accuracy of ±10 km in hypocentral depth estimate is expected.

Figure 3a shows an example of the J-array recording of a deep earthquake in Tonga. Because individual records are invariably contaminated by near-receiver reverberations and other noise, it is very hard to identify later-arriving phases that are the signature of mantle discontinuities. However, by stacking individual records the signal-to-noise ratio is significantly improved, facilitating the identification of later-arriving phases[18]. Seismograms with favourable signal-to-noise ratios are selected and processed with a 2-s low-pass filter. The seismograms are subsequently aligned such that the maximum amplitude of the direct P wave, which is normalized to unity, occurs at time zero (when necessary, the polarity of the seismogram is reversed). Approximately 40–120 station records with a window of 150 s are selected for stacking.

For an incoming wave with an apparent velocity v (or an equivalent, slowness p), the time lag τ_j for the jth station is
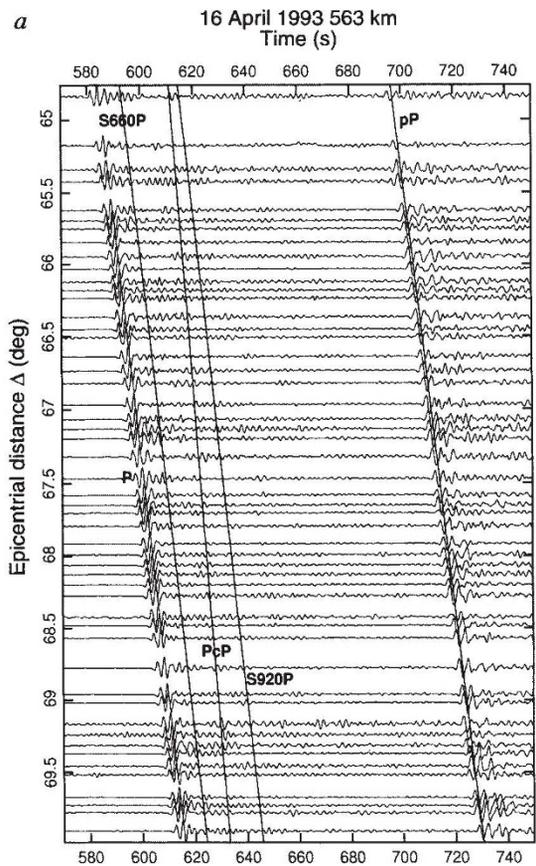
FIG. 3 a, Example of a portion of a seismic section for an event occurred in Tonga. Theoretical travel times of the iasp91 model for major expected bodywave arrivals are indicated by straight lines. Note that except for P and pP phases, other phases are difficult to identify in individual traces. b, A colour contour map of the stacked section of the seismograms shown in Fig. 3a. The slowness assumed for each stack is varied with respect to that of the direct P arrivals (defined as zero) in increment of 0.1 s deg^{-1} within the range of ±3 s deg^{-1}. The resulting 61 stacked waveforms are subsequently converted to amplitude envelopes using the Hilbert transform. The maximum amplitude is chosen from all 61 stacked traces and is used to normalize the traces in a unit of decibel. The final result is indicated by a colour contour map in which 'hotter' colour clusters represent greater energy and may represent possible phase arrivals. Note that many later phases between P and pP are present. Major phases are identified by arrows.
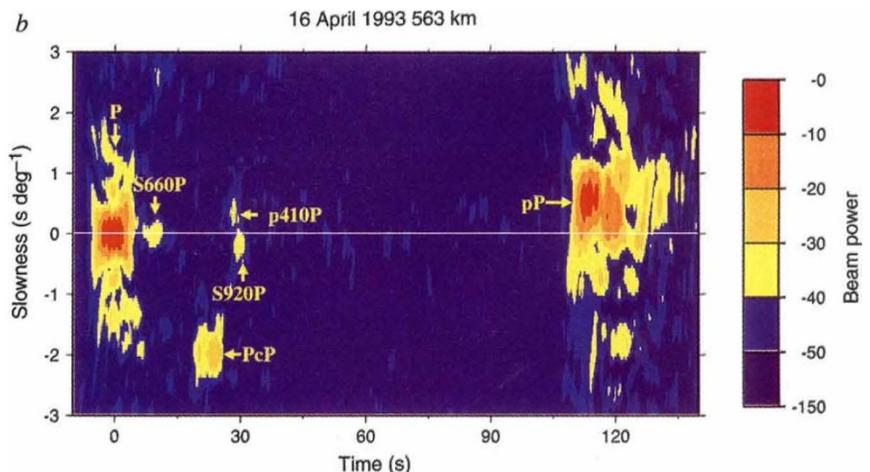
TABLE 1 Earthquakes and observed parameters

