ARTICLE

Evidence for a difference in rupture initiation between small and large earthquakes

S. Colombelli[1], A. Zollo[1], G. Festa[1] & M. Picozzi[1]

The process of earthquake rupture nucleation and propagation has been investigated through laboratory experiments and theoretical modelling, but a limited number of observations exist at the scale of earthquake fault zones. Distinct models have been proposed, and whether the magnitude can be predicted while the rupture is ongoing represents an unsolved question. Here we show that the evolution of P-wave peak displacement with time is informative regarding the early stage of the rupture process and can be used as a proxy for the final size of the rupture. For the analysed earthquake set, we found a rapid initial increase of the peak displacement for small events and a slower growth for large earthquakes. Our results indicate that earthquakes occurring in a region with a large critical slip distance have a greater likelihood of growing into a large rupture than those originating in a region with a smaller slip-weakening distance.

[1] Department of Physics, University of Naples Federico II, Complesso Universitario Monte S. Angelo, via Cinthia, 80126 Naples, Italy. Correspondence and requests for materials should be addressed to A.Z. (email: aldo.zollo@unina.it).

1

The slow, stable nucleation phase and its transition to a dynamic earthquake rupture have been thoroughly investigated by theoretical, numerical and laboratory studies to understand the preparatory phase of earthquakes and their potential to grow into a large event[1–6]. Being able to predict the final size of an earthquake while the rupture is ongoing on the fault plane is a highly debated scientific problem. Contrasting views have been proposed on whether the initial stage of the rupture process is different for small and large earthquakes[7–11]. According to the "cascade model", the rupture starts in the same way for small and large earthquakes. The local friction conditions and the geometrical discontinuities of the fault surface control the final size[12]. The immediate implication of this model is that a seismic event is a stochastic process for which the final size cannot be known until the rupture has finished and is even more indeterminate when only the beginning of the process is considered. However, recent theoretical studies, seismic observations and laboratory experiments have indicated a possible scaling between the characteristics of the initiation phase and the earthquake magnitude, suggesting a potential deterministic nature of seismic rupture[13–16]. Ishihara et al.[7] analysed the moment rate function (MRF) of a few Japanese earthquakes and provided evidence for their different rising slopes, suggesting a causal relation between the MRF and the final event magnitude. Using a broader dataset, Ellsworth and Beroza[9] found that the size and duration of the initiation phase scale with the final magnitude. Several authors, in contrast, did not observe a change in the P-wave onset or they provided an alternative interpretation of the differences in the nucleation phase due to the filtering procedure or propagation effects[8,10–11].

The interest in rupture initiation was raised again recently because of its immediate impact on the rapid assessment of earthquake size. The ability to correctly distinguish a small shock from a large event through the analysis of the first P-wave observation is crucial for risk mitigation actions triggered by earthquake early warning systems. Several authors demonstrated that the early portion of recorded P-waves contains information regarding the magnitude through its frequency content and/or amplitude[17–20]. Most of these studies concentrated on the analysis of a fixed P-wave time window (PTW; 3–4 s).

In the present study, we measure the peak displacement amplitude of filtered P-wave signals (P_d) over a progressively expanding PTW. Previous studies of the beginning of waveforms analysed either a few earthquakes recorded at a limited number of stations or earthquakes belonging to the same sequence. Under such conditions, recognizing and clearly discriminating between source and propagation effects, both of which can influence the initial shape of signals, can be difficult. Here, we use a larger, high-quality dataset of 43 moderate-to-strong Japanese events spanning wide magnitude (M) and distance (R) ranges ($4 \leq M \leq 9$; $0 \leq R \leq 500$ km). We analyse more than 7000 three-component waveforms recorded at 1,208 stations. The full list of analysed events is provided in Supplementary Table 1, while further details regarding the database selection are provided in the Supplementary Methods section. Figure 1 shows the map of stations used and the epicentres of the selected events.

Results

Time evolution of peak displacement. The number of analysed earthquakes ensures adequate data sampling in both magnitude and distance bins (Fig. 2a–c). For each event, the logarithm of P_d versus PTW curve (hereafter abbreviated as LPW curve) is obtained by averaging the data from all available stations at any time (typically more than 100; see the Methods section). We found that each curve progressively increases with time until a final, stable value is reached. The plateau level (PL) and saturation time generally increase with magnitude (Fig. 3a).

For a single station, the far-field source kinematic models indicate that the P_d versus time curve reproduces the shape of the
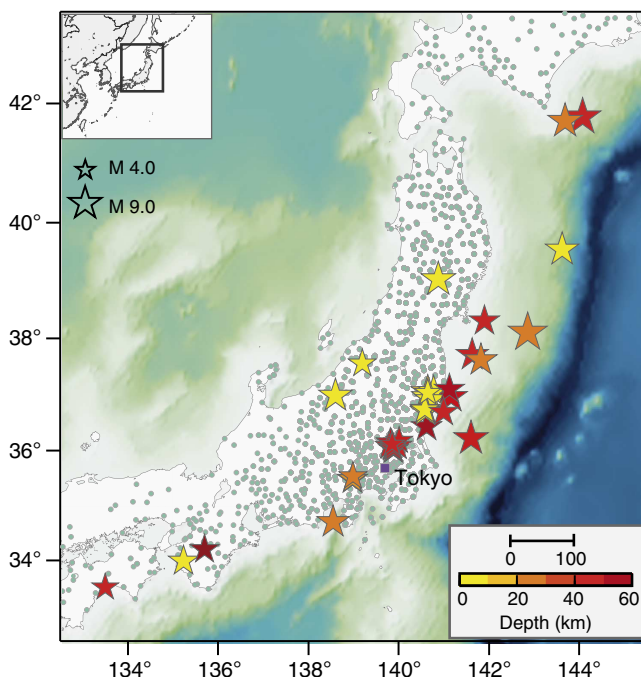

Figure 1 | Event map. The distribution of stations used in this study (small green circles) and the epicentral locations of the 43 selected events (coloured stars). The size of the star is proportional to the magnitude, and the colour represents the source depth. The black bar at the bottom right denotes a 100 km length.
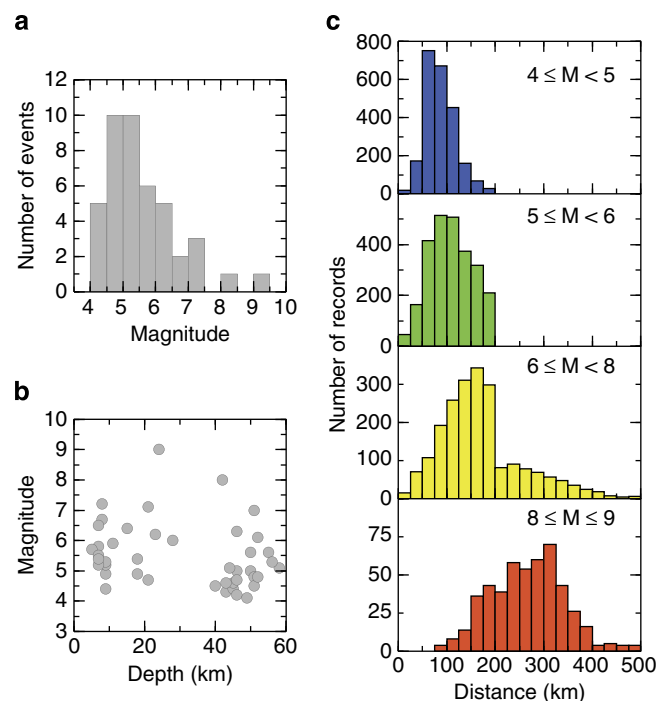

