nature geoscience

ARTICLES

https://doi.org/10.1038/s41561-018-0145-1

Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks

William L. Ellsworth[1*] and Fatih Bulut[2]

Much of what we know about the nucleation of earthquakes comes from the temporal and spatial relationship of foreshocks to the initiation point (hypocentre) of the mainshock. The 1999 M_w 7.6 Izmit, Turkey, earthquake was preceded by a 44-minute-long foreshock sequence, which built in intensity as the time of the mainshock approached. Here we apply a series of high-resolution methods to the sparse seismological observations to determine the spatial and temporal relation of the foreshocks to the mainshock hypocentre. We find that the foreshocks form a contiguous series of ruptures that progressed systematically from west to east towards the mainshock hypocentre, located at the extreme eastern edge of the foreshocks. The Izmit foreshock sequence occurred as a triggered cascade in which one foreshock loads the adjacent fault patch causing it to fail, with the mainshock initiation no different than another foreshock. We find no evidence to support a hypothesized precursory aseismic driving process. If we are to resolve the role of aseismic deformation and possible role of fluid overpressure in the earthquake nucleation process it will likely require measurements made in the near field of the hypocentre.

How do earthquakes begin[1–7]? This question lies at the heart of our understanding of earthquake physics, including the long-contentious issue of their predictability. Empirically, we know that the initial radiation of waves comes from a very limited area on the fault plane—the initial hypocentre—and grows at a magnitude-independent rate until rupture finiteness effects appear[8,9]. This is true regardless of whether the earthquake achieves a final magnitude M_w of −1 or 9, a range of 15 orders of magnitude in seismic moment. Theoretically, we know that a transition from a locked state at the hypocentre to a dynamically growing rupture takes place through the pre-seismic growth of slip instability over a finite area of the fault plane[10]. Laboratory experiments indicate that the instability might have dimensions as small as a few decimetres[11]. Even if the nucleation zone has dimensions of tens of metres, it would be essentially impossible to detect at the surface for hypocentres at typical crustal depths of 5–15 km due to its small strain amplitude relative to the background Earth noise[12].

Foreshocks, small earthquakes that sometimes occur in the vicinity of the initial hypocentre, provide one of the few windows we have into the nucleation process. Although less than one earthquake out of twenty is followed by a larger event[13], those that are foreshocks suggest the existence of an evolving nucleation process[4]. However, their interpretation in terms of physical processes taking place on the fault plane remains unresolved[7]. One view is that they are triggered by aseismic slip over an extended area surrounding the eventual mainshock hypocentre (preslip model[2]). If correct, the underlying aseismic slip might be a precursor to the earthquake[6]. Alternatively, they might occur by neighbour-to-neighbour stress transfer between one foreshock and another without an aseismic slip component (cascade model[2,14,15]). Under this interpretation, the foreshocks are no different than any other set of clustered earthquakes and the mainshock is just a random outcome of triggering[5].

One approach to discriminating between these end-member alternatives focuses on the spatial and temporal evolution of the foreshocks as revealed by their seismograms[4,7]. If aseismic slip drives the foreshock process, their hypocentres would be expected, on average, to populate the preslip zone. Foreshocks might also repeat, as asperities were reloaded by continuing preslip. Under the cascade hypothesis, foreshocks would cluster in close enough proximity to one another to permit event-to-event triggering from either dynamic and/or static stress changes.

Few foreshock sequences of large ($M > 7$) earthquakes have been captured by local seismic stations. One of the best examples is that of the 17 August 1999 M_w 7.6 Izmit earthquake (Fig. 1). This earthquake nucleated at a depth of 15 km and ruptured bilaterally for over 120 km along the North Anatolian Fault. A previous study[16] used waveform similarity among 18 foreshocks recorded by a single seismic station in the 44 minutes before the earthquake to conclude the foreshocks represented repeated rupture of the same fault patch and inferred that they were driven by aseismic slip. However, similarity in waveform only implies close proximity in space, not repeating rupture[17]. We investigate the foreshocks in this paper using seismograms from nine additional seismic stations and apply a suite of waveform-based methods to identify, precisely locate and quantify the foreshocks, and interpret their progression toward the Izmit earthquake hypocentre.

Izmit foreshocks revisited. Foreshock waveforms from 10 seismic stations located within 100 km of the mainshock are used to determine the evolution of the sequence in space and time, and to place the foreshocks into a common frame of reference with the mainshock hypocentre. The methods include template matching, cross correlation timing, precision hypocentre (point of rupture initiation) and hypocentroid (centre of moment release) determination, waveform-based clustering and spectral analysis for source parameter measurement (see Methods for details).