Date (yr mo d)	Time (h min s)	Lat.	Long.	Depth (km)‡	Region	Mag. mb	'660-km' (T P D)*			'920-km' (T O E D)†				S$_{660}$P/P (O C)§		S$_{920}$P/S$_{660}$P (O C)§	
1990 06 08	15 05 10.3	−18.70	181.10	518.	Tonga	5.9	13.5	−0.05	660.	35.0	−0.10	−0.14	920.	0.080	0.068	0.43	0.44
1990 07 22	09 26 18.0	−23.48	179.93	535.	Tonga	5.8	10.5	0.10	655.			−0.13		0.270	0.124		0.27
1993 08 07	17 53 28.5	−23.60	179.10	553.	Tonga	5.9	13.8	0.00	710.	31.8	−0.12	−0.12	925.	0.080	0.117	0.38	0.18
1991 12 03	10 33 42.0	−26.31	178.57	559.	Tonga	5.8	14.4	−0.10	730.	30.5	−0.10	−0.11	922.	0.100	0.163	0.25	0.21
1993 04 16	14 08 40.0	−17.40	181.10	563.	Tonga	5.9	9.7	−0.05	670.	31.1	−0.10	−0.13	925.	0.095	0.110	0.74	0.64
1992 08 30	20 09 06.9	−17.74	181.23	565.	Tonga	5.9	12.8	−0.05	710.	28.6	−0.12	−0.13	900.	0.021	0.083	0.95	0.24
1991 09 30	00 21 47.5	−20.88	181.41	595.	Tonga	6.3	8.0	−0.05	685.	25.7	−0.10	−0.11	900.	0.083	0.081	0.63	0.34
1993 01 19	14 39 26.4	38.57	133.52	434.	Japan Sea	6.0	20.0	−0.08	660.	43.8	−0.10	−0.13	943.	0.055	0.246	0.42	0.22
1991 06 07	11 51 27.6	−7.26	122.57	543.	Flores Sea	6.2	18.0	−0.10	720.	39.0	−0.28	−0.27	945.	0.070	0.096	0.31	0.34

* T, arrival time (s); P, slowness (s deg^{-1}); D, Depth (km).
† T, arrival time (s); O, observed slowness (s deg^{-1}); E, expected slowness (s deg^{-1}); D, depth (km).
‡ Depth from pP–P.
§ O, observed amplitude ratio; C, calculated amplitude ratio.

$\tau_j = D_j/v = D_j \cdot p$, where D_j is the epicentral distance of the jth station minus the epicentral distance of the centre of the array. Let x_{ij} represent the amplitude at the jth station at the ith time for the case with K stations. We use Nth root stacking[23]; for a given slowness p, an Nth root stack, $y(p)$, is given by

$$y_i(p) = R_i(p)\,|R_i(p)|^{N-1} \qquad (1)$$

where

$$R_i(p) = \frac{1}{K}\sum_{j=1}^{K} \text{sign}\,(x_{i+\tau_j,j})\,|x_{i+\tau_j,j}|^{1/N}$$

N can be assigned any value depending upon the signal quality. When $N=1$, equation (1) reduces to the usual linear slant stack[18]; here we use $N=2$. Figure 3b depicts a representative example of an Nth root stack for the records shown in Fig. 3a.

