GEOPHYSICS

Basal nucleation and the prevalence of ascending swarms in Long Valley caldera

Bing Q. Li[1,2]*, Jonathan D. Smith[1], Zachary E. Ross[1]

Earthquake swarms are ubiquitous in volcanic systems, being manifestations of underlying nontectonic processes such as magma intrusions or volatile fluid transport. The Long Valley caldera, California, is one such setting where episodic earthquake swarms and persistent uplift suggest the presence of active magmatism. We quantify the long-term spatial and temporal characteristics of seismicity in the region using cluster analysis on a 25-year high-resolution earthquake catalog derived using leading-edge deep-learning algorithms. Our results show that earthquake swarms beneath the caldera exhibit enlarged families with statistically significant tendency for upward migration patterns. The ascending swarms tend to nucleate at the base of the seismogenic zone with a spatial footprint that is laterally constrained by the southern rim of the caldera. We suggest that these swarms are driven by the transport of volatile-rich fluids released from deep volcanic processes. The observations highlight the potential for extreme spatial segmentation of earthquake triggering processes in magmatic systems.

INTRODUCTION

Earthquake swarms are a pervasive yet mysterious feature of magmatic systems, with a variety of nontectonic driving mechanisms including inflation of magma bodies (1, 2), caldera collapse (3), dike intrusions (4), and degassing processes (5). In some instances, swarms have served as precursors to volcanic eruptions (6, 7), while in other circumstances they are seemingly unrelated (8, 9). Zobin (10) gives a detailed outline of case studies from many magmatic swarming regions. For example, there appears to be notable depth segmentation between deep (10 to 20 km) and shallow events, often separated by an aseismic zone corresponding to viscous material within or near magma chambers. This depth segmentation is also evident as a temporal segmentation, where the initial stages of eruptive seismicity and steady-state seismicity include both deep and shallow events, whereas events tend to be clustered exclusively at shallow depths immediately before eruptions (11–15). Other studies focusing on the waveform similarity of earthquakes in volcanic settings find that noneruptive events tend to exhibit more similarity and temporal regularity than events before and during eruptions, which may be interpreted as eruptions presenting a substantial change in geologic and stress conditions (16–18). The early work of Benoit and McNutt (8) constructed a database of volcanic earthquake swarms for the period of 1979–1989, including 191 swarms preceding eruption activity from which they determined that the duration of the swarm is correlated with the time between eruptive episodes, where longer swarm activity leads to a long interevent time of eruptions. On the other hand, noneruptive earthquake swarms found in a number of settings (19–21) have been attributed to underlying mechanisms such as arrested magma intrusions, snowmelt, and deep hydrothermal activity. In particular, the Yellowstone caldera has exhibited a number of earthquake swarms, which have been linked to fluid migration, as determined by elevated b values near the resurgent dome (22), multiplet analysis (23), as well as spatiotemporal evolution and focal mechanisms (24, 25). The Long Valley caldera is an active volcanic system in California that has produced several notable noneruptive volcanic swarms. The most prominent of these began in 1978, where a period of relative quiescence spanning several decades was interrupted by a large earthquake swarm on the south rim, along with substantial uplift in the center of the caldera that continues episodically to the present day (26, 27). The renewed seismicity and uplift have raised concerns of an impending eruption due to the inflation of the magma chamber (26); however, some studies have shown that the seismicity in the caldera is consistent with regional-scale tectonic processes, being on the eastern side of the Sierra Nevada mountain range (28, 27). Geochemical studies suggest that no new magma is intruding into the system (29), while others have shown that earthquake swarms are triggered by low-viscosity, high-pressure hydrothermal fluids originating from an ancient crystallizing pluton underlying the caldera (30). The latter hypothesis is evidenced by a high-resolution seismicity analysis of two prominent swarms that occurred in 2014, which showed strongly fault-bounded seismicity with rapid migration rates suggesting transport of a low-viscosity fluid (31). Nevertheless, the question regarding the mechanism behind uplift remains open, given that fluid-driven mechanisms should exhibit episodes of subsidence as fluids escape the surface (32). The debate highlights the heterogeneous and enigmatic nature of volcano-tectonic systems and raises questions regarding how indirectly measured features such as seismicity may be interpreted in the context of the underlying geological processes.

To better understand the origins of seismicity in the Long Valley caldera, we reprocessed nearly 25 years of continuous seismic data, with deep-learning algorithms for earthquake monitoring (33–36). We created a high-resolution seismicity catalog and subsequently performed a cluster analysis to comprehensively quantify the spatial and temporal dynamics of seismicity within and outside of the caldera.

RESULTS

Overview of seismicity features

We present an earthquake catalog comprising 260,312 events spanning 1995–2019, including 169,638 events that were relocated

[1]Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA. [2]Department of Civil and Environmental Engineering, Western University, London, ON, Canada.
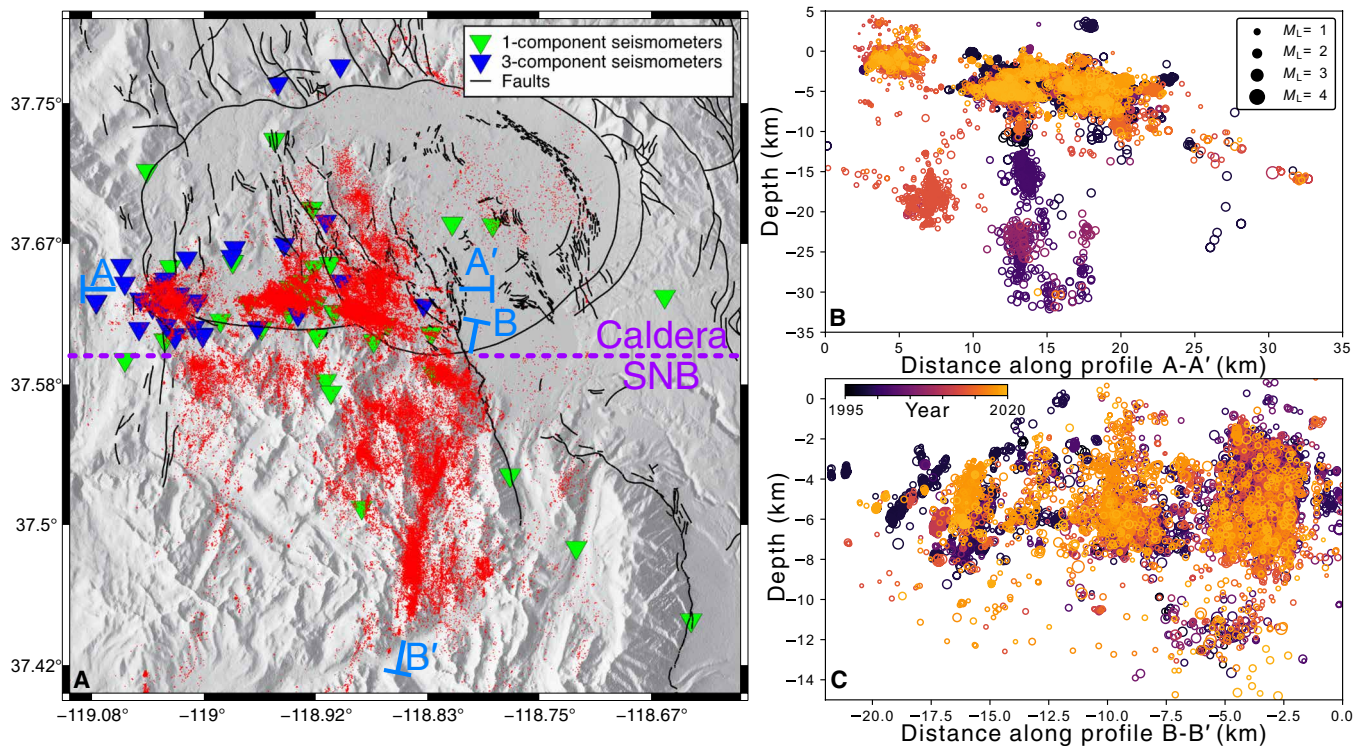*Corresponding author. Email: bing.li@uwo.ca