The foreshock sequence contained 26 events with magnitudes between M 0.9 and 2.8 and lasted just 44 minutes (Fig. 1, Supplementary Table 1). No additional foreshocks were found in the 12 months before the mainshock (see Methods). Results from cluster analysis, group the foreshocks into four episodes, each with one principal foreshock with $M \geq 2.0$ (Fig. 2). Hypocentroids of these

[1]Department of Geophysics, Stanford University, Stanford, CA, USA. [2]Geodesy Department, Kandilli Observatory and Earthquake Research Institute, Boğaziçi University, İstanbul, Turkey. *e-mail: wellsworth@stanford.edu

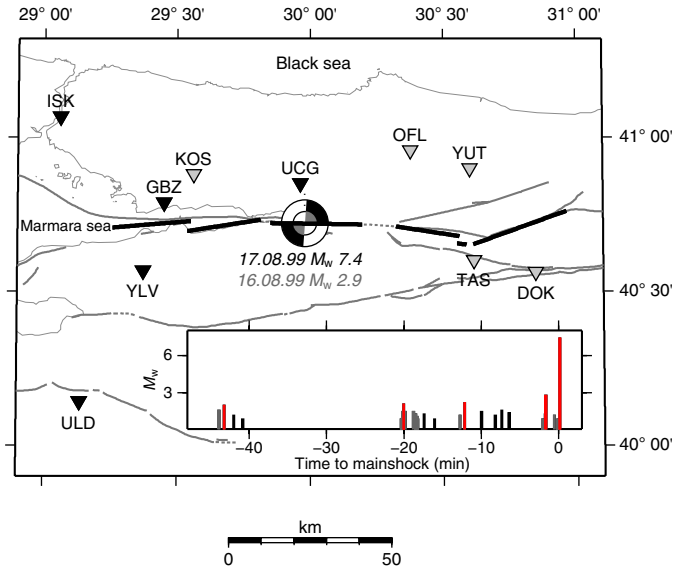

Fig. 1 | Location map for the Izmit earthquake showing seismographs (triangles) used to locate the four large foreshocks. Mainshock located with stations denoted by black triangles. Focal mechanism of the mainshock from centroid moment tensor (CMT) catalogue (compressional quadrants in black). Focal mechanism of the largest foreshock obtained using first motions at 19 regional seismographs (compressional quadrants in grey). Surface rupture in black[35] and other active faults in grey[36]. Inset: magnitude–time diagram. The four largest foreshocks and the mainshock (red), other foreshocks from ref. [16] (grey) and from this study (black).

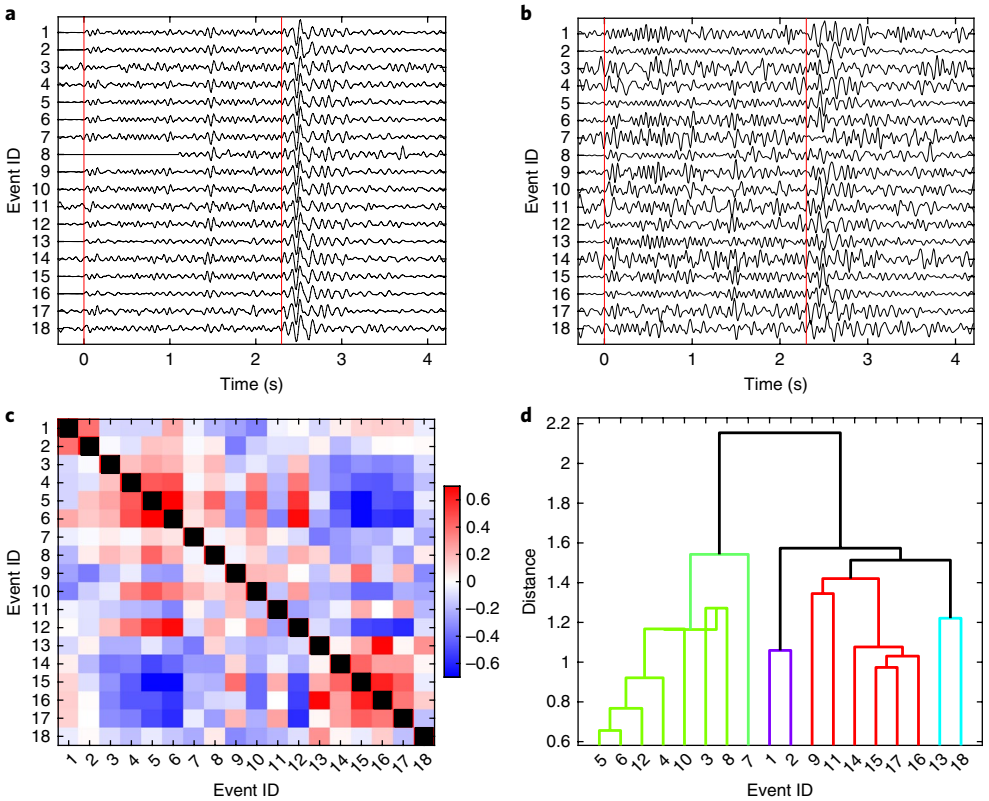

Fig. 2 | Cluster analysis of the four largest foreshocks. a, Seismograms of the foreshock sequence at UCG aligned on the S wave. The four largest foreshocks are number 2, 6, 13 and 16 (see Supplementary Table 2 for compete list of foreshocks). **b**, Residual seismograms after removal of common-mode signal. Vertical lines in **a** and **b** mark the average P and S arrival times. **c**, Pairwise cross-correlation coefficients for residual seismogram measured in the frequency band 3–18 Hz. **d**, Event family tree from **c** determined though cluster analysis.