The Nth root stack analysis has two advantages over examining individual seismograms: (1) the signal-to-noise ratio is improved significantly, and the later-arriving phase S$_d$P is therefore clear enough to be readily identified. Although many later arriving phases can be observed in linear stacks with $N=1$, by choosing $N=2$, small signals show up more clearly; (2) not just the arrival time but also the slowness of a particular phase is obtained, which aids correct identification of the phase. As described above, an S$_d$P phase is characterized by a smaller value of slowness than that of the direct P wave, which effectively eliminates the possibility of misidentification of direct P arrivals due to aftershocks. Slowness thus plays a critical role in the identification of later phases.

The 920-km discontinuity identified

In Fig. 4a, Nth root stacked traces of the seven deep events are plotted with respect to source depth. In this figure, many late-arriving phases are evident and the most prominent phase, which is denoted by a closed circle, is identified as the S–P conversion wave associated with the 660-km discontinuity (S$_{660}$P). The S$_{660}$P arrivals have a maximum delay of approximately 6 s from the theoretical time, which corresponds to a 60-km deepening of the 660-km discontinuity[18,20] (the result of our analysis of the geometry of the 660-km discontinuity in Tonga will be presented in a separate paper[24]). After S$_{660}$P, another phase (denoted by an open circle) is also notable (except for the event 22 July 1990). Two characteristics of this phase are observed: (1) this phase arrives ∼35 s after the P wave for a source depth of 518 km, and the relative arrival time decreases to ∼25 s for a source depth of about 600 km; (2) the slowness of this phase, shown in Fig. 4b, is determined to be slightly less than that of S$_{660}$P. Detailed observed travel-time and slowness values are listed in Table 1.

On the basis of these observations, this phase is identified using the iasp91 model as an S–P converted wave with a conversion depth of ∼920 km (S$_{920}$P). In other words, the existence of a discontinuity near 920-km depth beneath the Tonga subduction zone is strongly suggested. The observation that a

straight line fits through the S$_{920}$P arrivals in the depth section (Fig. 4a) indicates that this discontinuity is relatively flat in comparison with the 660-km discontinuity. The detailed conversion depths for each event are also shown in Table 1. We also note that several other large amplitude peaks, which may correspond to other phases, are observed in Fig. 4, but none of them lines up as nicely as S$_{920}$P. We may need to invoke a three-dimensional structure, such as the subducting slab, to explain some of those peaks (we see some evidence for p$_{520}$P[21], but not for p$_{210}$P[18]).

Relative amplitudes of S$_{920}$P to S$_{660}$P have a range of 0.25–0.95, with an average near 0.5 (Table 1). According to the iasp91 model, for a focal depth of about 580 km, a reflection of an upgoing p wave at the 410-km discontinuity, p$_{410}$P is characterized by about the same travel time as S$_{920}$P. Consequently, the overlap of these two waves may produce a large amplitude ratio for events 16 April 1993, 30 August 1992 and 30 September 1991. Omitting these events in the estimation of S$_{920}$P/S$_{660}$P, the ratio may be reduced to ∼0.4. Considering that S$_{920}$P and S$_{660}$P are characterized by only slightly different take-off angles during the initial radiation as S waves from the source, similar amplitudes are expected for these two phases before P wave conversion. It is therefore most likely that the S-wave velocity contrast (the amplitude of the S$_d$P phases depends almost entirely on S-velocity contrast) at the 920-km discontinuity is about 40% of that of the 660-km discontinuity. Theoretical ratios of S$_{660}$P/P and S$_{920}$P/S$_{660}$P are also given in Table 1. They are estimated from synthetic seismograms stacked in the same way as for the observed data. Reflectivity synthetic seismograms[25] are calculated assuming known focal mechanisms[26] for a modified iasp91 model, which has an extra discontinuity at a depth of 920 km with a velocity jump of 40 percent of the 660-km discontinuity (that is ∼2.4% S-wave velocity jump). On average, theoretical and observed S$_{920}$P/S$_{660}$P are quite similar, supporting the estimated size of the 920 km discontinuity. However, we could not identify S$_{920}$P for the 22 July 1990 event (Fig. 4a, b), although the theoretical calculation suggests that S$_{920}$P should be large enough to be observed. The large discrepancy between observed and theoretical S$_{660}$P/P suggests that the assumed focal mechanism may be incorrect. The expected arrival times of S$_{920}$P and s$_{410}$P are almost the same and the destructive overlap of two phases may be masking their presence in the stacked traces, because only a limited portion of J-array data are available for this event, reducing the slowness resolution in stacking. As to the thickness of the 920-km discontinuity, considering that S$_{920}$P is also observable in the period range as short as 1–2 s, the thickness of the 920-km discontinuity may be of the order of 10 km.