Figure 2 | Data histograms. (**a**) Histogram of the number of earthquakes as a function of magnitude. (**b**) Distribution of magnitude as a function of depth for the selected events. (**c**) Distribution of records as a function of distance for different magnitude classes.
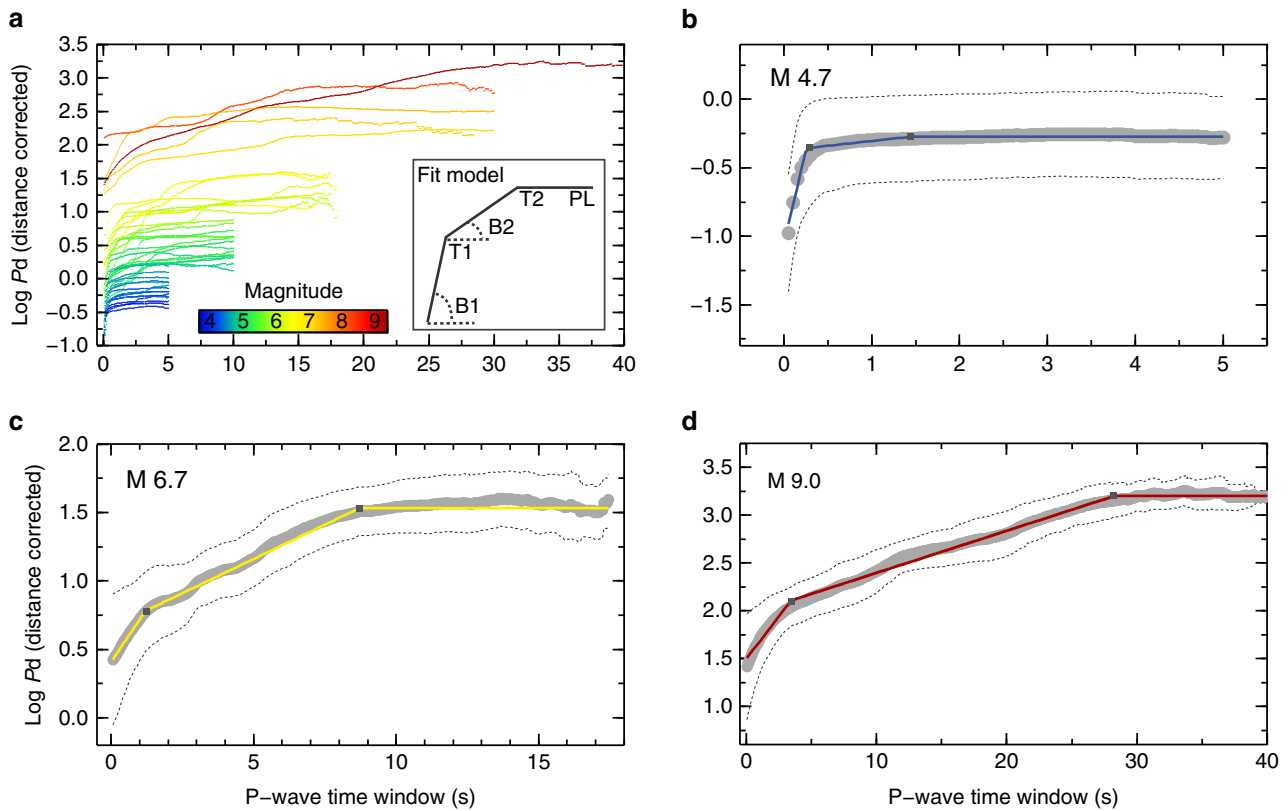
Figure 3 | LPW curves and model fit. (**a**) Average logarithm of the P_d versus PTW (LPW) curve for each analysed event. The y-axis represents the average logarithm of P_d, obtained by using all the available data at each time point, after correcting individual values for the distance effect (using equation 1). The insert box is a schematic representation of the piecewise linear fitting function. (**b,c,d**), Example of LPW curve for three representative events of magnitude M 4.7 (**b**), 6.7 (**c**) and 9 (**d**), respectively. In each panel, grey circles are the observed data with the associated ±1σ error bars (dashed lines). The solid line is the best fit function, obtained using the piecewise linear model and the dark squares mark T1 and T2 times of each event.

apparent source time function[21]. By averaging P_d among many stations, distributed over azimuth and distance, the resulting function approximates the MRF. Assuming a triangular shape for the MRF, the plateau of P_d versus time curves is reached at the peak of the MRF, and the corresponding time is a proxy for the source moment and duration[22]. Our observations indicate that the initial shape of the P_d versus time curves typically follows an exponential trend.

Model fit for parameter estimation. To reproduce the observed shape, we modelled the LPW curve using a piecewise linear function (Fig. 3a, small box). For each event, we fit the data and determined the corner time of the first and second straight line (T1 and T2, respectively), the slope of the two lines (B1 and B2, respectively) and the final PL. Examples of fitted curves for three representative events are shown in Fig. 3b–d. As expected, T2 and PL scale with the magnitude (Supplementary Fig. 1), which is an *a posteriori* validation of the hypothesis that averaged LPW curves approximate the MRF. Our main focus here is on T1 and B1 for their possible relation with the beginning of the rupture process. We found a significant log-linear correlation with magnitude for both T1 and B1. In particular, T1 increases with magnitude, whereas B1 is inversely proportional to the earthquake size (Fig. 4a,b). These results indicate that the evolution of the LPW curve at the early stage of the rupture is to some degree related to the final earthquake size. Figure 4c shows the observed LPW curve up to its T1 time for a few representative events and the expected slopes given the B1 versus magnitude trend (small box). The individual estimated parameters are remarkably

accurate in terms of errors, and their scaling with magnitude is robust (linear correlation coefficient $R^2 = 0.8$), although a clear variability appears in the plot of the whole dataset.

Tests and analyses of robustness. Different factors might have influenced the shape of the LPW curves. The earthquake depth, for example, may impact source processes and frictional properties, thus affecting our observation regarding the evolution of P-wave peak displacement with time. Among the other factors, the high-pass filtering operation and the propagation effects related to crustal structure, geometrical spreading and anelastic attenuation are the most relevant. We performed different tests and separately investigated the influence of these factors on the analysis (see Supplementary Discussion and Supplementary Figs 2–6 for details). We did not find any clear evidence that either the source depth or the other effects are responsible for the observed trends of T1 and B1 with magnitude. Finally, we tested the persistence of the B1 trend with magnitude by limiting all the LPW curves to a fixed time window. We selected 1 s for $M \geq 6$ and 0.1 s for $M < 6$, and for each event, we estimated the initial slope (by fitting data with a single straight line). Supplementary Fig. 7 shows that the slope estimates obtained within fixed time windows are consistent with those obtained by fitting the whole curves and that the trend of B1 with magnitude is still observed when only the very beginning of the curve is considered.

Discussion

Abrupt variations in the slip acceleration at the early stage of the rupture are associated with the unstable behaviour of the
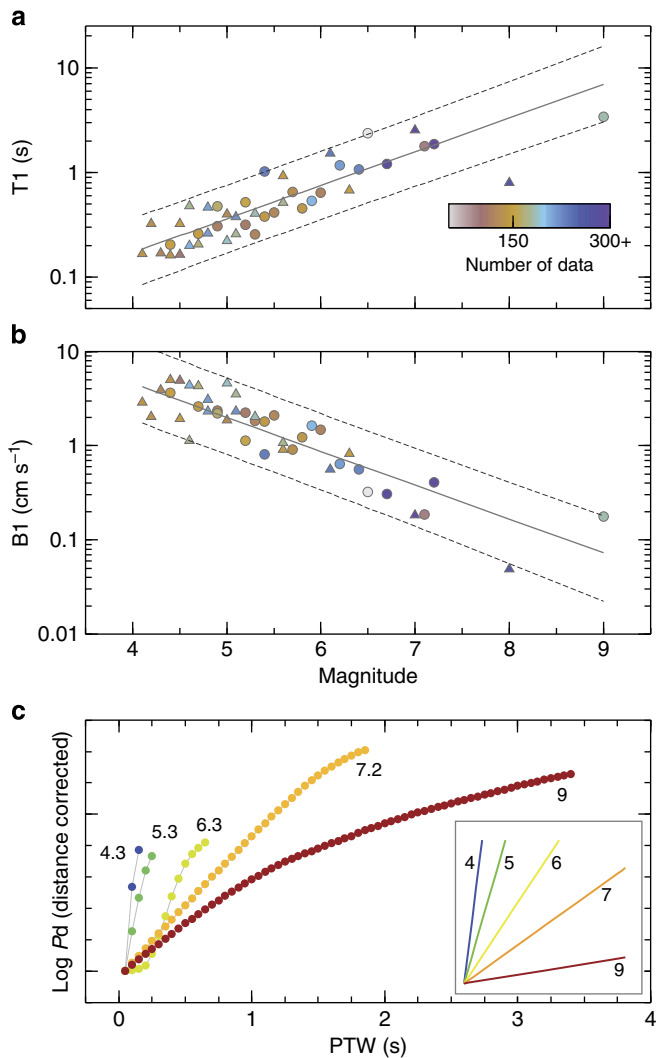
a

b

c

Figure 4 | Fit parameters T1 and B1. (**a**) First-corner time of the linear piecewise fit model as a function of magnitude for all the analysed events on a log-linear scale. The colour scale indicates the number of data points used for the average computation at time T1. (**b**) Slope of the first-line segment of the fit model as a function of magnitude for all the analysed events on a log-linear scale. For each individual B1 value, the colour is the same as that used in **a** for the corresponding T1 data point. In **a** and **b**, small circles represent the crustal events (source depth < 30 km), and triangles represent the subcrustal events (source depth ≥ 30 km). The solid grey line shows the linear best fit function, and the dashed lines are the 95% confidence intervals for new observations given the existing fit to the data. Error bars for individual data points are smaller than the symbol size. (**c**) Magnification of the LPW curve until time T1 for representative events. A common initial value has been assigned to compare the shape of the curves. The insert box shows the expected initial slope of the LPW curve for different magnitudes given the observed trend of B1 with magnitude.

nucleation phase, before the onset of dynamic rupture propagation[23]. The characteristic time for the nucleation of large earthquakes is from one to a few seconds[15], which is comparable to our estimates of the T1 time from the LPW curves (Fig. 4a). In the framework of linear slip-weakening models, analytical solutions of the nucleation phase indicate that the exponential growth is mainly controlled by the weakening rate[24,25], which is the ratio between the strength drop of the fault related to the friction and the characteristic slip-weakening distance (D_c) over which the energy is dissipated. During

nucleation, the initial stress is close to the threshold in the slipping zone. We can therefore replace the strength drop with dynamic stress drop in our interpretation. Observations indicate a decrease in B1 of approximately one order of magnitude in the range M 4.5–7 (Fig. 4b). In the same range, this decrease would correspond either to a decrease in the initial stress or to an increase in the critical slip distance, both of one order of magnitude. Far-field measurements indicate an unclear scaling, or eventually a weak increase, in the stress drop with magnitude in this range[26].