four largest foreshocks, identified hereafter by their hour and minute of occurrence at 23:18 (M_w 2.0), 23:41 (M_w 2.1), 23:49 (M_w 2.2) and 23:59 (M_w 2.9), locate between 125 and 27 m west of the mainshock hypocentre (Supplementary Table 1). Both the centroid focal depths of the foreshocks and the mainshock hypocentral depth are within ±30 m. The foreshocks' hypocentroids systematically migrate from

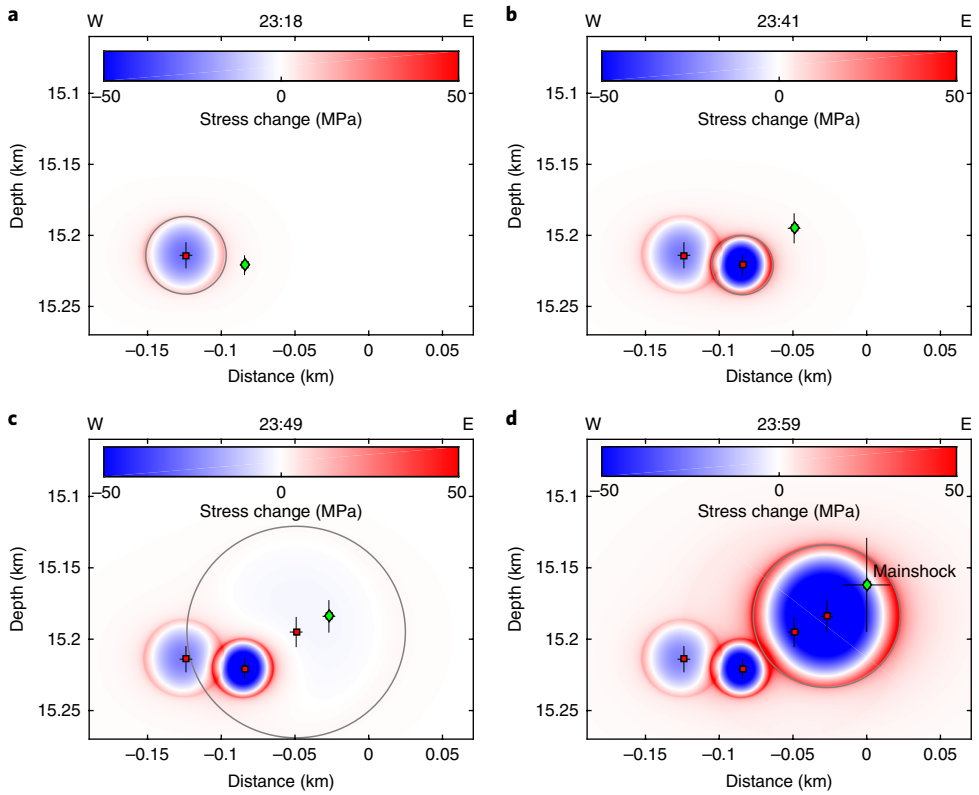

Fig. 3 | East–west cross sections of the evolving shear stress changes on the fault plane during the foreshock sequence to the 1999 Izmit earthquake. **a**–**d**, Each frame corresponds to the origin time of one of the four largest foreshocks and adds the stress change from that event. Stress change based on a previously reported source model[18]. Red squares and error-bars show hypocentroids and 2σ location uncertainty of events occurring through the time indicated in the figure heading. Green diamonds denote the hypocentroid of the next foreshock in the sequence (**a**–**c**) and the mainshock hypocentre (**d**).

west to east along the fault, approaching the mainshock hypocentre just minutes before it initiated the M_w7.6 earthquake at 00:01:38, a sequence of events suggestive of a triggered cascade.

Although the smaller foreshocks could not be located, they form an integral part of the failure process as it progresses from west to east along the fault plane. Because the foreshock waveforms are slightly dissimilar from one another (Fig. 2b) it is unlikely that any are repeating earthquakes. The low cross-correlation values of the residual seismogram pairs provide quantitative evidence that the small foreshocks occur at spatially distinct locations (Fig. 2c). In other words, we find no evidence for "repetitive seismic bursts" proposed previously[16]—with one key exception discussed below.