Fig. 1. Map and cross-section views of relocated seismicity. (**A**) Map view of all relocated seismicity. Light blue bars indicate locations of depth cross sections A-A′ (5-km projection distance) and B-B′ (1-km projection distance), and purple dashed line indicates the dividing line between the caldera and the Sierra Nevada block (SNB). Black lines denote locations of known faults, including the caldera rim. One- and three-component seismometer locations are indicated in inverted triangles. (**B**) Depth cross-section A-A′ along the southern rim of the caldera. Note the base of the shallow seismogenic zone dips gently from approximately 5 km in the west to 10 km in the east. (**C**) Depth cross-section B-B′ along the seismicity in the SNB. Depths are referenced to sea level.

with waveform cross-correlation methods (Fig. 1). The map view of seismicity demonstrates a north-south striking of the faults in the Sierra Nevada block (SNB), which transitions to east-west and northwest-southeast striking features in the caldera. The depth cross-section A-A′ indicates that the bulk of the seismicity occurs in the top 5 to 10 km, with the base of this seismogenic zone sharply delineated by a subhorizontal seismicity surface. This transition gently dips from a depth of approximately 5 km in the west to 10 km in the east, and the B-B′ depth cross section suggests that the boundary is flat lying in the north-south direction. This seismicity structure has been interpreted as being a thermally controlled brittle-ductile transition (*37*) overlying an ancient pluton consisting of crystalline mush (*30*). The A-A′ depth cross section also shows isolated clusters of earthquakes at depths greater than 15 km, which may be related to dike intrusions or fluid pressure pulses (*38*).

Cluster dynamics

To study the spatiotemporal behavior of seismicity over the 25-year period, we divide the earthquake catalog into clusters using a nearest-neighbor approach (*39*). This method computes a space-time distance between pairs of events and identifies the nearest neighbor (parent) for each event, resulting in a single tree structure spanning the entire catalog. The distribution of nearest-neighbor distances (NNDs) in tectonic regions for large space-time windows commonly exhibits two modes: one that occurs at short space-time distances and has been interpreted as due to clustering processes such as aftershocks or aseismic forcing, and the other that has roughly a

Weibull distribution and is consistent with a Poisson process (*40*). An example is shown for the entire southern California region in Fig. 2A, where the distance is decomposed into spatial and temporal components (see Materials and Methods). For our study area, we separate events within the caldera (Fig. 2C) from those immediately outside to the south (Fig. 2B). The NND distribution for events in the latter region exhibits the two-mode pattern similar to the southern California results, corresponding to Poissonian and clustered behavior (*41*). However, the NND distribution within the caldera is markedly different and is nearly unimodal; the fraction of all earthquakes that can be viewed as triggered (not background) is an astonishing 68% compared to 34% in adjacent SNB. This indicates that the caldera earthquakes predominantly occur within sequences, in contrast to the seismicity in the SNB, which has many more background events relative to the total number that occurred. To better understand these observations, we next examine seismicity behavior of earthquake clusters.

The NND diagrams provide a simple means to extract seismicity clusters by breaking the links of the spanning tree whenever the NND exceeds some threshold (see Materials and Methods). Applying this procedure results in 10,750 clusters within the caldera and 21,096 clusters in the SNB. The large number of clusters allows detailed examination of the cluster patterns between the two regions. Figure 3 shows cumulative distributions for the number of events in a cluster. The clusters in the caldera are seen to be statistically larger than those clusters immediately to the south. This observation is consistent with the general findings of Zaliapin and Ben-Zion (*42*),
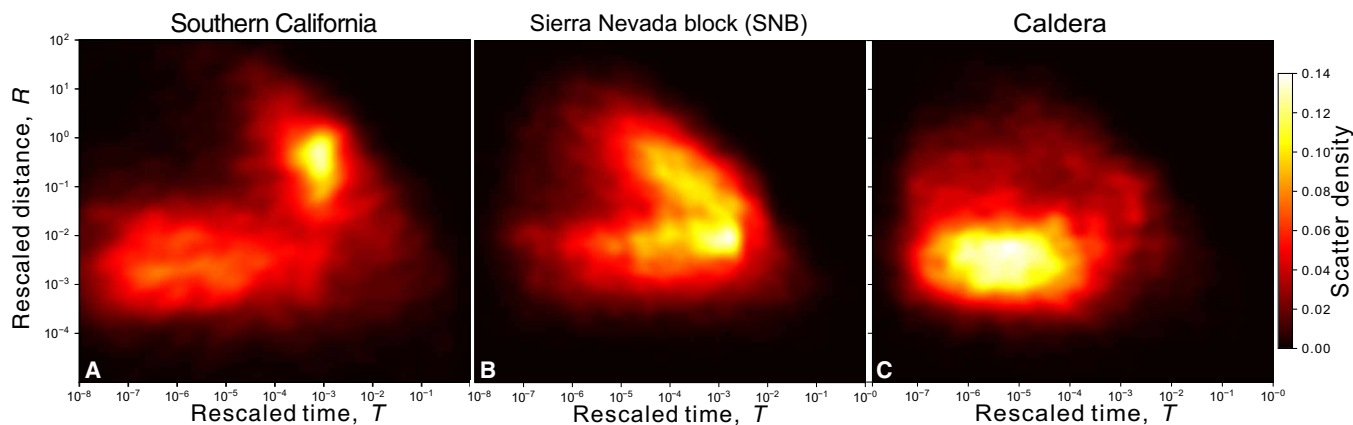
Fig. 2. Normalized spatial and temporal distribution of NND η shown in color. Presented for (**A**) southern California (SCSN), $M_c = 3$, $n = 13{,}044$; (**B**) the SNB, $n = 34{,}649$; and (**C**) the caldera, $n = 31{,}574$.
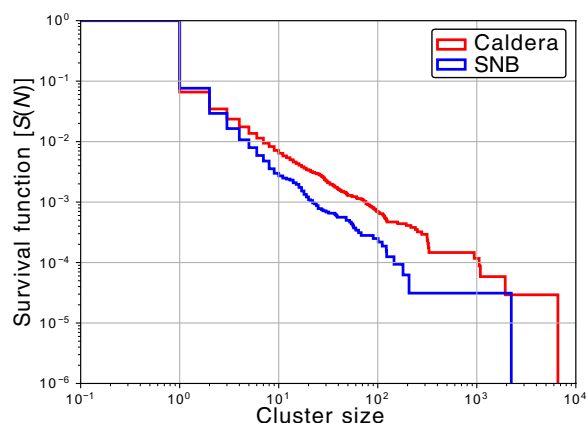