To see whether the inferred 920-km discontinuity is simply a local feature, we analysed deep earthquakes that occurred in Japan sea (19 January 1993) and in Flores sea (7 June 1991) recorded respectively by the Southern California Seismographic Network and the J-array. Figure 5 shows the stacked results for these two events. Following the direct P arrivals, two notable
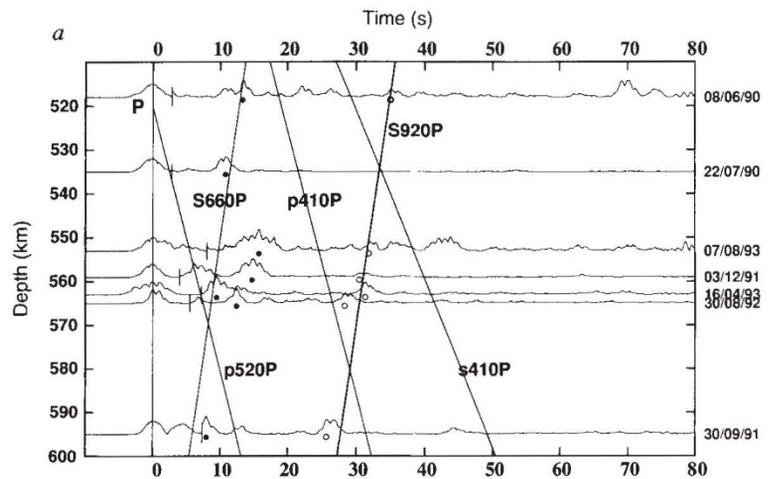
FIG. 4 a, Depth section of the stacked traces for a slowness of -0.1 s deg^{-1}. To show later arriving phases more clearly, traces to the right of the vertical line are enlarged so that the amplitude of S$_{660}$P becomes the same as that of direct P. Theoretical travel times of the iasp91 model for major expected bodywave arrivals at the epicentral distance $\Delta = 70°$ are indicated by straight lines. The thick line indicates the theoretical arrival time of S$_{920}$P. Event dates are given in the order day/month/year. b, Colour contour maps of the