Although these estimates refer to the whole rupture process, an initial decrease is unlikely to result in a final constant value, independent of magnitude. In contrast, assuming for D_c a characteristic value of 0.5–1 m for an $M = 7$ event[27], the corresponding value for an $M = 4.5$ event would be at the scale of 5–10 cm. Extrapolating this value to negative magnitudes, micrometric-to-millimetric values for D_c would be obtained, consistent with the estimates from laboratory experiments. Such a scaling indicates the availability of a larger fracture energy in the initial stage of the rupture for larger magnitude events. In addition, the nucleation size is directly linked to the weakening rate for linear slip-weakening models[28], which suggests a larger size and a longer characteristic time of the nucleation of larger events, as confirmed by the positive scaling of T1 with magnitude.

Increasing in the breakdown energy with magnitude is commonly interpreted in earthquake fracture mechanics as a homogenized effect of all competing dissipative mechanisms on the fault plane, such as friction, off-fault damage, thermal pressurization and melting[29]. Extending this interpretation to the initial stage of the rupture, possible changes in D_c along the fault can be related to a characteristic length of the asperities[3] or to an intrinsic stiffness of the fault zone, which elastically accumulates energy during tectonic fault loading. Assuming a variable, fractal distribution of the D_c on the fault, with some space coherency, we argue that when an earthquake rupture develops over a region with a low D_c value, its probability of rupturing adjacent regions is low. In contrast, when a coherent rupture nucleates in a region of larger D_c, the probability that it will grow into a larger size event increases. In this context, the variation of D_c and of the probability that an initial earthquake rupture will grow justifies the dispersion of B1 values with respect to magnitude.

Finally, the log-linear relationship between the MRF and B1 indicates the exponential growth of the nucleation phase. Although the observed T1 values are comparable to the nucleation times, we cannot exclude the possibility that they involve the initial acceleration of the dynamic rupture phase. In addition, larger breakdown energy prevents a rapid acceleration of the rupture at this initial stage, which also limits the moment rate increase with time.

The observed variation in B1 with magnitude may have an important impact on the rapid assessment of the earthquake size. Nevertheless, the slope measured in this work is derived from the average trend of hundreds of records, spanning wide azimuth and distance ranges, whereas in real-time, the availability of data depends on the source-station geometry. Further analyses are therefore required to simulate the real-time data streaming and to assess the minimum near-source data necessary to constrain the initial slope. Assuming that the slope is well-constrained from near-source data, a short time window (approximately 1 s, Supplementary Fig. 7) is sufficient for an evaluation of the earthquake size.

Methods
Data processing and average peak displacement computation. To obtain displacement waveforms, the mean value and linear trend are first removed

from the raw acceleration data. These records are then double integrated, and a non-causal, high-pass Butterworth filter with a cutoff frequency of 0.075 Hz is applied to remove the artificial long-period trend and baselines introduced by the double-integration operation. The P-wave onset times were manually selected from each vertical component of the accelerometer records.

We measured the peak displacement amplitude on the filtered P-wave signals (P_d) over a progressively expanding PTW, starting from 0.05 s after the P-wave onset time and continuing until the expected arrival of the S-phase[30]. For each event, we obtained the LPW curve by averaging all the available data at each time window after correcting the observed P_d values at different stations for the geometrical attenuation effect. To compare the observed P_d values at different stations, the geometrical attenuation effect was corrected with the following attenuation relationship:

$$\log P_d = A + B M + C \log R \qquad (1)$$

where M is the event magnitude, P_d is measured in cm (in a fixed time window of 3 s) and the hypocentral distance R is measured in km. In equation (1), a single term ($\log R$) is used to account for both the geometrical spreading and the anelastic attenuation. The magnitudes used here are provided by Japan Meteorological Agency and are assumed to be equivalent to moment magnitudes. For earthquakes with $M < 7$, we used only data up to a maximum hypocentral distance of 200 km, whereas for $M \geq 7$ events, we included records up to 500 km. We preliminarily estimated the coefficients A, B and C of equation (1) for each distance range through a non-weighted linear regression analysis. Then, depending on the magnitude and the distance range, we used the appropriate coefficients to compare the observed P_d values at different stations. Supplementary Table 2 shows the coefficients A, B and C for each magnitude and distance range.

To avoid the inclusion of S-waves in the considered PTW, we computed the expected S-wave arrival time at each station (T_s) using the following relation:

$$T_s - T_p = b R \qquad (2)$$

where T_p is the observed P-wave onset time, R is the hypocentral distance in km and $b = 0.088$. The coefficient b was derived through a linear regression analysis after manually selecting the S-wave arrival times for a limited number of records from our dataset.

For each event, we set a maximum PTW to be explored, based on the expected duration of the source from kinematic models for the corresponding magnitude, and considered at least twice the expected duration. To compute the average LPW curve, we also required at least five data points to be used in each PTW. The average computation stops either when the number of data points is not sufficient or when the PTW reaches its maximum time limit, whichever condition occurs first.

To model the LPW curves, we used a piecewise, linear fit function inverted for five unknown parameters: first- and second-corner times (T1 and T2, respectively), the slope of the two straight lines (B1 and B2, respectively) and the final PL.

References

1. Ohnaka, M. & Yamashita, T. A cohesive zone model for dynamic shear faulting based on experimentally inferred constitutive relation and strong motion source parameters. *J. Geophys. Res.* **94,** 4089–4104 (1989).
2. Ohnaka, M. Nonuniformity of the constitutive law parameters for shear rupture and quasistatic nucleation to dynamic rupture: a physical model of earthquake generation processes. *Proc. Natl Acad. Sci. USA* **93,** 3795–3802 (1996).
3. Ohnaka, M. A physical scaling relation between the size of an earthquake and its nucleation zone size. *Pure Appl. Geophys.* **15,** 2259–2282 (2000).
4. Rubin, A. & Ampuero, J.-P. Self-similar slip pulses during rate-and-state earthquake nucleation. *J. Geophys. Res.* **114,** B11305 (2009).
5. Nielsen, S., Taddeucci, J. & Vinciguerra, S. Experimental observation of stick-slip instability fronts. *Geophys. J. Int.* **180,** 697–702 (2010).
6. Kaneko, Y. & Ampuero, J.-P. A mechanism for preseismic steady rupture fronts observed in laboratory experiments. *Geophys. Res. Lett.* **38,** L21307 (2011).
7. Ishihara, Y., Fukao, Y., Yamada, I. & Aoki, H. Rising slope of moment rate functions: the 1989 earthquake off East Coast of Honshu. *Geophys. Res. Lett.* **19,** 873–876 (1992).
8. Abercrombie, R. & Mori, J. Local observations of the onset of a large earthquake: 28 June 1992 Landers, California. *Bull. Seismol. Soc. Am.* **84,** 725–734 (1994).
9. Ellsworth, W. L. & Beroza, G. C. Seismic evidence for an earthquake nucleation phase. *Science* **268,** 851–855 (1995).
10. Mori, J. & Kanamori, H. Initial rupture of earthquakes in the 1995 Ridgecrest, California sequence. *Geophys. Res. Lett.* **23,** 2437–2440 (1996).
11. Scherbaum, F. & Bouin, M.-P. FIR filter effects and nucleation phases. *Geophys. J. Int.* **130,** 661–668 (1997).
12. Wesnousky, S. G. Predicting the endpoints of earthquake ruptures. *Nature* **444,** 358–360 (2006).
13. Rice, J. R. In *Physics of the Earth's Interior* (eds Dziewonski, A. M. & Boschi, E.) 555–649 (Italian Physical Society, 1980).
14. Iio, Y. Observations of the slow initial phase generated by microearthquakes: implications for earthquake nucleation and propagation. *J. Geophys. Res.* **100,** 15333–15349 (1995).
15. Beroza, G. C. & Ellsworth, W. L. Properties of the seismic nucleation phase. *Tectonophysics* **261,** 209–227 (1996).
16. Olson, E. & Allen, R. M. The deterministic nature of earthquake rupture. *Nature* **438,** 212–215 (2005).
17. Allen, R. M. & Kanamori, H. The potential for earthquake early warning in Southern California. *Science* **3,** 685–848 (2003).
18. Zollo, A., Lancieri, M. & Nielsen, S. Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion. *Geophys. Res. Lett.* **33,** L23312 (2006).
19. Zollo, A., Lancieri, M. & Nielsen, S. Reply to comment by P. Rydelek *et al.*, on "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records". *Geophys. Res. Lett.* **34,** L20303 (2007).
20. Wu, Y.-M. & Zhao, L. Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophys. Res. Lett.* **33,** L16312 (2006).
21. Sato, T. & Hirasawa, T. Body wave spectra from propagating shear cracks. *J. Phys. Earth* **21,** 415–431 (1973).
22. Duputel, Z., Tsai, V. C., Rivera, L. & Kanamori, H. Using centroid time-delays to characterize source durations and identify earthquakes with unique characteristics. *Earth Planet. Sci. Lett.* **374,** 92–100 (2013).
23. Yamashita, T. & Ohnaka, M. Nucleation process of unstable rupture in the brittle regime: a theoretical approach based on experimentally inferred relations. *J. Geophys. Res.* **96,** 8351–8367 (1991).
24. Dascalu, C., Ionescu, I. R. & Campillo, M. Fault finiteness and initiation of dynamic shear instability. *Earth Planet. Sci. Lett.* **177,** 163–176 (2000).
25. Ampuero, J. P., Vilotte, J. P. & Sánchez-Sesma, F. J. Nucleation of rupture under slip dependent friction law: simple models of fault zone. *J. Geophys. Res.* **107,** 2324 (2002).
26. Oth, A., Bindi, D., Parolai, S. & Di Giacomo, D. Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: insights from KiK-net data in Japan. *Geophys. Res. Lett.* **37,** L19304 (2010).
27. Tinti, E., Bizzarri, A. & Cocco, M. Modeling the dynamic rupture propagation on heterogeneous faults with rate- and state-dependent friction. *Ann. Geophys.* **48,** 327–345 (2005).
28. Uenishi, K. & Rice, J. R. Universal nucleation length for slip-weakening rupture instability under nonuniform fault loading. *J. Geophys. Res.* **108,** 2042 (2003).
29. Lapusta, N. & Rice, J. R. Nucleation and early seismic propagation of small and large events in a crustal earthquake model. *J. Geophys. Res.* **108,** 2205 (2003).
30. Colombelli, S., Zollo, A., Festa, G. & Kanamori, H. Early magnitude and potential damage zone estimates for the great Mw 9 Tohoku-Oki earthquake. *Geophys. Res. Lett.* **39,** L22306 (2012).

