Methods

Experimental Setup

A biaxial apparatus, CrackDyn, was used to perform the experiments (Fig. 1a). The apparatus housed two polymethylmethacrylate (PMMA) plates (40 x 10 x 1 cm and 45 x 10 x 1.8 cm). Experiments were conducted on PMMA rather than natural rocks because it provides three major advantages for scaling laboratory observations to natural fault systems. First, due to its lower elastic stiffness compared to rocks, the state-evolution slip distance, L, and cohesive zone size, X_c, are smaller in PMMA. As a result, a laboratory-scale PMMA fault is dynamically representative of a much larger natural fault, whereas a laboratory rock fault essentially remains of the same scale as its natural counterpart. Second, the reduced elastic properties and fracture energy, G_c, of PMMA ensure that nucleation lengths, ℓ_c and ℓ_∞, remain smaller than the total fault dimension in the experiment, which is rarely the case in rocks, where nucleation patches can be comparable to or larger than the laboratory fault. Finally, the birefringent properties of PMMA allow for photoelastic visualization, enabling direct observation of nucleation growth and subsequent rupture propagation using high-speed imaging.

The plates were characterized by a static Young's Modulus, E, of 3 GPa and a Poisson's ratio, ν, of 0.35. A normal load was applied by three vertical pistons via steel sample holders. The pistons were supplied by a ENERPAC P141 hydraulic pump. Similarly, a shearing load was applied by a single horizontal piston. This piston was supplied by a TOP Industrie PMHP-35-1000 hydraulic pump. The force applied by each piston was recorded by a Scaime K13 load cell at 500 Hz. Local strains were recorded by thirteen 350-ohm strain gauge rosettes (Micro-Measurements C5K-06-S5198-350-33F) located at 3 mm from the sample-sample interface. The strain gauges recorded at 2 MHz and were amplified by a factor of 10 by Elsys SGA-2 MK2 amplifiers. Ten Philtec model D100-E2H2PQTS displacement sensors were fixed across the sample-sample interface and recorded the local displacement with a recording frequency of 2 MHz. Thirteen Brüel & Kjær type 8309 accelerometers were glued horizontally approximately 30 mm from the sample-sample interface and recorded at 2 MHz. These signals were amplified by a Nexus conditioning amplifier 2692. Finally, an effiLux EFFI-LINE3-WTR-600-000-POL-PWR-C light source was used to shine linearly-polarized light through the sample. This light then passed through a second linear polarizer before reaching a Phantom TMX 6410 high-speed camera, which recorded 1280 x 32 pixels at 1 MHz. Considering that PMMA is a birefringent material, this allowed for the use of photoelasticity to track changes in stress across the sample-sample interface[11,40–42]. A piezoelectric sensor attached to one of the PMMA samples was used as a trigger for an oscilloscope (Picoscope 4224 A) which generated a TLL-like signal which triggered the camera and allowed for the synchronization of the other acquisition systems.

Experimental Approach

Experiments were begun by applying either 100, 150, 200, 250, or 300 bar nominal normal stress (i.e., the pressure indicated on the analogue gauge of the pump supplying the vertical pistons). Next, the shear stress was increased by setting a constant flow rate, 3 $\frac{\mathrm{cm}^3}{\mathrm{min}}$, on the pump supplying the horizontal piston. These conditions were kept constant until enough slip had been accumulated such that the displacement sensors were out of range (approximately 0.5 mm). A 1.3 x 0.3 x 10 cm stopper was placed in between the lower-most PMMA sample and the horizontal piston. By adjusting this stopper prior to the experiment such that it was either in a raised or lowered position, the local loading conditions could be altered such that a richer variety events could be produced. Each of the five normal stresses were tested with both stopper positions, resulting in 10 total experiments and 94 dynamic events.

Data Treatment

As the strain gauges were set up in a quarter bridge configuration, the strain, ε, could be found for the i-th strain gauge as,

$$\varepsilon_i = \frac{-4\mathrm{U}_i}{\mathrm{U}_{\mathrm{ex}}\left(\mathrm{G_f}\mathrm{G_{amp}}\left(1+\frac{2\mathrm{U}_i}{\mathrm{U_{ex}}\mathrm{G_{amp}}}\right)\right)}, \tag{2}$$

where U_i is the voltage reading of the individual strain gauge, $\mathrm{U_{ex}}$ is the excitation voltage, $\mathrm{G_{amp}}$ is the amplification gain, and $\mathrm{G_f}$ is the gauge factor of the strain gauge.

In this three-strain-gauge rosette, the strain gauges were oriented at 45° from one another, such that the principal strains were then found as,

$$\varepsilon_{xx} = \varepsilon_{315} + \varepsilon_{45} - \varepsilon_{\mathrm{v}}, \quad \varepsilon_{yy} = \varepsilon_{\mathrm{v}}, \quad \varepsilon_{xy} = \frac{\varepsilon_{45} - \varepsilon_{315}}{2}. \tag{3}$$

when ε_{v} is the vertical direction, and ε_{315} and ε_{45} are oriented 315° and 45° clockwise from vertical (Fig. 1a). Here, x refers to the horizontal direction and y refers to the vertical direction.