Fig. 3. Survival function for cluster size. Presented for clusters in (red) the caldera and (blue) the SNB.

who noted that seismicity clusters in areas with elevated heat flow were statistically larger than those in areas with lower heat flow.

The tree structures associated with a set of clusters have been used to study topological aspects of seismicity. Using the term leaf to describe an event with no offspring, we next focus on a quantity called the leaf depth, d, which is defined as the number of generations separating a given leaf and the root of the tree. The average leaf depth over the tree has been shown to be useful for distinguishing swarm-type clusters from more typical mainshock-aftershock clusters (*42*), with swarms tending to exhibit larger values, indicating that each event tends to trigger few offspring, resulting in long chain–like tree structures. Mainshock-aftershock sequences tend to have a small number of events that trigger most of the events in the sequence, resulting in trees with few generations. Figure 4A shows a map of seismicity clusters that have mean leaf depth $d \geq 5$, while Fig. 4B separately indicates clusters with $d < 5$. More than 95% of the clusters with $d \geq 5$ are located within the caldera directly; however, the number of clusters with $d < 5$ is similar between the caldera and the adjacent SNB. Thus, the spatial distribution of swarm-like clusters terminates abruptly at the southern rim of the caldera.

The question of whether swarms tend to migrate upward or downward is important for understanding the physical processes and conditions underlying swarm evolution. Basic considerations

of the diffusion problem lead to the expectation that swarms should be upward along a hydraulic gradient [e.g., (*43*)]. The handful of well-documented swarms in the Long Valley caldera and SNB, however, have exhibited both upward (*38*, *31*) and downward (*44*) migration patterns; the limited number of examples here is too small to make any kind of general statements about swarm behavior in the caldera. Given the large number of clusters in our dataset, we can test whether there is a statistical tendency for swarms to migrate upward. We define a vertical mean migration distance $\Delta \overline{Z} = Z_{root} - \overline{Z}$ for each cluster, where Z_{root} is the depth of the first (root) event in the cluster, $\overline{Z}$ is the cluster mean depth, and positive $\Delta \overline{Z}$ distances indicate ascending migration. To determine whether clusters systematically exhibit ascending migration within and outside of the caldera, we conduct one-sided t tests with the null hypothesis that clusters in the caldera and the SNB exhibit $\Delta \overline{Z} = 0$, i.e., that clusters on average do not exhibit ascending migration. The results in Table 1 show that we cannot reject the null hypothesis for clusters in the SNB. In the caldera, we can reject the null hypothesis at 10% significance in the caldera, indicating that clusters in the caldera have statistically significant tendency to migrate upward in the 25-year study period. We do note that the histogram of $\Delta \overline{Z}$ shown in fig. S6 indicates that the difference between the caldera and the SNB is somewhat subtle. The mean migration distance for the caldera is 85 m and 13 m for the SNB. The SEs are 49 m for the caldera and 115 m for the SNB. We find that of the caldera clusters, 36% exhibit ascending migration greater than 100 m, while 23% exhibit descending migration greater than 100 m. Similarly, the proportions in the SNB are 45% ascending >100 m, 33% descending >100 m. Given the location errors, it is difficult to resolve these patterns for individual clusters; however, the one-sided t test shows that the migrations are systematically positive in the caldera.

Given the finding that the caldera swarms are statistically likely to migrate upward, we can examine the spatial distribution of the clusters exhibiting strong migration. Figure 5A shows the nucleation site of clusters with $\Delta \overline{Z} \geq 0.5$ km, while Fig. 5B separately shows clusters with $\Delta \overline{Z} < 0.5$ km. The depth cross sections (Fig. 5, B and D) suggest that the ascending clusters in the caldera nucleate near the base of the seismogenic zone, whereas non-ascending clusters are distributed across the entire seismogenic zone. This can again be systematically investigated using a one-sided t test, with the null hypothesis that ascending clusters ($\Delta \overline{Z} \geq 0.5$ km) and non-ascending
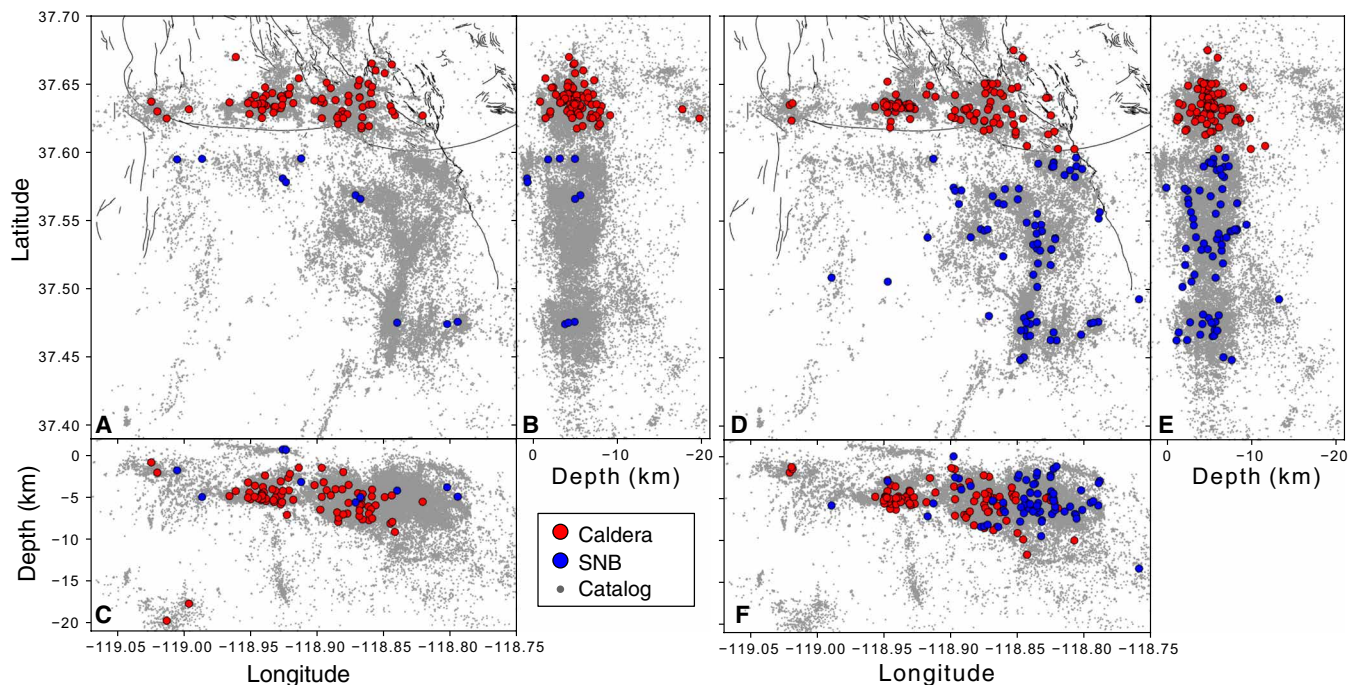
Fig. 4. Spatial distribution of clusters with large ($d \geq 5$) and small ($1 \leq d < 5$) mean leaf depth. (**A** to **C**) Map and north-south and east-west views of large leaf depth root event locations. (**D** to **F**) Small leaf depth root event locations.