Acknowledgements

Author contributions

S.C. implemented and applied the methodology and wrote the related text. G.F. and A.Z. contributed to the methodological development and interpretation of the results. M.P. contributed to the discussion of methodological aspects. All authors contributed to the manuscript redaction and final revision.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interest.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Colombelli, S. *et al.* Evidence for a difference in rupture initiation between small and large earthquakes. *Nat. Commun.* 5:3958 doi: 10.1038/ncomms4958 (2014).
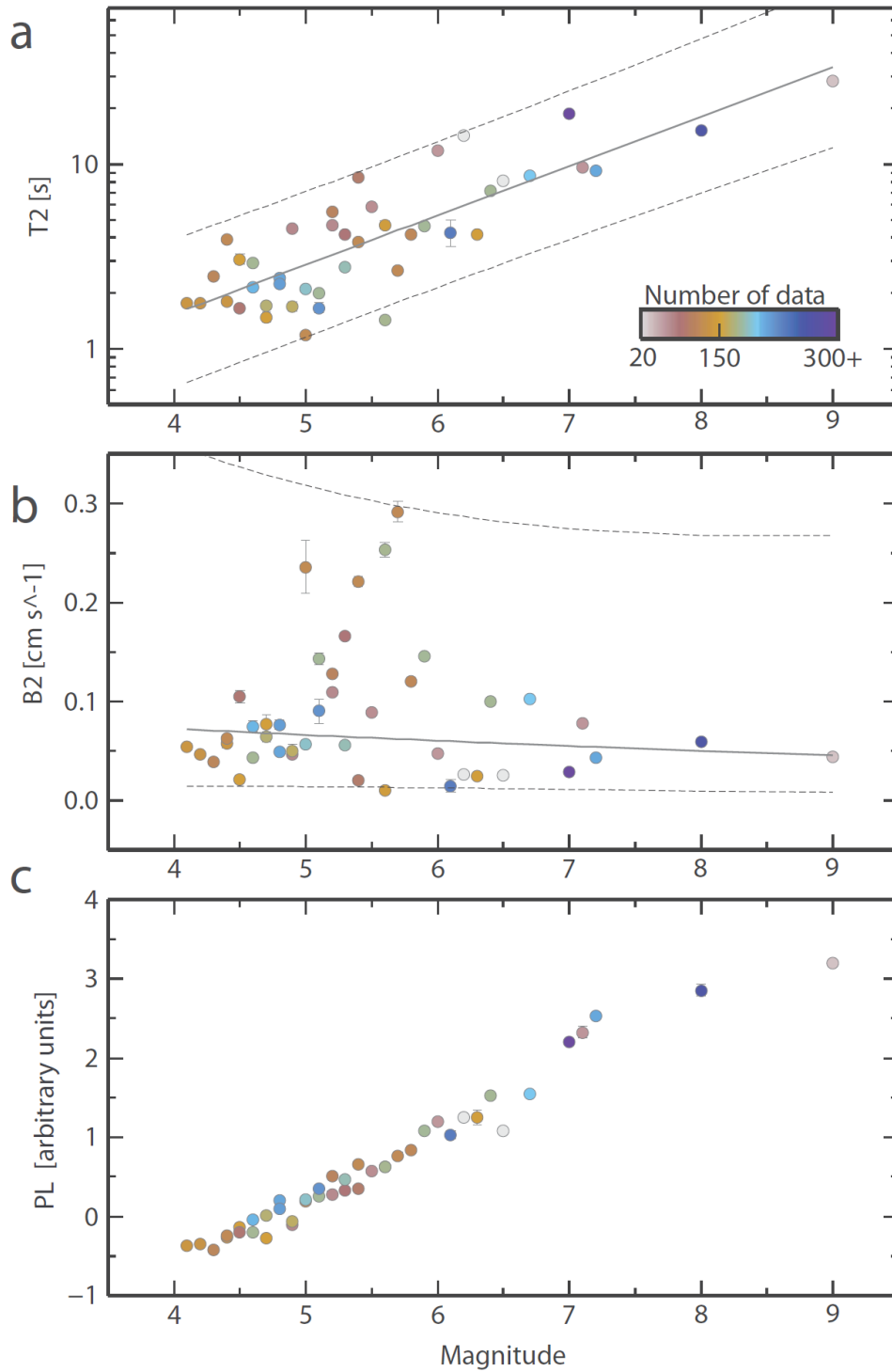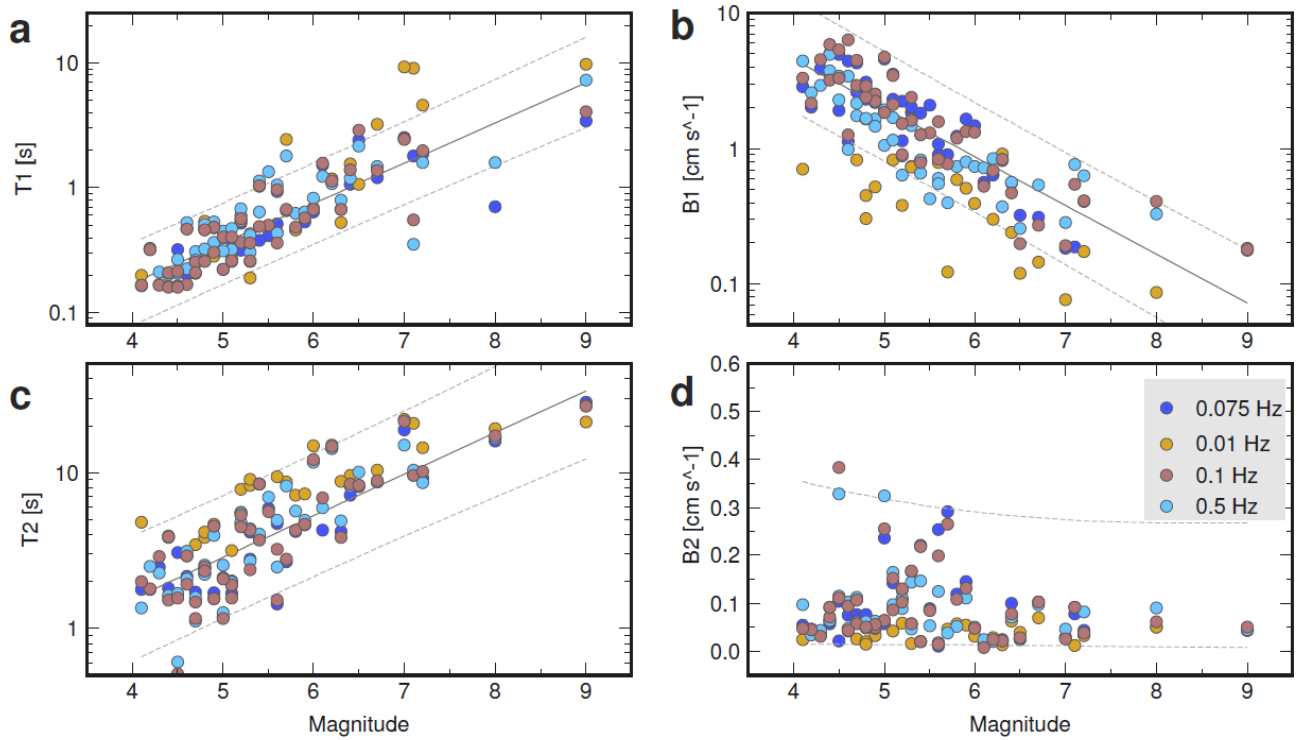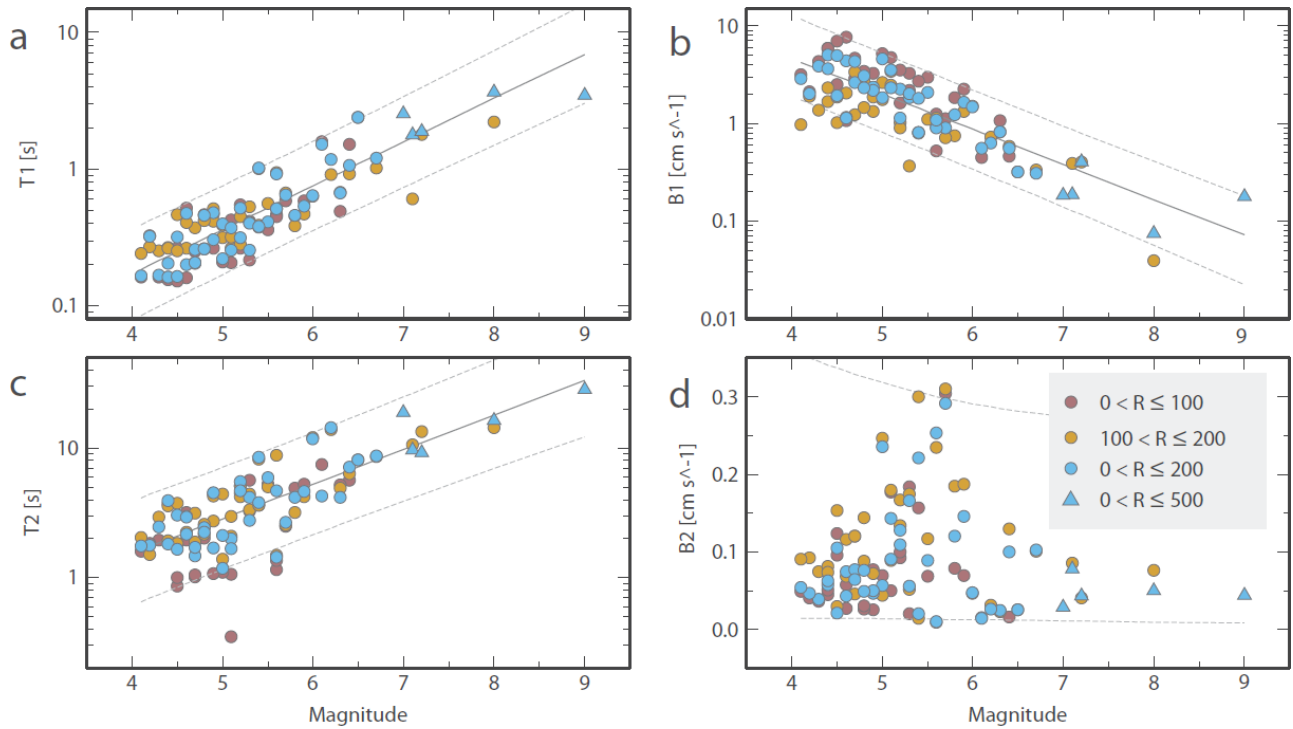
Supplementary Figure 1. Fit parameters T2, B2 and PL. a, Second corner time of the linear piecewise fit model (T2) as a function of magnitude for all the analysed events, on a log-linear scale. The colour scale indicates the number of data used for the average computation at the time T2. **b,** Slope of the second line piece of the fit model (B2) as a function of magnitude for all the analysed events, on a linear scale. For each individual B2 value, the colour is the same as used in panel a for the corresponding T2 data point. In panels a and b, the solid grey line shows the linear best fit function and the thin dashed lines are the 95% confidence interval for new observations, given the existing fit to the data. Error bars for individual data points are also shown in the plot (only the largest errors are visible). **c,** Saturation level (PL) of the LPW curve for each event. The saturation level is measured at the time T2.
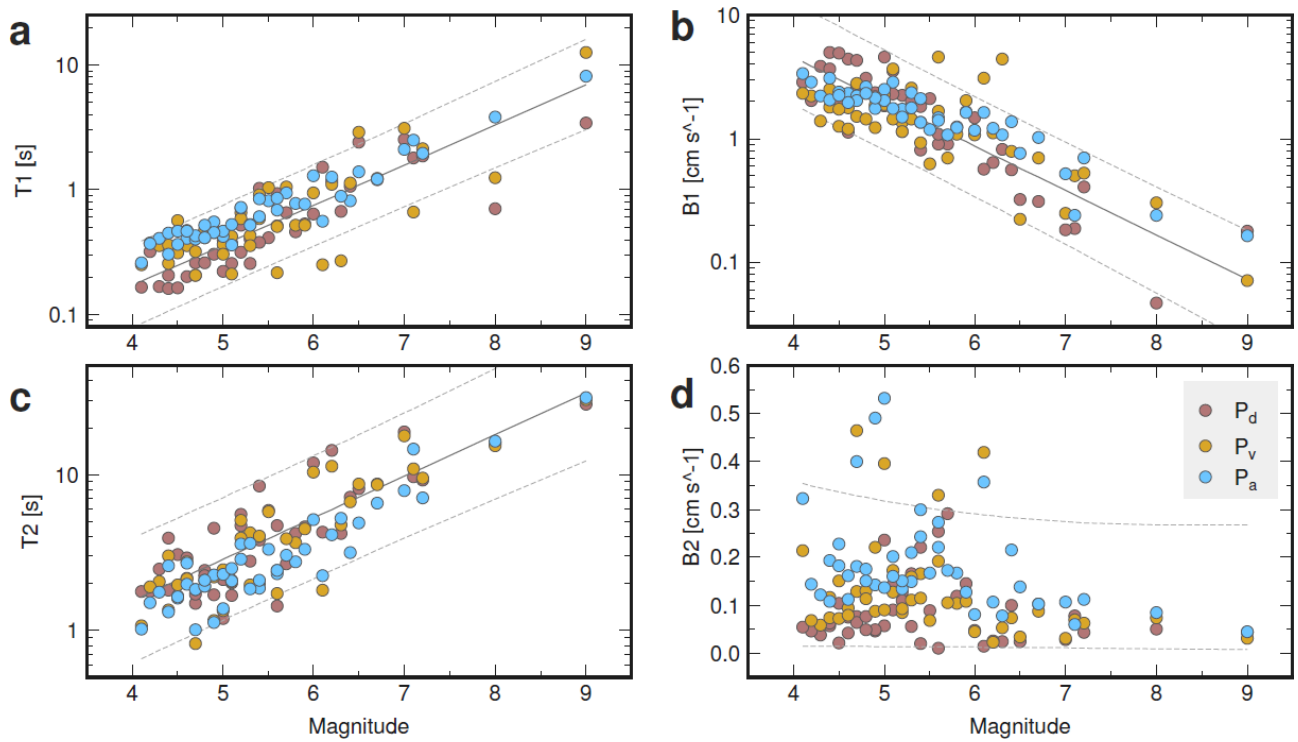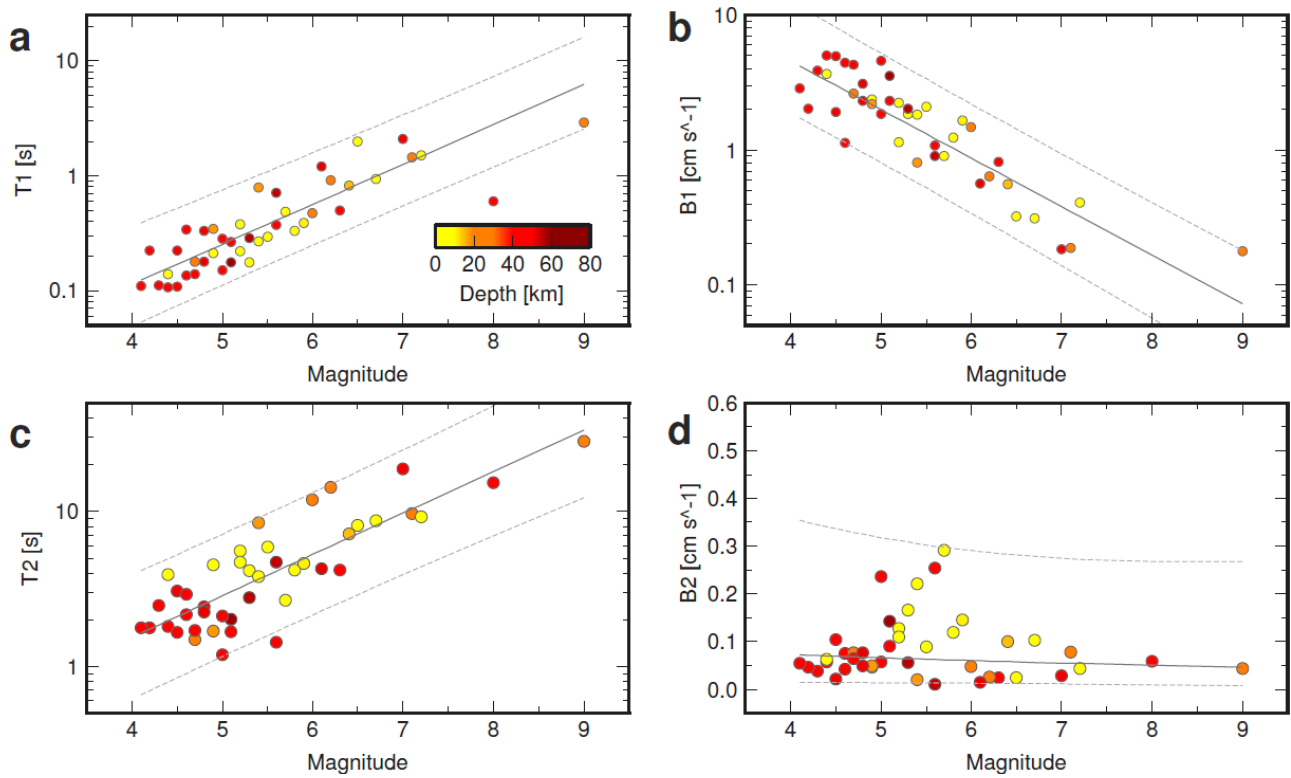