Foreshock interactions. We model the static stress change associated with the foreshocks to better understand the interactions among the foreshocks and mainshock hypocentre using source dimensions determined by analysis of shear waves using a spectral ratio method and a previously reported dynamically consistent source model[18] (see Methods, Supplementary Fig. 1, Supplementary Table 2). Several of the largest foreshocks have surprisingly large stress drops, of up to 83 MPa or nearly half of the theoretically available shear stress for a strike slip rupture at 15 km depth[19]. We model the shear stress redistribution using a function reported previously[20] to avoid nonphysical singularities at the rupture edge. Each foreshock drops the shear stress within its source radius; exterior to the rupture the shear stress increases but decays rapidly. Studies of earthquake triggering elsewhere indicate that the stress change can trigger earthquakes out to several source radii[21].

Figure 3 summarizes the stages of the foreshock sequence. In each panel the hypocentroids of the current and previous events are marked with a red square, and the hypocentroid of the next event by a green diamond (also see Supplementary Figs. 2–4). Observe that each succeeding hypocentroid lies to the east of the previously ruptured areas on the fault plane. The first three large foreshocks abut with minor overlap, suggesting that they represent a cascade (Fig. 3a–c). The fourth large foreshock largely re-ruptures the same area as the third foreshock (Fig. 3d), although it should be noted that we could not resolve if they are coplanar or lie on subparallel planes offset by several metres. Regardless of their exact alignment or choice of rupture model, the weak stress drop of the third event followed by the largest of the fourth produces a concentrated stress increase on the adjacent unruptured parts of the fault (Supplementary Table 2).

The mainshock nucleation point (Fig. 3d) locates near the edge of the last large foreshock. Taking into account the uncertainty in the mainshock hypocentre and the simplicity of the assumed slip model, we cannot be sure if the mainshock initiated where slip had just occurred, or just beyond. Overall, the foreshocks migrate with a stress triggering front from west to east towards the mainshock, driven by the four large foreshocks. We note that studies of the first 10 seconds of the rupture process of the mainshock indicate that it first ruptured principally to the east[22,23]. It seems likely that the accumulated stress drop of the foreshocks influenced the initial rupture growth.

The 23:49 foreshock. Among the four large foreshocks, the 23:49 event stands out as anomalous with a stress drop an order of magnitude smaller than the other three and an average displacement of just 4 mm, compared to 17, 39 and 100 mm, for 23:18, 23:41 and 23:59, respectively (Supplementary Table 2). The 23:49 foreshock also exhibits rupture directivity not seen in the other foreshocks with corner frequency measurements varying systematically with azimuth (Fig. 4 and Supplementary Table 3). At station KOS

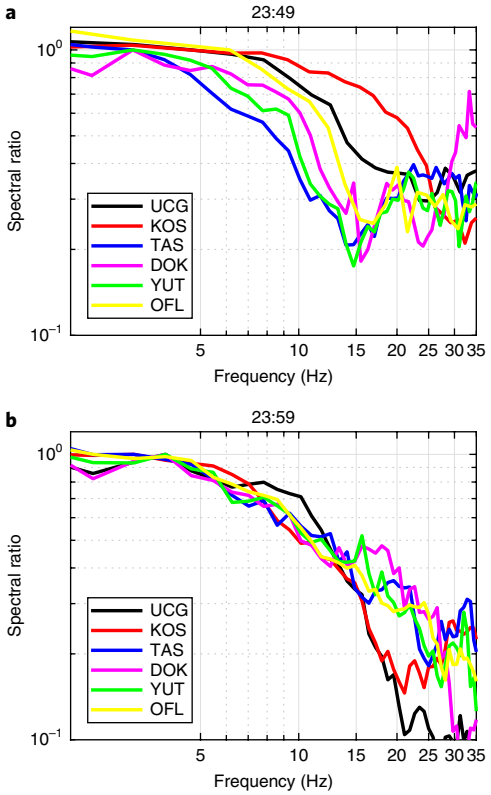

Fig. 4 | P wave spectral ratios for the 23:49 and 23:59 foreshocks computed using event 23:41 as the denominator and normalized at 1 Hz. Event 23:49 shows significant azimuthal dependence in corner frequencies. For the 23:49 event, KOS and UCG show significantly higher corner frequencies than TAS, DOK, YUT and OFL, indicating that rupture began to the east of the hypocentroid and propagated to the west. The 23:59 event has almost the same corner frequency at every station, indicating bilateral rupture.

(azimuth 286) the P wave corner frequency is 19 Hz whereas at the opposite azimuths (stations TAS, DOK and YUT) the corner frequency ranges from 9 to 11 Hz. This implies westward directivity[24] for 23:49. Its 23 Hz S wave corner frequency at station KOS is higher than the P corner (19 Hz), providing a second line of evidence for directivity to the west. This places its hypocentre to the east of its hypocentroid. The larger foreshock, 23:59, ruptures symmetrically as it has almost identical corner frequency at each station (Fig. 4), placing its hypocentre near its hypocentroid.

The clustering of the minor foreshocks with the 23:49 event also stands out, as the foreshocks closest to 23:49 in time order cluster with the fourth large foreshock (23:59). Only the final foreshock at 00:01:30 (14 s before the mainshock) clusters with the 23:49 event. The westward directivity of the 23:49 rupture and absence of directivity in 23:59 suggests that the 23:49 hypocentre locates on the eastern edge of the sequence along with the final foreshock that clusters with it (Supplementary Table 1). The initiation of the mainshock has also been interpreted as consisting of two immediate foreshocks in the final 0.15 s, further emphasizing the cascading nature of the Izmit earthquake nucleation process[16].