individual events. Colour scaling is similar to Fig. 3b, but modified slightly for each event to help observing major arrivals. Expected arrival times of S$_{920}$P are indicated by vertical red arrows. White arrows indicate S$_{920}$P.
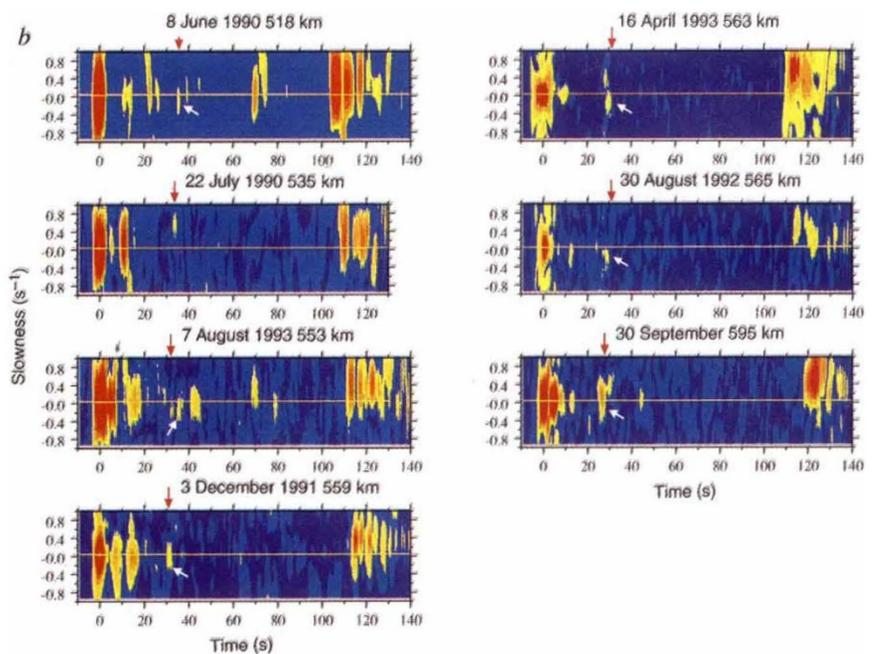
FIG. 5 Stacked section of two events in other subduction zones. The red arrows indicate expected arrival times of S$_{945(3)}$P as in Fig. 4b. For the Flores sea event, the expected slowness of S$_{945}$P is -0.27 s deg^{-1}. This is smaller than that for the other events because the epicentral distance of the J-array, $\Delta = 45°$, is smaller for this event. The '920-km discontinuity' is also present in these regions.
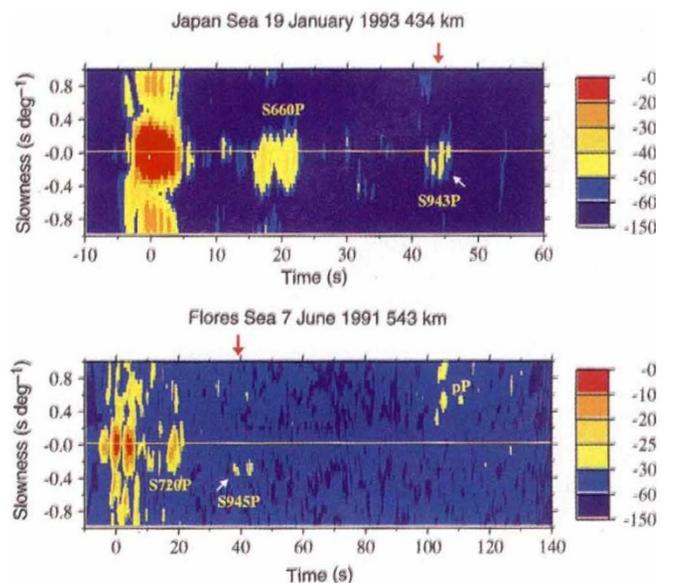
phases with small negative slowness are evident. Their slowness indicates that these two phases are S–P converted waves. The second phases correspond to conversion depths of 943 km and 945 km for the Japan and Flores seas respectively.