Supplementary Figure 2. Filter effect on fit parameters T1, B1, T2, B2. The figure shows the four parameters resulting from the fitting procedure as a function of magnitude using filtered displacement waveforms (brown circles), unfiltered velocity waveforms (yellow circles) and unfiltered acceleration records (light blue circles). **a,** First corner time of the linear piecewise fit model (T1) as a function of magnitude for all the analysed events, on a log-linear scale. **b,** Slope of the first line piece (B1) as a function of magnitude for all the analysed events, on a log-linear scale. **c,** Second corner time of the linear piecewise fit model (T2) as a function of magnitude for all the analysed events, on a log-linear scale. **d,** Slope of the second line piece (B2) as a function of magnitude for all the analysed events, on a linear scale. In all panels the solid grey line shows the linear best fit function when the P_d data are used and the thin dashed lines are the corresponding 95% confidence bounds for new observations.
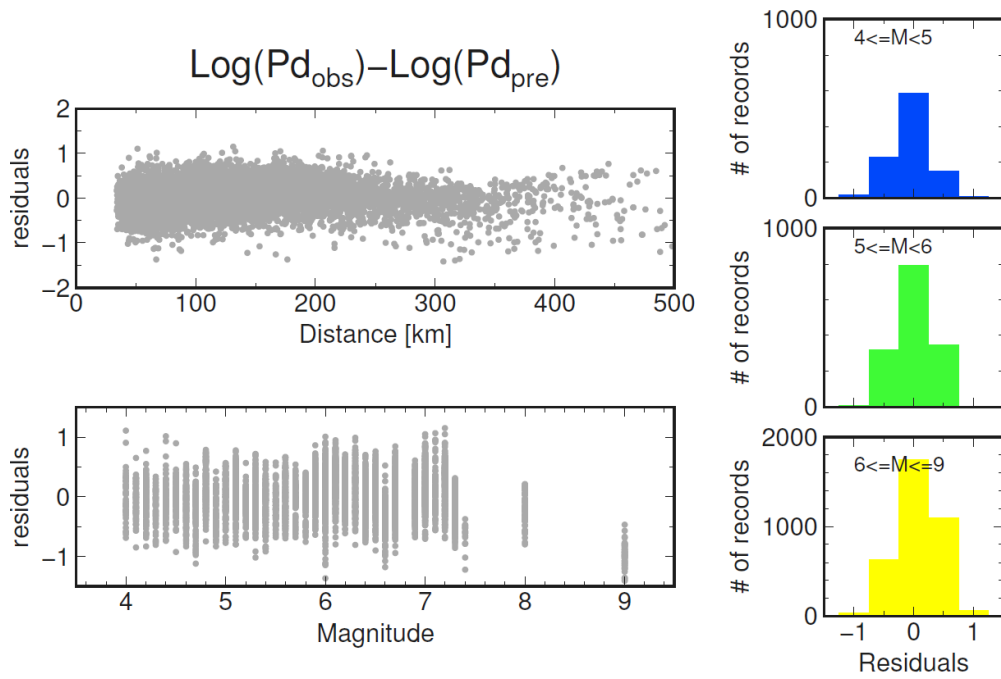
Supplementary Figure 3. Filter effect on fit parameters T1, B1, T2, B2. The figure shows the four parameters resulting from the fitting procedure as a function of magnitude using different cut-off frequencies for the high-pass filtering operation on displacement waveforms: 0.075 Hz (dark blue circles), 0.01 Hz (yellow circles), 0.1 Hz (brown circles) and 0.05 Hz (light blue circles). **a,** First corner time of the linear piecewise fit model (T1) as a function of magnitude for all the analysed events, on a log-linear scale. **b, S**lope of the first line piece (B1) as a function of magnitude for all the analysed events, on a log-linear scale. **c,** Second corner time of the linear piecewise fit model (T2) as a function of magnitude for all the analysed events, on a log-linear scale. **d,** Slope of the second line piece (B2) as a function of magnitude for all the analysed events, on a linear scale. In all panels the solid grey line shows the linear best fit function when the cut-off frequency of 0.075 Hz is used and the thin dashed lines are the corresponding 95% confidence bounds for new observations.