Mechanisms underlying the cascade to failure. The brief foreshock sequence of the Izmit earthquake generally conforms to the event-to-event triggering of the cascade model. If triggering were the underlying process, why would the same fault area rupture twice within 10 minutes in the very low stress drop 23:49 and highest stress drop 23:59 foreshocks? We consider three hypotheses: (1) aseismic slip, (2) fault-valve behaviour and (3) rupture heterogeneity.

If aseismic slip accompanied the foreshocks, it likely amounted to no more than a few centimetres of fault displacement, as slip in the foreshocks would at a minimum keep pace with the aseismic slip because the foreshocks release both the short-term slip increment and the accumulated long-term slip deficit. The small foreshocks accompanying the 23:18 and 23:41 events ceased within minutes (Fig. 1 insert), implying that any creep accompanying those foreshocks decayed rapidly and was spatially restricted to their immediate rupture area. Because these foreshocks ruptured distinct areas on the fault, if they were loaded by creep the fault displacement would amount to at most a few centimetres. Foreshocks to the east soon followed, leading to the 23:49 event. Aseismic slip within the rupture area of 23:49 would decrease the shear stress and move the patch further from failure, so this seems unlikely as a mechanism promoting re-rupture. If creep occurred exterior to the foreshock rupture areas, it could reload the 23:49 rupture, but wouldn't explain the decaying aftershock sequence. The absence of any other repeating foreshocks argues against a process driven by aseismic slip (fault creep) as envisioned previously[16].

Re-rupture could occur without the need to restore the stress drop by reducing the effective normal stress by elevating the pore pressure within the foreshock rupture. More generally, fluid pressure could be driving the entire foreshock sequence. The geology of exhumed faults frequently contains evidence for large fluid pressure fluctuations when rupture opens a fault valve[25]. Fluid overpressure commonly drives seismic swarms in volcanic and geothermal settings. The Izmit foreshocks bear a strong similarity to the propagation of rupture in the 2014 Long Valley, California, swarms, where burst of earthquakes expanded the activated fault area as it propagated along the fault plane in neighbour-to-neighbour triggering[26]. In situ experiments in a mine in France reactivated a fault through controlled fluid injection that resulted in a large increase in hydraulic diffusivity ($>10^3\,m^2\,s^{-1}$) of the fault due to shear slip of $<100\,\mu m$ (ref. [27]). The 23:49 average rupture displacement of 4 mm would have enhanced the permeability of the fault zone over its approximately 200 m long rupture, providing a pathway for fluid overpressure to rapidly invade its rupture. A pressure rise of about 2 MPa would be needed to bring the rupture area back to the failure condition.

But did fluids play a role at Izmit? A previous study[16] identified a weak, low-frequency increase in seismic noise at UCG coincident with the first foreshock and proposed that it was the signature of fault creep. The available seismic recordings from the other stations studied here can neither confirm nor refute whether the noise originated in the foreshock zone or was unrelated to the earthquake. If it were related, it is unlikely that waves excited by fluid flow into the just-faulted areas could explain the observation[28].

One feature of the foreshocks that we can be confident in is the spatial roughness of slip in each earthquake. Whereas a constant stress drop crack is often used to model slip as smoothly varying within the rupture, both theoretical considerations and observations tell us that neither the slip nor the resulting stress distribution are smooth[20,29–32]. A constant stress drop crack model for the M4.3 foreshock of the 1992 M6.1 Joshua Tree, California, earthquake placed the mainshock hypocentre within foreshock rupture[4] where stresses decreased, while a more accurate finite fault reconstruction of the foreshock slip placed the hypocentre outside of the rupture where stress increased[33]. If the 23:59 foreshock nucleated where the shear stress increased as a result of the 23:49 rupture, it would simply be a triggered earthquake with a hypocentre interior to the 23:49 rupture, and there is no need to invoke either aseismic slip or fault-valve behaviour to explain it.

The nucleation of the Izmit earthquake and its foreshock sequence can be most simply explained as a cascade of triggered events in which the last one got away and grew to a M_w7.6 earthquake. This behaviour has been proposed as a general property of foreshock sequences in which any earthquake can trigger another[5]. While aseismic slip (unobservable) or fault-valve behaviour (at best weakly supported by the evidence) are possible contributors, the

tight spatial grouping of the foreshocks, their systematic progression from west to east, and absence of ubiquitous repeating foreshocks all support the cascade hypothesis. While re-rupture of the faulted area of the penultimate large foreshock (23:49) by the largest foreshock (23:59) stands out as the single issue for the cascade model, it is not without precedent. A previous study[4] reported cases of re-rupturing foreshocks, albeit with analysis methods that were less precise than those used here.