Global feature?

We have shown above the evidence for a 920-km discontinuity in the deep mantle beneath three subduction zones. It is natural to wonder whether or not it is a global feature. Global confirmation of the 920-km discontinuity is, however, limited by the geographical distribution of the occurrence of deep earthquakes and seismic arrays. Consequently, it is difficult to determine whether this discontinuity exists globally or is restricted to the vicinity of subducting slabs.

A straightforward conclusion which may be drawn from present work is that there is a seismic discontinuity with an S-wave velocity jump of a few per cent at a depth around 920 km beneath subduction zones. If such a discontinuity cannot be observed beneath different tectonic regions other than subduction zones, the most likely explanation for the discontinuity is that it represents the bottom of the subducted slab's garnet layer[27]. Considering that the depths are almost constant beneath three subduction zones, it is unlikely that the layer is dynamically supported as depicted by Ringwood and Irifune[27]; instead, it should be fairly flat and the depth is thermodynamically controlled. If a stability condition of materials that constitute such a layer is experimentally determined, we should be able to constrain the temperature of the mantle at a depth of ~920 km, which we have no good way to estimate otherwise.

On the other hand, we must also pay attention to the fact that suggestion of either first-order or second-order discontinuities at these depths is not completely new. Earlier this century, Gutenberg noted a second-order discontinuity (a discontinuous velocity gradient) at a depth of 950 km, which defines the bottom of the transition zone[28]. As mentioned earlier, Bullen also viewed this depth as the base of the transition zone[1]. Changes of slowness of P waves bottoming at these depths have been reported by many researchers[11,29-31]. Reflected phases from this general depth range have also been suggested previously beneath different tectonic regions: subduction zones[13-15], oceanic ridges[32] and continents[9], although the evidence for the presence of a discontinuity is much weaker than that presented by this study. Recently, analysing precursors to the SS phase, Petersen et al.[13] showed good evidence for a reflector at a depth of 900 km beneath the Kuril–Kamchatka subduction zone. Although the authors do not mention it, the recent work by Gurrola et al. (in their Fig. 14a)[33] also suggests a presence of such discontinuity beneath western Russia.

While the question is by no means settled, we thus favour the possibility that the 920-km discontinuity may be a global feature, and that it may represent the bottom of the transition zone, below which the lower mantle should be defined. The geophysical consensus on the depth of the top of the lower mantle has shallowed considerably from ~950 km to ~660 km in the past 30 years[6]. This was probably so because the 660-km discontinuity has attracted so much attention on the part of geophysical community. The confirmation of the existence of the 920-km discontinuity may deepen the depth of the top of the lower mantle back to its original placement by Bullen and Gutenberg. We may also need to introduce primed symbols for Bullen's region C, as Bullen himself noted[1].

The consequence of the presence of a 920-km discontinuity, if its global nature is ultimately confirmed, would be of great importance for many problems in geodynamics. With currently accepted mineralogical models for the upper mantle, if we assume that upper and lower mantles are chemically homogeneous, no known phase transformation exists below a depth of 720 km (ref. 34). It has been suggested that the change in crystal symmetry of perovskite might account for the existence of a discontinuity at about ~900 km (refs 15, 35, 36), but the recent experimental and theoretical studies seem to indicate that orthorhombic perovskite is stable in most of the lower mantle[37,38]. Funamori and Yagi[37] showed that orthorhombic perovskite is stable up to 1,900 K at 36 GPa. This condition is close to that expected for a depth of 920 km with a 'cold geotherm' which can be obtained assuming the whole mantle is convecting as a unit[39]. It may still be possible that orthorhombic perovskite transforms into a higher symmetry phase at a higher temperature with a 'hot geotherm' to account for the presence of the 920-km discontinuity. Even in this case, a presence of a thermal boundary layer at this depth range would be required, leading to the idea of a two-layered convecting mantle[39]. If the possibility of the change in crystal symmetry of perovskite is completely ruled out from experimental work, we may need to invoke the idea that the lower mantle is chemically distinct from the upper mantle (that is, the 920-km discontinuity is a chemical boundary, while the 660-km discontinuity is a phase boundary), even though some slab may penetrate through the boundary[40,41]. So in any case, the implication of the presence of a global 920-km discontinuity would be enormous for our understanding of the mantle dynamics, and further detailed seismological and mineralogical studies of the lower part of the mantle transition zone below the 660-km discontinuity may help to answer the long-standing unsolved problem of Earth sciences: does the Earth's mantle convect in one layer or two? □