Supplementary Figure 4. Distance effect on fit parameters T1, B1, T2, B2. The figure shows the four parameters resulting from the fitting procedure as a function of magnitude using different distance ranges for the average log P_d computation: 0-100 km (brown circles), 100-200 km (yellow circles), 0-200 km (light blue circles) and 0-500 km (light blue triangles). **a,** First corner time of the linear piecewise fit model (T1) as a function of magnitude for all the analysed events, on a log-linear scale. **b,** Slope of the first line piece (B1) as a function of magnitude for all the analysed events, on a log-linear scale. **c,** Second corner time of the linear piecewise fit model (T2) as a function of magnitude for all the analysed events, on a log-linear scale. **d,** Slope of the second line piece (B2) as a function of magnitude for all the analysed events, on a linear scale. In all panels the solid grey line shows the linear best fit function when the cumulative ranges are used (0-200 km for magnitudes < 7 and 0-500 km for magnitudes ≥ 7) and the thin dashed lines are the corresponding 95% confidence bounds for new observations.

Supplementary Figure 5. Source depth effect on fit parameters T1, B1, T2, B2. The figure shows the four parameters resulting from the fitting procedure as a function of magnitude with a colour scale representing the source depth. **a,** First corner time of the linear piecewise fit model (T1) as a function of magnitude for all the analysed events, on a log-linear scale. **b,** Slope of the first line piece (B1) as a function of magnitude for all the analysed events, on a log-linear scale. **c,** Second corner time of the linear piecewise fit model (T2) as a function of magnitude for all the analysed events, on a log-linear scale. **d,** Slope of the second line piece (B2) as a function of magnitude for all the analysed events, on a linear scale. In all panels the solid grey line shows the linear best fit function when the cumulative ranges are used (0-200 km for magnitudes < 7 and 0-500 km for magnitudes ≥ 7) and the thin dashed lines are the corresponding 95% confidence bounds for new observations.