We were fortunate to have just enough seismic recordings of the Izmit earthquake to find all the foreshocks with $M \gtrsim 1.0$, precisely locate and measure the source parameters of the $M \geq 2.0$ foreshocks, and place the mainshock hypocentre in the same frame of reference. But this is far from the state of the art today. Advances in seismicity analysis when combined with better instrumentation permit us to look much more completely into seismic sequences with a magnitude of completeness less than M 0 (ref. [26]). Yet, even today, we would benefit from imaging the rupture in each event to better constrain the stress evolution, a prospect that is within sight but still beyond our reach. To resolve the role of aseismic deformation and fluid overpressure in the earthquake nucleation process measurements made in the ultra-near field of earthquakes are critical[34]. Yet, even today, we would benefit from seismic imaging of the rupture in each foreshock to better constrain the stress evolution, a prospect that is within sight if just beyond our reach.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41561-018-0145-1.

Received: 29 September 2017; Accepted: 30 April 2018;
Published online: 04 June 2018

References

1. Jones, L. M. & Molnar, P. Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults. *J. Geophys. Res. Solid Earth* **84**, 3596–3608 (1979).
2. Ellsworth, W. L. & Beroza, G. C. Seismic evidence for a earthquake nucleation phase. *Science* **268**, 851–855 (1995).
3. Abercrombie, R. E. & Mori, J. Occurrence patterns of foreshocks to large earthquakes in the western United States. *Nature* **381**, 303–307 (1996).
4. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Detailed observations of California foreshock sequences: implications for the earthquake initiation process. *J. Geophys. Res.* **101**, 22371–22392 (1996).
5. Helmstetter, A., & Sornette, D. Foreshocks explained by cascades of triggered seismicity. *J. Geophys. Res.* https://doi.org/10.1029/2003JB002409 (2003).
6. Bouchon, M., Durand, V., Marsan, D., Karabulut, H. & Schmittbuhl, J. The long precursory phase of most large interplate earthquakes. *Nat. Geosci.* **6**, 299–302 (2013).
7. Mignan, A. The debate on the prognostic value of earthquake foreshocks: a meta-analysis. *Sci. Rep.* **4**, 4099 (2014).
8. Uchide, T. & Ide, S. Scaling of earthquake rupture growth in the Parkfield area: self-similar growth and suppression by the finite seismogenic layer. *J. Geophys. Res. Solid Earth* **115**, B11302 (2010).
9. Noda, S. & Ellsworth, W. L. Scaling relation between earthquake magnitude and the departure time from P wave similar growth. *Geophys. Res. Lett.* **43**, 9053–9060 (2016).
10. Dieterich, J. H. Pre-seismic fault slip and earthquake prediction. *J. Geophys. Res.* **83**, 3940–3948 (1978).
11. Beeler, N. M. Review of the physical basis of laboratory-derived relations for brittle failure and their implications for earthquake occurrence and earthquake nucleation. *Pure Appl. Geophys.* **161**, 1853–1876 (2004).
12. Johnston, M. J. S., Borcherdt, R. D., Linde, A. T. & Gladwin, M. T. Continuous borehole strain and pore pressure in the near field of the 28 September 2004 M 6.0 Parkfield, California, earthquake: Implications for nucleation, fault response, earthquake prediction, and tremor. *Bull. Seismol. Soc. Am.* **96**, S56–S72 (2006).
13. Christophersen, A. & Smith, E. G. Foreshock rates from aftershock abundance. *Bull. Seismol. Soc. Am.* **98**, 2133–2148 (2008).
14. Wyss, M. & Brune, J. N. Regional variations of source properties in southern California estimated from the ratio of short-to long-period amplitudes. *Bull. Seismol. Soc. Am.* **61**, 1153–1167 (1971).
15. Fukao, Y. & Furumoto, M. Hierarchy in earthquake size distribution. *Phys. Earth Plant. Inter.* **37**, 149–168 (1985).
16. Bouchon, M. et al. Extended nucleation of the 1999 M_w 7.6 Izmit earthquake. *Science* **331**, 877–880 (2011).
17. Ellsworth, W. L. in *Urban Disaster Mitigation: The Role of Science and Technology* (eds Cheng, F. Y. & Sheu, M.-S.) 1–14 (Elsevier, Oxford, 1995).
18. Madariaga, R. Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.* **66**, 639–666 (1976).
19. Huang, Y., Ellsworth, W. L. & Beroza, G. C. Stress drops of induced and tectonic earthquakes in the central United States are indistinguishable. *Sci. Adv.* **3**, e1700722 (2017).
20. Andrews, D. J. A stochastic fault model: 1. static case. *J. Geophys. Res.* **85**, 3867–3877 (1980).
21. Shearer, P. M. Space-time clustering of seismicity in California and the distance dependence of earthquake triggering. *J. Geophys. Res. Solid Earth* **117**, B10306 (2012).
22. Bouchon, M. et al. Space and time evolution of rupture and faulting during the 1999 Izmit (Turkey) earthquake. *Bull. Seismol. Soc. Am.* **92**, 256–266 (2002).
23. Harris, R. A., Dolan, J. F., Hartleb, R. & Day, S. M. The 1999 Izmit, Turkey, earthquake: a 3D dynamic stress transfer model of intraearthquake triggering. *Bull. Seismol. Soc. Am.* **92**, 245–255 (2002).
24. Kaneko, Y. & Shearer, P. M. Seismic source spectra and estimated stress drop from cohesive-zone models of circular subshear rupture. *Geophys. J. Int.* **197**, 1002–1015 (2014).
25. Sibson, R. H. Implications of fault-valve behavior for rupture nucleation and recurrence. *Tectonophysics* **211**, 283–293 (1992).
26. Shelly, D. R., Ellsworth, W. L. & Hill, D. P. Fluid-faulting evolution in high definition: connecting fault structure and frequency-magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm. *J. Geophys. Res. Solid Earth* **121**, 1776–1795 (2016).
27. Guglielmi, Y., Cappa, F., Avouac, J.-P., Henry, P. & Elsworth, D. Seismicity triggered by fluid injection-induced aseismic slip. *Science* **348**, 1224–1226 (2015).
28. Dunham, E. M. & Ogden, D. E. Guided waves along fluid-filled cracks in elastic solids and instability at high flow rates. *J. Appl. Mech.* **79**, 031020 (2012).
29. Boatwright, J. A dynamic model for far-field acceleration. *Bull. Seismol. Soc. Am.* **72**, 1049–1068 (1982).
30. Gusev, A. A. Descriptive statistical model of earthquake source radiation and its application to an estimation of short-period strong motion. *Geophys. J. Int.* **74**, 787–808 (1983).
31. Mai, P. M. & Beroza, G. C. A spatial random field model to characterize complexity in earthquake slip. *J. Geophys. Res.* https://doi.org/10.1029/2001JB000588 (2002).
32. Abercrombie, R. E. Stress drops of repeating earthquakes on the San Andreas fault at Parkfield. *Geophys. Res. Lett.* **41**, 8784–8791 (2014).
33. Mori, J. Rupture directivity and slip distribution of the M 4.3 foreshock to the 1992 Joshua Tree earthquake, Southern California. *Bull. Seismol. Soc. Am.* **86**, 805–810 (1996).
34. Savage, H. M. et al. Scientific exploration of induced seismicity and stress (SEISMS). *Sci. Drill.* **23**, 57–63 (2017).
35. Barka, A. et al. The surface rupture and slip distribution of the 17 August 1999 Izmit earthquake (M 7.4), North Anatolian fault. *Bull. Seismol. Soc. Am.* **92**, 43–60 (2002).
36. Saroglu, F., Emre, O. & Kuscu, I. *Active Fault Map of Turkey* (General Directorate of Mineral Research and Exploration, Ankara, 1992).