Table 1. One-sided t test for the expected value of migration distance $\Delta \bar{Z}$ in the caldera and SNB.

Null hypothesis H_0	Sample size N	t statistic	P value	Decision
$\Delta \bar{Z}_{caldera} = 0$	159	1.97	0.0255	Reject at 10% significance
$\Delta \bar{Z}_{SNB} = 0$	78	0.0797	0.468	Accept

clusters ($\Delta \bar{Z} < 0.5$ km) have the same global mean depth (averaged over all clusters within each group). The null hypothesis is rejected with 10% significance when comparing ascending clusters in the caldera to non-ascending clusters in the caldera, as well as comparing ascending clusters in the caldera to all other clusters in our catalog (Table 2). This indicates that ascending clusters in the caldera nucleate deeper on average than the non-ascending clusters. Examples of the spatiotemporal evolution of ascending clusters are shown in Fig. 5 (E and F), and an example of a descending cluster is shown in Fig. 5G.

DISCUSSION

Our results show that seismicity in the Long Valley caldera includes a substantial fraction of episodic swarm-type events characterized by short normalized interevent times and distances, relatively large cluster sizes, and large leaf depths denoting a long chain–like topology, where each event is linked to only a small number of subsequent events (Figs. 2C and 6B). These swarm-type clusters, which statistically exhibit a tendency for ascending migration patterns, are markedly distinct from typical tectonic clusters such as in the adjacent SNB, where individual mainshocks trigger large families of aftershocks

with substantial branching (e.g., each event directly triggers many aftershocks) (Figs. 2B and 6C). In addition, the ascending swarm-type clusters occur almost exclusively within the caldera (Figs. 4A and 5A), with a sharp demarcation at the caldera rim suggesting that these swarms are specifically triggered by volcanic processes underlying the caldera (Fig. 6A). The observations further provide a line of evidence that the caldera rim is a key boundary for these volcanic processes, which may not extend beyond it.

Previous geological (*30*), geochemical (*29*), and seismological studies (*31*) provide evidence that these volcanic swarms are triggered by the transport of hydrothermal fluids originating from below the seismogenic zone, which are produced as part of the degassing of an ancient caldera-wide pluton. These hydrothermal fluids could account for the uplift seen at the resurgent dome and may be relieved by transport along steeply dipping ring faults along the south rim of the caldera. These ring faults were likely formed or activated in response to the complex tectonic stresses on the edge of the Sierra Nevada mountains (*28*, *45*), leading to elevated shear stresses that trigger tectonic-type double-couple focal mechanisms seen alongside non–double-couple mechanisms in the caldera (*46*). This elevated shear stress is also seen in our results, where the clusters in the caldera, although dominated by the short normalized interevent
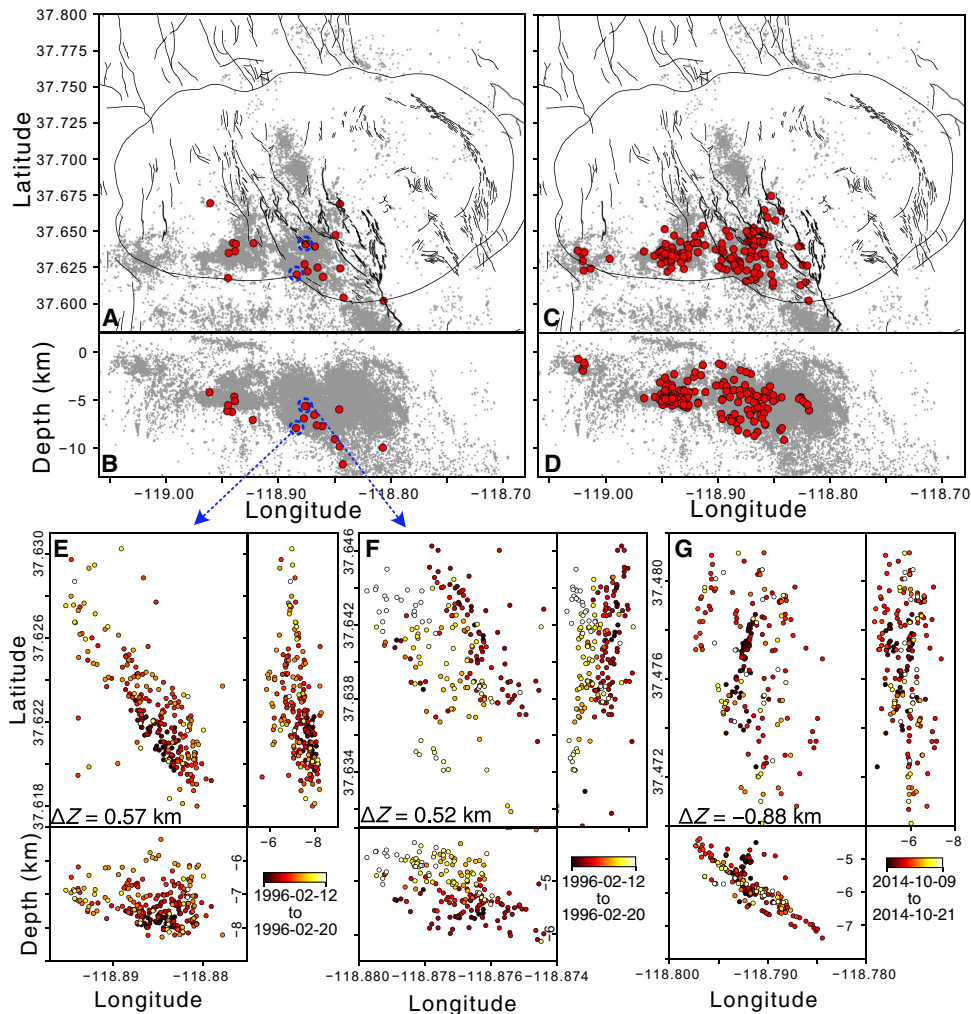
Fig. 5. Spatial distribution of clusters with large ($\Delta \overline{Z} > 0.5$ km) and small ($\Delta \overline{Z} < 0.5$ km) mean migration distance. (**A** and **B**) Map and east-west views of large migration distance root event locations. Blue circles and arrows show nucleation points of clusters shown in (E) and (F). (**C** and **D**) Small migration distance root event locations. Note that while the nonmigrating clusters occur throughout the seismogenic zone, the upward migrating clusters occur exclusively near its base. (**E** and **F**) Examples of ascending clusters in plan and section views, where color indicates time. (**G**) Example of a descending cluster in plan and section views, where color indicates time.