Supplementary Figure 6. Residual distribution. The figure shows the difference between the logarithm of the observed P_d and the logarithm of the predicted P_d as a function of distance (top) and magnitude (bottom). Right panel shows the histograms of the residuals in different magnitude classes.
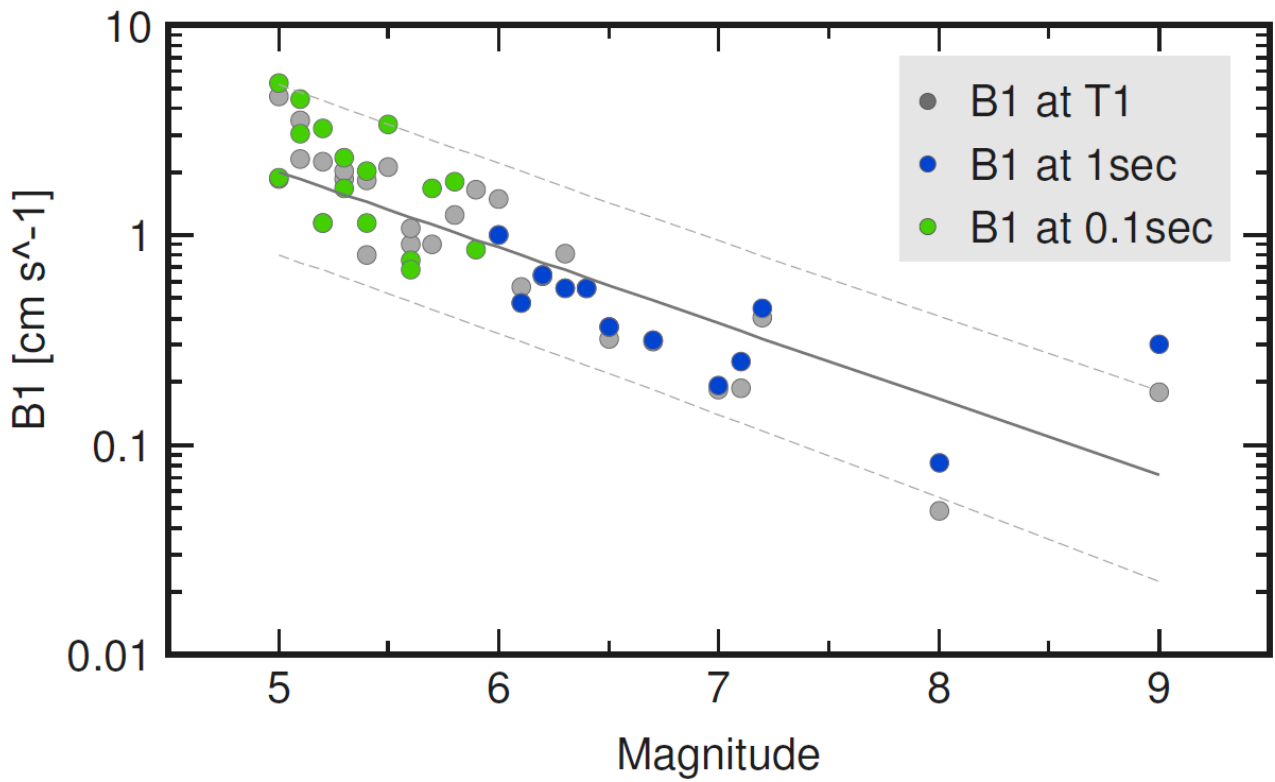
Supplementary Figure 7. Slope estimates within a fixed time window. The figure shows the initial slope of the LPW curves, obtained within a fixed time window as a function of magnitude. The slope is obtained using a single linear fit function. For $M \geq 6$ events (blue circles), the slope is measured within 1 second, whereas for $M < 6$ (green circles) the slope is measured within 0.1 second. Light grey circles represent the slopes estimated with the 3-piecewise model, up to the first corner time T1.

Supplementary Tables

Supplementary Table 1. List of the analysed events. The table includes the list of the selected earthquakes and their source parameters. For each event, from left to right, the table shows: date, time, location coordinates (longitude, latitude), source depth and magnitude (as provided by the JMA, Japan Meteorological Agency).

Supplementary Table 1: List of the analysed events					
Date (YYYY-MM-DD)	**Time (HH:MM:SS)**	**Lon (°)**	**Lat (°)**	**Depth (km)**	*M*
2003-09-26	04:50:00	144.07	41.78	42	8.0
2003-09-26	06:08:00	143.69	41.71	21	7.1
2008-05-08	01:45:00	141.61	36.23	51	7.0
2008-06-14	08:43:00	140.88	39.03	8	7.2
2010-03-29	10:29:00	139.88	36.03	43	4.3
2010-05-01	18:20:00	139.19	37.56	9	4.9
2010-07-21	06:19:00	135.69	34.21	58	5.1
2010-10-24	13:50:00	139.85	36.06	45	4.4
2010-11-05	19:14:00	139.84	36.06	45	4.6
2011-03-11	14:46:00	142.86	38.10	24	9.0
2011-03-12	03:59:00	138.60	36.99	8	6.7
2011-03-19	18:56:00	140.57	36.78	5	5.8
2011-03-23	07:12:00	140.79	37.08	8	5.0
2011-04-01	20:58:00	142.16	39.33	45	5.9
2011-04-11	17:16:00	140.67	36.94	6	6.4
2011-04-11	20:42:00	140.63	36.97	11	5.7
2011-04-12	14:07:00	140.64	37.05	15	5.0
2011-04-26	21:12:00	139.98	36.08	46	6.0
2011-06-18	20:31:00	141.82	37.62	28	5.5
2011-06-23	06:51:00	142.59	39.95	36	5.6
2011-07-08	03:35:00	141.13	37.10	55	6.3
2011-07-25	03:51:00	141.63	37.71	46	6.2
2011-07-31	03:54:00	141.22	36.90	57	6.1
2011-09-15	17:00:00	141.48	36.26	51	5.2
2011-09-17	04:26:00	143.09	40.26	7	4.4
2011-09-21	22:30:00	140.58	36.74	9	5.3
2011-11-20	10:23:00	140.59	36.71	9	4.9
2012-01-12	12:20:00	141.30	36.97	33	5.4
2012-01-28	07:39:00	138.98	35.49	18	4.7
2012-01-28	07:43:00	138.98	35.49	18	4.7
2012-01-29	16:46:00	138.98	35.54	21	5.2
2012-02-11	10:27:00	139.79	36.09	46	5.3

2012-03-01	07:32:00	140.63	36.44	56	5.4
2012-03-27	20:00:00	142.33	39.81	21	4.8
2012-05-05	18:56:00	137.17	35.19	45	6.5
2012-05-18	17:19:00	139.84	36.13	51	5.1
2012-06-01	17:48:00	139.88	36.03	44	4.5
2012-06-09	04:07:00	138.37	35.02	26	4.2
2012-06-17	16:13:00	140.01	36.18	51	4.8
2012-07-12	13:54:00	139.88	36.08	46	4.6
2012-07-16	04:31:00	139.82	36.15	52	4.1
2012-08-03	22:19:00	139.90	36.03	43	5.6
2012-08-30	04:17:00	139.85	36.12	49	4.5

Supplementary Table 2. Coefficients of equation (1) for each magnitude and distance range.
Table 2 shows the coefficients A, B, C of equation (1) of the main text for each magnitude and corresponding distance range. Equation (1) is used to compare the observed P_d values at different stations and to correct them for the geometrical attenuation effect.

Supplementary Table 2. Coefficients of equation (1) for each magnitude and distance range		
Coefficient	$M<7$ & $0<R\leq200$ km	$M\geq7$ & $0<R\leq500$ km
A	-2.89 ± 0.12	-2.24 ± 0.04
B	0.62 ± 0.01	0.59 ± 0.01
C	-1.25 ± 0.06	-1.51 ± 0.02

Supplementary Discussion

Resolution and tests

To assess the reliability and robustness of the results, we identified different factors that could have influenced the observed shape of the log P_d vs. time curves. Among them, the high-pass filtering operation, the propagation effects and the source depth, are the most relevant. Hence, we separately investigated each of these factors to understand their influence on the analysis.

Filter effect

Because the corner frequency of the event depends on the magnitude, the high-pass filtering operation may distort the recorded waveforms, eventually resulting in an artificial variation of the initial slope of LPW curves. We performed two tests to exclude that the observed trend of T1, B1, T2 and B2 with magnitude is an artificial effect of the high-pass filtering operation.

First we repeated the entire analysis using unfiltered velocity and acceleration records. With the same approach as used for P_d, we measured the initial Peak Velocity (P_v) and the initial Peak Acceleration (P_a) and obtained the corresponding log. P_v vs. PTW and log. P_a vs. PTW curves. We fitted these curves and estimated T1, B1, T2, B2 parameters. Supplementary Figure 2 shows that the scaling of these parameters with magnitude is preserved even when unfiltered acceleration or velocity records are used.

A second test consists in the use of different cut-off frequencies for the high-pass filtering operation on displacement waveforms. We tested 0.01 Hz, 0.075 Hz, 0.1 Hz and 0.5 Hz. For each filter we obtained the log P_d vs. time curves and estimated T1, B1, T2 and B2 parameters, as described above. Supplementary Figure 3 shows the parameters resulting from the fitting procedure as a function of magnitude using different cut-off frequencies for the filter. The trend of parameters with magnitude is preserved when different filters are used.