Acknowledgements

We thank the seismic network operators of Kandilli Observatory and Earthquake Research Institute, the Marmara Research Center, GFZ Potsdam and Turkish Disaster Affairs for access to their seismic recordings of the Izmit earthquake foreshock sequence. R. Abercrombie, G. Beroza, M. Brehme, E. Dunham and F. Tilmann provided valuable input.

Author contributions

The authors shared the analysis and writing equally.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41561-018-0145-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to W.L.E.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Seismic waveforms. We analysed seismograms from 10 stations at distances of 10 to 100 km from the mainshock epicentre (Fig. 1). They include stations ISK and YLV operated by Kandilli Observatory and Earthquake Research Institute, stations UCG, KOS and GBZ operated by the Marmara Research Center, and stations OFL, YUT, DOK, TAS and ULD operated by Turkish Disaster Affairs in cooperation with GFZ-Potsdam. Stations ISK and YLV recorded continuous seismic data. All of the others recorded in triggered mode. All stations were recorded at 100 Hz sampling rate except ISK (20 Hz). Stations UCG and KOS have only a vertical component, the rest have three components. Stations observing specific events are identified in Supplementary Table 1.

The closest station to the earthquakes, UCG, analysed previously[16] recorded in triggered mode and captured many, but not all of the foreshocks (14 of 20 with $M \geq 1.0$). It is located 10 km north of the epicentre of the east–west striking and near-vertical North Anatolian Fault. For this geometry, arrival time differences between P and S waves change very slightly with position along the fault plane and consequently are insensitive to location differences within the plane of the vertical North Anatolian Fault. Seismograms from nearby positions on the fault plane would be expected to have highly similar waveforms in the bandwidth of the data for hypocentral separation of up to several hundred metres[37]. The additional stations are thus vital for determining the relative location of the foreshocks and the mainshock hypocentre. Recordings of the mainshock initiation are available from only six stations due to the disruption of the power grid by the earthquake. Nevertheless, they proved sufficient to accurately determine the locations of the mainshock hypocentre relative to the four largest foreshocks.