Table 2. One-sided *t* test for the mean depth of ascending ($\Delta \overline{Z} > 0.5$ km) and non-ascending ($\Delta \overline{Z} < 0.5$ km) clusters. Sample size column indicates *N* for the two populations considered. $\overline{Z}_{\text{remaining}}$ refers to the mean depth of the population consisting of SNB clusters and non-ascending caldera clusters.

Null hypothesis H_0	Sample size N	t statistic	P value	Decision
$\overline{Z}_{\text{caldera,ascending}} = \overline{Z}_{\text{caldera,non-ascending}}$	18/141	1.43	0.0778	Reject at 10% significance
$\overline{Z}_{\text{SNB,ascending}} = \overline{Z}_{\text{SNB,non-ascending}}$	16/62	0.758	0.226	Accept
$\overline{Z}_{\text{caldera,ascending}} = \overline{Z}_{\text{remaining}}$	18/219	1.75	0.0409	Reject at 10% significance

times and distances, nevertheless contain a number of clusters with shorter leaf depths and non-ascending behavior (Fig. 4D). The hypothesis of combined fluid-driven and tectonic stress changes may be further supported by the general lack of seismicity east of the resurgent dome (Fig. 1A), which is more distal from the stress changes induced by the Hilton Creek fault and thus less likely to

form or activate faults. As a result, the resurgent dome likely consists of relatively intact material forming a central piston-type block similar to that observed at Kilauea (*47*).

We further show that swarms exhibiting large positive migration distances (ascending, $\Delta \overline{Z} > 0.5$ km) occur at greater depths than other earthquake clusters in Long Valley (Fig. 5). The root events
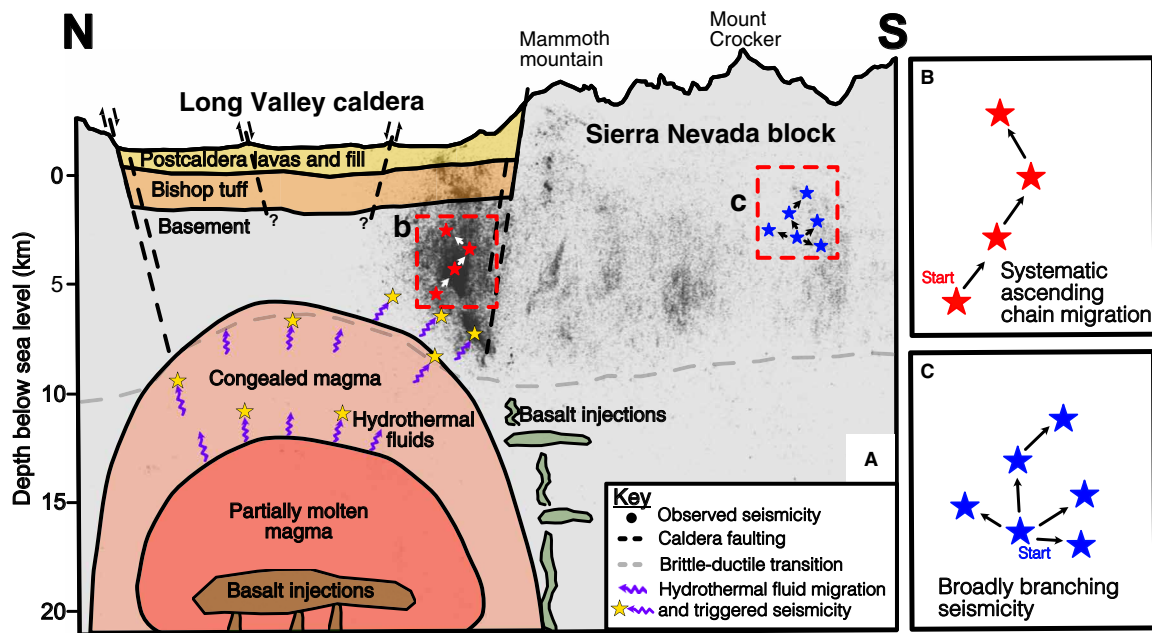
Fig. 6. Schematic north-south depth cross section. (**A**) Earthquake locations in black circles, inferred fault lines in dashed black lines, and inferred brittle-ductile transition in dashed gray lines. Black lines denote inferred geological boundaries. (**B**) Topological representation of long chain–like earthquake clustering commonly found in the caldera. (**C**) Topological representation of tectonic-type clustering found commonly in the adjacent SNB.

for these ascending clusters are evenly spaced across the south rim of the caldera but are not present at Mammoth mountain on the southwest corner of the caldera, suggesting that these root event locations exclusively represent nucleation sites of hydrothermal fluid conduits originating from the ancient crystal mush. On the other hand, the seismicity under Mammoth mountain may be attributed to a different triggering mechanism, as suggested by the presence of large and small leaf depth clusters (Fig. 4) while lacking upward migrating swarms (Fig. 5A). Here, the seismicity may be more closely related to changes in an active magma chamber (38, 48), which is localized to the active volcanism at Mammoth and the Mono-Inyo chain extending to the north.

Seismicity patterns globally have long been observed to exhibit extraordinary variability in their behavior, even on relatively short length scales; however, historically, it was unclear how much of this complexity was due to observational error and/or the non-uniqueness associated with geophysical inverse problems. The recent improvements to seismicity catalogs resulting from increased detection sensitivity (49–51) and picking accuracy (52, 53) have repeatedly shown that such complexity appears to be genuine, especially for swarms (3, 43, 54). The highest-resolution observations of recent years have further highlighted the importance of spatially variable fault zone properties and processes in underlying such dynamic behavior. Here, we demonstrate that these enhanced catalogs, which can be readily scaled to long study periods owing to its underlying automated algorithms, can be robustly probed for long-term patterns using statistical techniques. Our results, which broadly characterize fluid-triggered and regional tectonic seismicity under complex thermal-hydro-mechanical conditions, provide a reference point for regions with sparse network coverage and/or shorter time scales. Thus, our results represent a substantial step forward in our understanding of subsurface dynamics in magmatic regions.