Distance effect

The initial P-wave amplitude may be strongly affected by all the propagation effects related to crustal structure, geometrical spreading and anelastic attenuation. In addition to this, we must consider that in our analysis, due to the natural amplitude decay with distance and magnitude, for the largest events P_d is measured up to larger hypocentral distances ($R>200$ km). Moreover the largest events in our database are mostly offshore earthquakes associated to the subduction zone. An intrinsic bias due to distance-attenuation and multipathing could therefore affect the results of this analysis. A further source of error can be the distance correction of P_d, which is necessary to compare records at different stations. As a matter of fact, an inadequate distance correction (i.e., an inadequate coefficient C of equation (1) of the main text) could influence the initial slope variation of log P_d curves.

To understand if these effects are contributing to (and are responsible of) the different initial slope, we looked at different events in the same distance range. We prior determined the coefficients A, B, C of equation (1) for each distance range (i.e., 0-100km, 0-200km, 100-200km, 0-500km); we than corrected P_d using the appropriate coefficient for each distance and fitted the LPW curves to determined T1 and B1. Supplementary Figure 4 shows the two parameters as a function of magnitude and of the distance range. Assuming that T1 and B1 depend on both magnitude and distance, we would expect to see an evident (and systematic) variation when a different distance range is used. This effect, instead, is not evident from the plot, suggesting that the observed trend with magnitude is not controlled by the distance range. The trend of T1 and B1 with magnitude is preserved even when restricting the average to a smaller distance range, although it is evident but less pronounced when looking at larger distances, due to the poor magnitude coverage.

Furthermore, if the anelastic attenuation was responsible of the observed T1 and B1 trend with *M* we would expect an increase of the initial slope with magnitude, since higher frequencies attenuate more rapidly, as opposed to the observed behaviour. A more realistic situation in our analysis is that, because higher frequencies are mostly recorded at small distances whereas lower frequencies are mainly associated to long distances, the overall anelastic attenuation effect is averaged out. Another issue is the azimuthal variation in displacement amplitude which in case of only few stations available may introduce variability of the log P_d vs. time curves. The average on a large number of data (i.e. hundreds of stations, as done in our analysis) is generally expected to reduce the influence of the radiation pattern.

Finally, we considered that averaging the displacement amplitude over a wide range of distances may result in the inclusion of different P-arrivals within the considered PTW (i.e., Pn, Pg phases) thus possibly affecting the slope estimate. The results of Supplementary Figure 4 show that the trend of T1 and B1 with magnitude is preserved even when restricting the average to a smaller distance range where a single P-wave arrival is expected.

Source depth effect

The earthquake depth might impact on source processes and frictional properties and therefore on the observation regarding the evolution of P-wave peak displacement with time. The scatter plot of Figure 2b of the main text shows that there is no magnitude-depth dependence for the earthquakes in our database. Thus, the observed T1, B1, T2 and B2 vs. magnitude trend should not be influenced by the frictional processes occurring at different depths. To clarify this point, in Supplementary Figure 5 we plotted the fit parameters as a function of magnitude, with a colour scale representing the event depth. We did not find any clear evidence that the source depth may affect the initial P-wave displacement evolution.

Residual analysis

To compare the observed P_d at different stations, a standard Ground Motion Prediction Equation (in the form of equation 1 of the main text) has been used. We are aware that more complex GMPEs may better describe the amplitude decay. However, in the context of real-time applications (for which the P_d parameter was first introduced), the generally adopted decay law[1-2] has the same form of the one used in this study.

To show that the attenuation relationship does not introduce any bias, we plotted the residuals LogP_d(observed)-LogP_d(predicted) as a function of distance and magnitude and the histograms of the residual in different magnitude classes. Supplementary Figure 6 (top panel) shows that residuals are distributed around zero in the whole distance range and no evident bias appears neither from the scatter plot, nor from the histograms (right panel). A similar behaviour is observed in the plot of logP_d(observed)-LogP_d(predicted) as a function of magnitude (bottom panel). Also in this case we found that residuals are distributed around zero, except for the case of the *M*=9.0 Tohoku-Oki earthquake.

For this single event, the deviation from the zero value is related to the problem of parameter saturation, when a short P-wave time window is used to estimate the magnitude. The saturation effect of the initial amplitude is a well-known problem and has been extensively discussed in the literature[3-6]. The saturation is likely due to the use of only a few seconds of the P-wave, which cannot capture a sufficient history of the rupture process for a large earthquake[7], but it can be overcome when a larger time window is used[8]. Thus, for the case of the Tohoku-Oki event, the use of a fixed P-wave time window of 3 seconds, results in significantly lower initial peak displacement values. It is worth to note, however, that no evident bias appears from the residual vs. distance plot.

To correct P_d values for the distance effect in Figure 3a of the main text, only the distance contribution is used. If a bias effect with magnitude would exist, it will not affect the slope of LPW curves.

Dependency of B1 from T1: analysis at a fixed time window
The most relevant result of our paper is the observed variation in the initial slope of P_d for different magnitude earthquakes (i.e., the trend of B1 vs. *M*). Without using any fitting procedure, this result appears evident from a visual inspection of the curves and it is even more evident when comparing the way the initial P_d increases for different magnitude class events (Figure 4c of the main text).

However, the use of a 3-piecewise linear function to fit the data might have introduced some bias in the estimation of B1 and T1. We then tested the persistence of the B1 trend for increasing magnitude also by limiting all the LPW curves to a fixed time window. We selected 1 s for $M \geq 6$ and 0.1 s for $M < 6$, and for each event, we estimated the initial slope (by fitting data with a single straight line). Supplementary Figure 7 shows that the slope estimates obtained within fixed time are consistent with those obtained by fitting the whole curves and that the trend of B1 with magnitude is still observed when only the very beginning of the curves is considered.

Supplementary Methods

Database selection
The database consists of 43 moderate-to-strong earthquakes in the magnitude range $4 \leq M \leq 9$, located all over the Japanese active seismic region. To avoid any bias related to specific criteria for data selection (tectonic environment, dominant fault mechanism, aftershock/mainshock, event depth, …) we selected earthquakes occurred all over Japan, in different tectonic environments, both crustal and subcrustal events, as shown in Figure 1 and Figure 2b of the main text. We tried to avoid the inclusion of aftershocks belonging to the same sequence and preferred earthquakes recorded at the largest number of stations. The data set includes different types of fault mechanism with events that are distributed in depth between 5 and 60 km. Among the 12 earthquakes with $M \geq 6.0$ we included the 2003, M 8.0 Tokachi-Oki event and the 2011, M 9.0 Tohoku-Oki mega thrust earthquake. The full list of analysed events is provided in Supplementary Table 1.

We selected data up to a maximum hypocentral distance of 500 km, for a total of 7514, 3-component waveforms, recorded at 1208 stations. Acceleration waveforms have been provided by K-Net and Kik-Net online databases (http://www.kyoshin.bosai.go.jp/).

For earthquake waveforms extracted from K-Net and KiK-Net catalogues, the information concerning the event locations and magnitudes are provided by the JMA. By a comparison between the JMA magnitude (M_{JMA}) and the moment magnitude (M_w), different authors[9-11] have found negligible differences between the two scales the magnitude range 4 to 8. In this work we simply denote magnitude as M.

Supplementary References

1) Wu, Y.-M., & Zhao, L. Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophys. Res. Lett*. **33**, L16312 (2006).
2) Lancieri, M. & Zollo, A. A Bayesian approach to the real-time estimation of magnitude from the early P and S wave displacement peaks. *J. Geophys. Res*. **113** B12302.
3) Kanamori, H. Real-Time Seismology and Earthquake Damage Mitigation. *Annu. Rev. Earth Planet. Sci*. **33**, 195-214 (2005).
4) Rydelek, P. & Horiuchi, S. Is earthquake rupture deterministic?. *Nature* **442** (2006).
5) Rydelek, P. & Wu, C. & Horiuchi, S. Comment on "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records" by Aldo Zollo, Maria Lancieri, and Stefan Nielsen. *Geophys. Res. Lett.* **34** (2007).
6) Zollo, A. & Lancieri, M. & Nielsen, S. Reply to comment by P. Rydelek et al. on "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records,". *Geophys. Res. Lett*. **34** (2007).
7) Festa, G. & Zollo, A. & Lancieri, M. Earthquake magnitude estimation from early radiated energy. *Geophys. Res. Lett*. **35,** L22307 (2008).
8) Colombelli, S. & Zollo, A. & Festa, G. & Kanamori, H. Early magnitude and potential damage zone estimates for the great M_w 9 Tohoku-Oki earthquake. *Geophys. Res. Lett*. **39,** L22306 (2012).
9) Dziewonski, A. M. & Chou, T.-A. & Woodhouse, J. H. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res*. **86**, 2825–2852 (1981).
10) Katsumata, A. Comparison of magnitudes estimated by the Japan Meteorological Agency with moment magnitudes for intermediate and deep earthquakes. *Bull. Seism. Soc. Am.* **86**, 832-842 (1996).
11) Oth, A. & Bindi, D. & Parolai, S. & Di Giacomo, D. Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: insights from KiK-net data in Japan. *Geophys. Res. Lett*. **37**, L19304 (2010).