1. Bullen, K. E. An Introduction to the Theory of Seismology 3rd edn (Cambridge Univ. Press, 1963).
2. Dziewonski, A. M. & Anderson, D. L. Phys. Earth planet. Inter. 25, 297–356 (1981).
3. Kennett, B. L. N. & Engdahl, E. R. Geophys. J. Int. 105, 429–465 (1991).
4. Ringwood, A. E. Composition and Petrology of the Earth's Mantle (McGraw-Hill, New York, 1975).
5. Ito, E. & Takahashi, E. J. geophys. Res. 94, 10637–10646 (1989).
6. Anderson, D. L. Theory of the Earth (Blackwell Scientific, Boston, 1989).
7. Gutenberg, B. Trans. Am. geophys. Un. 39, 86–489 (1958).
8. Repetti, W. C. thesis, St. Louis Univ. (1930).
9. Hoffman, J. P., Berg, J. W. & Cook, K. L. Bull seism. Soc. Am. 51, 17–27 (1961).
10. Vvedenskaya, A. V. & Balakina, L. M. Izv. Akad. Nauk SSSR Geophys. Ser. 7, 1138–1146 (1959).
11. Kanamori, H. Bull. Earth. Res. Inst. Tokyo Univ. 45, 657–678 (1967).
12. Johnson, L. R. Bull seism. Soc. Am. 59, 973–1008 (1969).
13. Petersen, N., Gossler, J., Kind, R., Stammler, K. & Vinnik, L. Geophys. Res. Lett. 20, 281–284 (1993).
14. Revenaugh, J. & Jordan, T. H. J. geophys. Res. 94, 5787–5813 (1989).
15. Revenaugh, J. & Jordan, T. H. J. geophys. Res. 96, 19763–19780 (1991).
16. Wicks, C. W. Jr & Richards, M. A. EOS 74, 550 (1993).
17. Shearer, P. M. & Masters, T. G. Nature 355, 791–796 (1992).
18. Vidale, J. E. & Benz, H. M. Nature 356, 678–682 (1992).
19. Benz, H. M. & Vidale, J. E. Nature 365, 147–150 (1993).
20. Wicks, C. W. Jr & Richards, M. A. Science 261, 1424–1427 (1993).
21. Shearer, P. M. Nature 344, 121–126 (1990).
22. J-Array Group J. Geomag. Geoelect. 45, 1265–1274 (1993).

23. Kanasewich, E. R. Time Sequence Analysis in Geophysics (Univ. Alberta Press, Edmonton, 1973).
24. Niu, F.-L. thesis Univ. Tokyo (1994).
25. Kennett, B. L. N. in Seismological Algorithms (ed. Doornbos, D. J.) 237–259 (Academic, San Diego, 1988).
26. Dziewonski, A. M., Chou, T.-A. & Woodhouse, J. H. J. geophys. Res. 86, 2825–2853 (1981).
27. Ringwood, A. E. & Irifune, T. Nature 331, 131–136, (1988).
28. Gutenberg, B. Physics of the Earth's Interior (Academic, New York, 1959).
29. Corbishley, D. J. Geophys. J. R. astr. Soc. 19, 1–14 (1970).
30. Vinnik, L. P., Lukk, A. A. & Nikoleav, A. V. Phys. Earth planet. Inter. 5, 328–331 (1972).
31. Datt, R. & Muirhead, K. J. Phys. Earth planet. Inter. 15, 28–38 (1977).
32. Whitcomb, J. H. & Anderson, D. L. J. geophys. Res. 75, 5713–5728 (1970).
33. Gurrola, H., Minster, J. B. & Owens, T. Geophys. J. Int. 117, 427–440 (1994).
34. Irifune, T. Nature 370, 131–132 (1994).
35. Wang, Y., Guyot, F. & Liebermann, R. C. J. geophys. Res. 97, 12327–12347 (1992).
36. Wolf, G. H. & Bukowinski, M. S. T. Geophys. Res. Lett. 12, 809–812 (1985).
37. Funamori, N. & Yagi, T. Geophys. Res. Lett. 20, 387–390 (1993).
38. Stixrude, L. & Cohen, R. E. Nature 613, 613–616 (1993).
39. Jeanloz, R. & Morris, S. A. Rev. Earth planet. Sci. 14, 377–415 (1986).
40. Grand, S. P. J. geophys. Res. 99, 11591–11621 (1994).
41. Silver, P. G., Carlson, R. W. & Olson, P. A. Rev. Earth planet. Sci. 16, 477–541 (1988).