MATERIALS AND METHODS
Earthquake detection and location
Raw waveform data were downloaded from between 1995 and 2019 for stations within a box surrounding the Long Valley caldera (latitude, 37.2 °N to 38.2 °N; longitude, 119.4 °W to 118.4 °W), including the Incorporated Research Institutions for Seismology (IRIS), Northern California Earthquake Data Center (NCEDC), and Southern California Earthquake Data Center (SCEDC) data centers. The three-component waveform data were processed for P- and S-wave arrivals using the generalized phase detection (GPD) (33) method, which was trained using the NCEDC catalog picks. The one-component waveform data were similarly processed for phase picks using a modified GPD algorithm, which was trained with NCEDC catalog picks, where the training data consisted of 50% noise waveforms for regularization to minimize the number of false-positive phase picks.

Phase picks were then associated with the PhaseLink algorithm (34), where the model weights were first trained on a synthetic catalog of earthquake locations. This model was used to create an initial catalog of real earthquakes in the Long Valley region spanning 1995–2019, and then the initial catalog was used to improve the PhaseLink model weights using transfer learning. The events were initially located with the NonLinLoc package (35) using a three-dimensional S-wave velocity model (55) and $V_P/V_S = 1.68$, and then relocated using GrowClust (36). The final catalog consists of 260,312 events, of which 169,638 were relocated. Magnitudes were calculated using a local magnitude scale (56) and corrected to a duration magnitude scale using an empirical calibration suggested by the U.S. Geologic Survey (figs. S1A and S2) (31). Newly detected events with $M > 2$ are not assigned a magnitude in the catalog to avoid including events with clipped waveforms. The magnitude of completeness is assessed with guidance from the maximum

curvature method (*57*) using a 2000 event rolling window; we find $M_c = 1.0$ before 2011 and $M_c = 0.5$ after 2011 (fig. S3). We classify earthquakes with a latitude greater than 37.6° as originating from the caldera and latitude less than or equal to 37.6° as originating from the SNB (*28*). The regional catalog from the NCEDC covering the same period and spatial extents contains 131,912 events, of which 112,469 are also detected within our enhanced catalog totaling 255,650 events. Of the events common to the NCEDC and our catalog, 80% of epicentral locations were within 2.13 km, and 80% of depths were within 2.45 km (fig. S1B).

Clustering

The earthquakes are divided into clusters in the space-time-energy domain using the NND formulation proposed by Zaliapin and Ben-Zion (*39*), where each event *j* is assigned a parent event *i* separated by NND $\eta = R_{ij}T_{ij}$, where normalized distance R is defined as $R_{ij} = (r_{ij})^{d_f} 10^{-(1-q)bm_i}$, normalized time T is defined as $T_{ij} = t_{ij}10^{-qbm_i}$, r_{ij} is the hypocentral distance in kilometers, t_{ij} is the time between events in years, m_i is the magnitude, $d_f = 1.6$ is a fractal dimension of the earthquake hypocenter distribution, $b = 1$ is the slope of the Gutenberg-Richter frequency-magnitude distribution, and $q = 0.5$ is the relative weighting of the normalized time and distance. The method has been successfully applied to large earthquake catalogs spanning 111,981 events across southern California (*39*), as well as regional scale–induced earthquake sequences such as a catalog with around 8000 events at the Coso geothermal fields (*58*).

Within this framework, an earthquake catalog is assumed to contain two distinct populations consisting of (i) parent events demarked by large NND and (ii) child events at small NND. Given a population of η for the entire catalog, we then use a Gaussian mixture model to fit a bimodal distribution to η and determine $\eta_c = -6.2$ (fig. S4) as the best-fit parameter to separate the two populations. We then formulate clusters by removing all links larger than η_c such that any event with $\eta > \eta_c$ is considered the first, or root, event in a cluster. Any child events of the root event, and their subsequent children, are considered within this cluster. The leaf is then defined as an event without a child, with leaf depth defined as the number of links between the leaf and the root event. The branching ratio is defined as the number of children linked to a parent event.

Within each cluster, outliers are identified and removed if the event depth exceeds the inner fence and is more than 2.5 km from the median depth of the cluster. The upper and lower inner fences are defined as $Q3 + 1.5 * (Q3 - Q1)$ and $Q1 - 1.5 * (Q3 - Q1)$, respectively, where $Q1$ and $Q3$ are the 25th and 75th percentiles, respectively. Of the 65,656 events in the clustered catalog, 573 were determined to be outliers and removed. The survival function is calculated as $S(N) = 1 - \text{cdf}(N)$, where $\text{cdf}(N)$ is the empirical cumulative density function of the population of cluster family sizes.

Migration distance

We define the migration statistic of each cluster as the parameter $\Delta \overline{Z} = Z_{root} - \overline{Z}$, where Z_{root} is the depth of the first event in a cluster, and $\overline{Z}$ is the mean depth of the cluster. Because depth is positive downward, a positive value of $\Delta \overline{Z}$ indicates upward migration of the cluster. We test whether clusters in (i) the caldera and (ii) the SNB exhibit upward migration using a one-sided *t* test, with the null hypothesis that the mean migration is 0.

We further define clusters with $\Delta \overline{Z} > 0.5$ km as ascending clusters and additionally test the hypothesis that the ascending clusters

systematically nucleate deeper than non-ascending clusters, again using a one-sided *t* test with the null hypothesis that the mean depths of ascending clusters are statistically indistinguishable from the mean depths of non-ascending clusters. This test is used to compare (i) ascending with non-ascending clusters in the caldera, (ii) ascending with non-ascending clusters in the SNB, and (iii) ascending clusters in the caldera with all other clusters.