Earthquake locations. The four large foreshocks (23:18, 23:41, 23:49 and 23:59) and the mainshock (00:01) were located using double-difference method[38]. To eliminate any possibility of clock errors, we used differential S–P time measurements to locate the foreshocks. The mainshock location used only P wave arrival times. Differential travel times for the foreshocks are measured in two steps. In the first step, time domain cross-correlation is employed to determine time delays at precision of the sampling rate (100 s^{-1}). Cross-correlated time windows contain waveforms 0.64 s before and 1.28 s after P and S wave arrivals. In the second step, the phase of the cross spectrum of paired waveforms is used to measure differential times to subsample precision[39]. Note that the differential times are relative to the centroid time of each event, not its hypocentral time. Consequently, we determine the relative locations of the hypocentroids[17]. Differential travel times between the mainshock and the 23:59 foreshock were measure by correlating the on-scale portion of the mainshock P wave with the same interval of the foreshock P wave. Location uncertainties (2σ) come from the hypoDD program in singular value decomposition (SVD) mode[40] and were checked by a double-difference grid search scheme. This step was crucial especially for the mainshock location because it was recorded by a limited number of stations. Hypocentroid uncertainty for four foreshocks average 4 and 11 m in horizontal and vertical directions, respectively. For the mainshock, the 2σ hypocentre uncertainties are 16 and 49 m in horizontal and vertical directions, respectively.

Template matching. Station ISK is the only seismograph in the region with continuous recordings covering the foreshock period and 20 months before the mainshock. Continuous recordings at YLV were also available for the day before the mainshock. These data were searched using template matching to verify whether there are further foreshocks using the penultimate large foreshock 23:59. Eight additional foreshocks were found in the 44 minutes of previously identified foreshocks but none before the first foreshock seen at UCG at 23:17, 16 August 1999 (Supplementary Fig. 5).

Cluster analysis. Although precise locations could not be obtained for the smaller foreshocks, the association of the fourteen small foreshocks recorded at UCG with the four largest foreshocks could be determined using their seismograms (Fig. 4a). We analysed waveform similarity among the 18 foreshocks using SVD of the time-aligned waveforms. In the first step, SVD is used to remove the common mode signal from the seismograms (Fig. 2b). In the second step, residual seismograms are cross-correlated pairwise to generate a correlation-coefficient table (Fig. 2c). Finally, cluster analysis is performed to objectively sort the events into groups using cross correlation coefficients as the distance metric (Fig. 2d, Supplementary Table 1). Event 7 occurred 12 s after the M 2.0 23:41 event and was assigned to the same cluster.

Earthquake magnitudes. We used the amplitude of the common mode signal derived from the SVD analysis above to measure the relative seismic moments of each event[26]. When converted to moment magnitude, our values agree to 0.1 or better with those reported previously[16]. For consistency, we adopt those magnitudes in Supplementary Table 1 for earthquakes common to both studies.

Earthquake source parameters. The multi-window spectral ratio (MWSR) technique was used to determine the seismic moment, source dimension and stress drop of the four large foreshocks[41]. We used a previously reported source model[18] for calculation of source dimensions. MWSR uses small earthquakes as empirical Green's functions to suppress propagation effects including scattering and attenuation, site effects and sensor characteristics. The analysis was performed by fitting the Boatwright spectral model[42] to the spectral ratios (Supplementary Fig. 1). Source dimensions were derived from the shear wave corner frequency measurements (Supplementary Table 3) using two radially propagating circular source models[18,42,43] (Supplementary Table 2).

Code availability. Computer code used to analyse the data in this study is available from the corresponding author on request.

Data availability. The data that supports the findings of this study are available from the corresponding author on request. The original seismograms are available on request from the respective network operators identified in the 'Seismic waveforms' section.

References

37. Geller, R. & Mueller, C. Four similar earthquakes in Central California. *Geophys. Res. Lett.* **7**, 821–824 (1980).
38. Waldhauser, F. & Ellsworth, W. L. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bull. Seismol. Soc. Am.* **90**, 1353–1368 (2000).
39. Poupinet, G., Ellsworth, W. L. & Frechet, J. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras fault, California. *J. Geophys. Res.* **89**, 5719–5731 (1984).
40. Waldhauser, F. *hypoDD--A Program to Compute Double-Difference Hypocenter Locations (hypoDD version 1.0-03/2001)* Report 01-113 (US Geological Survey, 2001).
41. Imanishi, K. & Ellsworth, W. L. in *Earthquakes: Radiated Energy and the Physics of Faulting, 1* (eds Abercrombie, R. et al.) 81–90 (Geophysical Monograph Series 170, American Geophysical Union, Washington DC, 2013).
42. Boatwright, J. A spectral theory for circular seismic sources: simple estimates of source dimension, dynamic stress drop, and radiated seismic energy. *Bull. Seismol. Soc. Am.* **70**, 1–28 (1980).
43. Sato, T. & Hirasawa, T. Body wave spectra from propagating shear cracks. *J. Phys. Earth* **21**, 415–431 (1973).