REFERENCES AND NOTES

1. S. Cesca, J. Letort, H. N. T. Razafindrakoto, S. Heimann, E. Rivalta, M. P. Isken, M. Nikkhoo, L. Passarelli, G. M. Petersen, F. Cotton, T. Dahm, Drainage of a deep magma reservoir near Mayotte inferred from seismicity and deformation. *Nat. Geosci.* **13**, 87–93 (2020).
2. M. T. Gudmundsson, K. Jónsdóttir, A. Hooper, E. P. Holohan, S. A. Halldórsson, B. G. Ófeigsson, S. Cesca, K. S. Vogfjörd, F. Sigmundsson, T. Högnadóttir, P. Einarsson, O. Sigmarsson, A. H. Jarosch, K. Jónasson, E. Magnússon, S. Hreinsdóttir, M. Bagnardi, M. M. Parks, V. Hjörleifsdóttir, F. Pálsson, T. R. Walter, M. P. J. Schöpfer, S. Heimann, H. I. Reynolds, S. Dumont, E. Bali, G. H. Gudfinnsson, T. Dahm, M. J. Roberts, M. Hensch, J. M. C. Belart, K. Spaans, S. Jakobsson, G. B. Gudmundsson, H. M. Fridriksdóttir, V. Drouin, T. Dürig, G. Aðalgeirsdóttir, M. S. Riishuus, G. B. M. Pedersen, T. van Boeckel, B. Oddsson, M. A. Pfeffer, S. Barsotti, B. Bergsson, A. Donovan, M. R. Burton, A. Aiuppa, Gradual caldera collapse at Bárdarbunga volcano, Iceland, regulated by lateral magma outflow. *Science* **15**, aaf8988 (2016).
3. D. R. Shelly, W. A. Thelen, Anatomy of a caldera collapse: Kīlauea 2018 summit seismicity sequence in high resolution. *Geophys. Res. Lett.* **46**, 14395–14403 (2019).
4. J. Woods, C. Donaldson, R. S. White, C. Caudron, B. Brandsdóttir, T. S. Hudson, T. Ágústsdóttir, Long-period seismicity reveals magma pathways above a laterally propagating dyke during the 2014-15 Bárðarbunga rifting event, Iceland. *Earth Planet. Sci. Lett.* **490**, 216–229 (2018).
5. S. Abramenkov, N. M. Shapiro, I. Koulakov, I. Abkadyrov, Clustering of long-period earthquakes beneath Gorely Volcano (Kamchatka) during a degassing episode in 2013. *Geosciences* **10**, 230 (2020).
6. S. D. Malone, C. Boyko, C. S. Weaver, Seismic precursors to the mount St. Helens eruptions in 1981 and 1982. *Science* **221**, 1376–1378 (1983).
7. S. F. M. Chastin, I. G. Main, Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions. *Geophys. Res. Lett.* **30**, 1671 (2003).
8. J. P. Benoit, S. R. McNutt, Global volcanic earthquake swarm database and preliminary analysis of volcanic earthquake swarm duration. *Annals Geophys.* **39**, 221 (1996).
9. R. White, W. McCausland, Volcano-tectonic earthquakes: A new tool for estimating intrusive volumes and forecasting eruptions. *J. Volcanol. Geotherm. Res.* **309**, 139–155 (2016).
10. V. Zobin, *Introduction to Volcanic Seismology* (Elsevier, ed. 3, 2016), vol. 6.
11. W. Thelen, M. West, S. Senyukov, Seismic characterization of the fall 2007 eruptive sequence at Bezymianny volcano, Russia. *J. Volcanol. Geotherm. Res.* **194**, 201–213 (2010).
12. D. Dzurisin, J. W. Vallance, T. M. Gerlach, S. C. Moran, S. D. Malone, Mount St. Helens reawakens. *Eos* **86**, 25 (2005).
13. Z. Jiménez, V. H. Espíndola, J. M. Espíndola, Evolution of the seismic activity from the 1982 eruption of El Chichon volcano, Chiapas, Mexico. *Bull. Volcanol.* **61**, 411 (1999).
14. F. W. Klein, P. Einarsson, M. Wyss, The Reykjanes peninsula, Iceland, earthquake swarm of September 1972 and its tectonic significance. *J. Geophys. Res. Solid Earth* **82**, 865–888 (1977).
15. H.-C. Pu, C. H. Lin, Y. C. Huang, L. C. Chang, H. F. Lee, P. L. Leu, N. C. Hsiao, C. H. Chang, Y. H. Yeh, The volcanic earthquake swarm of October 20, 2009 in the Tatun area of northern Taiwan. *Terrestrial Atmos. Oceanic Sci.* **25**, 625 (2014).
16. W. Thelen, S. Malone, M. West, Multiplets: Their behavior and utility at dacitic and andesitic volcanic centers. *J. Geophys. Res. Solid Earth* **116**, B08210 (2011).
17. A. D. Miller, R. C. Stewart, R. A. White, R. Luckett, B. J. Baptie, W. P. Aspinall, J. L. Latchman, L. L. Lynch, B. Voight, Seismicity associated with dome growth and collapse at the Soufriere hills volcano, Montserrat. *Geophys. Res. Lett.* **25**, 3401–3404 (1998).
18. K. Umakoshi, N. Itasaka, H. Shimizu, High-frequency earthquake swarm associated with the May 1991 dome extrusion at Unzen volcano, Japan. *J. Volcanol. Geotherm. Res.* **206**, 70–79 (2011).
19. D. C. Roman, J. A. Power, Mechanism of the 1996-97 non-eruptive volcano-tectonic earthquake swarm at Iliamna volcano, Alaska. *Bull. Volcanol.* **73**, 143–153 (2011).
20. S. K. Ebmeier, J. R. Elliott, J. M. Nocquet, J. Biggs, P. Mothes, P. Jarrín, M. Yépez, S. Aguaiza, P. Lundgren, S. V. Samsonov, Shallow earthquake inhibits unrest near Chiles-Cerro negro volcanoes, Ecuador-Colombian border. *Earth Planet. Sci. Lett.* **450**, 283–291 (2016).

21. I. Park, A. Jolly, K. Y. Kim, B. Kennedy, Temporal variations of repeating low frequency volcanic earthquakes at Ngauruhoe volcano, New Zealand. *J. Volcanol. Geotherm. Res.* **373**, 108–119 (2019).

22. J. Farrell, S. Husen, R. B. Smith, Earthquake swarm and b-value characterization of the Yellowstone volcano-tectonic system. *J. Volcanol. Geotherm. Res.* **188**, 260–276 (2009).

23. F. Massin, J. Farrell, R. B. Smith, Repeating earthquakes in the Yellowstone volcanic field: Implications for rupture dynamics, ground deformation, and migration in earthquake swarms. *J. Volcanol. Geotherm. Res.* **257**, 159–173 (2013).

24. G. P. Waite, R. B. Smith, Seismic evidence for fluid migration accompanying subsidence of the Yellowstone caldera. *J. Geophys. Res.* **107**, 2177 (2002).

25. E. Russo, A. Tibaldi, G. P. Waite, F. L. Bonali, F. Massin, J. Farrell, Unraveling the complex deformation pattern at Yellowstone plateau through seismicity and fracture analysis. *Tectonophysics* **778**, 228352 (2020).

26. M. Battaglia, C. Roberts, P. Segall, Magma intrusion beneath Long Calley caldera confirmed by temporal changes in gravity. *Science* **285**, 2119–2122 (1999).

27. E. K. Montgomery-Brown, C. W. Wicks, P. F. Cervelli, J. O. Langbein, J. L. Svarc, D. R. Shelly, D. P. Hill, M. Lisowski, Renewed inflation of Long Valley Caldera, California (2011 to 2014). *Geophys. Res. Lett.* **42**, 5250–5257 (2015).

28. S. Prejean, W. Ellsworth, M. Zoback, F. Waldhauser, Fault structure and kinematics of the Long Valley Caldera region, California, revealed by high-accuracy earthquake hypocenters and focal mechanism stress inversions. *Bull. Seismol. Soc. Am.* **107**, 1–19 (2002).

29. G. Suemnicht, B. Kennedy, W. Evans, Helium isotope systematics of Long Valley Caldera, in *Proceedings World Geothermal Congress 2015* (International Geothermal Association, 2015), vol. 1, pp. 1–6.

30. W. Hildreth, Fluid-driven uplift at Long Valley Caldera, California: Geologic perspectives. *J. Volcanol. Geotherm. Res.* **341**, 269–286 (2017).

31. D. R. Shelly, W. L. Ellsworth, D. P. Hill, 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).

32. D. P. Hill, E. Montgomery-Brown, D. R. Shelly, A. Flinders, S. Prejean, Post-1978 tumescence at Long Valley caldera, California: A geophysical perspective. *J. Volcanol. Geotherm. Res.* **400**, 106900 (2020).

33. Z. E. Ross, M.-A. Meier, E. Hauksson, T. H. Heaton, Generalized seismic phase detection with deep learning. *Bull. Seismol. Soc. Am.* **108**, 2894–2901 (2018).

34. Z. E. Ross, Y. Yue, M.-A. Meier, E. Hauksson, T. H. Heaton, PhaseLink: A deep learning approach to seismic phase association. *J. Geophys. Res. Solid Earth* **124**, 856–869 (2019).

35. A. Lomax, J. Virieux, P. Volant, C. Berge-Thierry, Probabilistic earthquake location in 3D and layered models, in *Advances in Seismic Event Location*, C. Thurber, N. Rabinowitz, Eds. (Kluwer, 2000), vol. 18, pp. 101–134.

36. D. Trugman, P. M. Shearer, GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish springs and Sheldon, Nevada, earthquake sequences. *Seismol. Res. Lett.* **88**, 379–391 (2017).

37. D. Hill, Temperatures at the base of the seismogenic crust beneath Long Valley caldera, California, and the Phlegrean Fields caldera, Italy, in *Volcanic Seismology*, P. Gasparini, R. Scarpa, K. Aki, Eds. (Springer, 1992), vol. 3, pp. 432–461.

38. A. J. Hotovec-Ellis, D. R. Shelly, D. P. Hill, A. M. Pitt, P. B. Dawson, B. A. Chouet, Deep fluid pathways beneath mammoth mountain, California, illuminated by migrating earthquake swarms. *Sci. Adv.* **4**, eaat5258 (2018).

39. I. Zaliapin, Y. Ben-Zion, Earthquake clusters in southern California I: Identification and stability. *J. Geophys. Res. Solid Earth* **118**, 2847–2864 (2013).

40. I. Zaliapin, A. Gabrielov, V. Keilis-Borok, H. Wong, Clustering analysis of seismicity and aftershock identification. *Phys. Rev. Lett.* **101**, 018501 (2008).

41. I. Zaliapin, Y. Ben-Zion, Discriminating characteristics of tectonic and human-induced seismicity. *Bull. Seismol. Soc. Am.* **106**, 846–859 (2016).

42. I. Zaliapin, Y. Ben-Zion, Earthquake clusters in southern California II: Classification and relation to physical properties of the crust. *J. Geophys. Res. Solid Earth* **118**, 2865–2877 (2013).

43. Z. E. Ross, E. S. Cochran, D. T. Trugman, J. D. Smith, 3D fault architecture controls the dynamism of earthquake swarms. *Science* **368**, 1357–1361 (2020).

44. E. K. Montgomery-Brown, D. R. Shelly, P. A. Hsieh, Snowmelt-triggered earthquake swarms at the margin of Long Valley Caldera, California. *Geophys. Res. Lett.* **46**, 3698–3705 (2019).

45. D. P. Hill, E. Montgomery-Brown, Long valley caldera and the Ucerf depiction of Sierra Nevada range-front faults. *Bull. Seismol. Soc. Am.* **105**, 3189–3195 (2015).

46. D. Dreger, H. Tkalčić, M. J. Johnston, Dilational processes accompanying earthquakes in the Long Valley Caldera. *Science* **288**, 122–125 (2000).

47. R. S. Matoza, P. G. Okubo, P. M. Shearer, Comprehensive high-precision relocation of seismicity on the island of Hawai'i 1986–2018. *Earth Space Sci.* **8**, e2020EA001253 (2021).

48. A. Pitt, D. Hill, Long-period earthquakes in the Long Valley Caldera region, eastern California. *Geophys. Res. Lett.* **21**, 1679–1682 (1994).

49. D. R. Shelly, J. L. Hardebeck, Illuminating faulting complexity of the 2017 Yellowstone maple creek earthquake swarm. *Geophys. Res. Lett.* **46**, 2544–2552 (2019).

50. R. J. Skoumal, J. O. Kaven, J. I. Walter, Characterizing seismogenic fault structures in Oklahoma using a relocated template matched catalog. *Seismol. Res. Lett.* **90**, 1535–1543 (2019).

51. Z. E. Ross, D. T. Trugman, E. Hauksson, P. M. Shearer, Searching for hidden earthquakes in southern California. *Science* **364**, 767–771 (2019).

52. W. Zhu, G. C. Beroza, Phasenet: A deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.* **216**, 261–273 (2019).

53. S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, G. C. Beroza, Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* **11**, 3952 (2020).

54. J. Woods, T. Winder, R. S. White, B. Brandsdóttir, Evolution of a lateral dike intrusion revealed by relatively-relocated dike-induced earthquakes: The 2014–15 Bárðarbunga–Holuhraun rifting event, Iceland. *Earth Planet. Sci. Lett.* **506**, 53–63 (2019).

55. A. Flinders, D. R. Shelly, P. B. Dawson, D. P. Hill, B. Tripoli, Y. Shen, Seismic evidence for significant melt beneath the Long Valley Caldera, California, USA. *Geology* **46**, 799–802 (2018).

56. L. K. Hutton, D. M. Boore, The ML scale in southern California. *Bull. Seismol. Soc. Am.* **77**, 2074–2094 (1987).

57. J. Woessner, S. Wiemer, Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bull. Seismol. Soc. Am.* **95**, 684–698 (2005).

58. M. Schoenball, N. C. Davatzes, J. M. G. Glen, Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso geothermal field. *Geophys. Res. Lett.* **42**, 6221–6228 (2015).

Citation: B. Q. Li, J. D. Smith, Z. E. Ross, Basal nucleation and the prevalence of ascending swarms in Long Valley caldera. *Sci. Adv.* **7**, eabi8368 (2021).