These strains were then converted to stress, σ, considering Hooke's Law in plane stress conditions,

$$\begin{bmatrix}\sigma_{xx}\\ \sigma\\ \tau\end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix}1-\nu & \nu & 0\\ \nu & 1-\nu & 0\\ 0 & 0 & 1-2\nu\end{bmatrix}\begin{bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{xy}\end{bmatrix}. \tag{4}$$

The values of stress were filtered with a low-pass filter at 10kHz.

The camera recorded greyscale images in a 1280x32 matrix where the values ranged from zero (black) to a maximum value (white). Taking the first 20 frames as reference, the future frames were compared to this reference to detect changes in stress during rupture propagation[41,42]. One of the 32 rows (corresponding to the 1280 pixels closest to the sample-sample interface) was isolated and used for this analysis. The accelerometers' readings were converted to accelerations using each sensors' individually-calibrated sensitivities. Finally, the displacement sensors' readings were converted from voltage to gap distance using the calibration provided by Philtec. They were then filtered with a low-pass filter at 10 kHz. The nucleation duration was then defined as the interval between the foreshock recorded by the accelerometers and the transition to dynamic rupture, marked by rupture velocities exceeding 10 m/s as recorded by high-speed imaging.

Foreshocks and nucleation length scaling

In the present experimental dataset, the onset of nucleation is systematically associated with a foreshock. This feature is specific to this series of experiments and likely reflects the heterogeneous stress distribution along the fault, which varies between PMMA plates. In other laboratory studies, nucleation has been observed to occur without any identifiable foreshock. In such cases, the influence of foreshocks on the nucleation process cannot be investigated, and the nucleation size is, to a first order, defined by its theoretical asymptotic value, ℓ_∞[6], for under-stressed faults.

The coexistence of foreshocks and a larger nucleation front can be rationalized by considering the scale dependence of nucleation. At the scale of micron-sized asperities, local stress concentrations lead to significantly higher effective normal stresses than the macroscopically-applied value. Because the theoretical nucleation size scales inversely with normal stress[6],

$$\ell_\infty \approx \frac{1}{\pi}\frac{1}{(1-a/b)^2}\frac{\mu' L}{b\sigma} = \frac{1}{\pi}\frac{1}{(1-a/b)^2}\ell_b, \tag{5}$$

where ℓ_b is the elasto-frictional length scale, these local stress amplifications reduce the critical nucleation length at the asperity scale, enabling small slip instabilities that manifest as foreshocks. At the same time, the macroscopic nucleation process spans a much larger portion of the fault, controlled by the bulk stress state and the effective average σ, and can therefore extend over dimensions much greater than the asperity scale. This multiscale interplay explains how foreshocks, driven by localized asperity failure, can coexist with a larger quasi-static nucleation front that evolves toward dynamic rupture, as recently highlighted in the homogenized regime[10].

Transient sliding velocity

While the onset of stable slip is systematically triggered by a foreshock, the associated sliding velocity is too small to be found as an instantaneous measurement. Instead, slip is seen to increase exponentially with time (Fig. S3) and can fit with an exponential function. The derivative of this function is found just after the foreshock to provide the transient sliding velocity. A similar procedure is applied to the data digitized for natural earthquakes (Fig. S4, Tab. S1). Note that, in the case of data for natural earthquakes, these data rely on either inversion of the fault slip based on geodetic data or the supposition that the slip during repeater earthquakes is representative of the aseismic slip of the patch[43]. Both methods are subject to a number of assumptions and potential errors.

Foreshock magnitude

The foreshocks at the origin of all observable nucleation phases in these experiments are located using error minimization based on arrival times. This localization allows for the assessment of the event magnitude via an averaged acceleration, A_{fs}, found as[44],

$$A_{\mathrm{fs}} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(\frac{r_i}{0.01}A_i^{\max}\right)^2}, \tag{6}$$

where k is the number of accelerometers, r is the distance from the sensor to the event location (normalized by a reference sphere of 0.01 m), and $A_i^{\max}$ is the maximum acceleration measured by the sensor during the event.

Unified Framework for Earthquake Nucleation from Laboratory to Nature

The crack's Equation of Motion (EoM) stems from the Griffith's energy balance which together with a set of simplifying assumptions can allow for the solving of the crack front propagation[45]. Recently, the EoM framework has been applied to crack-like slip transients on rate-and-state frictional faults[6,9,46]. Here, we follow the EoM treatment by Garagash[9] who considers a plane-strain or anti-plane shear crack of half length ℓ, propagating at instantaneous velocity v_r along a planar 1D fault, characterized by the ambient sliding velocity V_0 and 'overstress', the difference between the initial stress (ratio) $f_0 = \tau_0/\sigma$ and the steady-state level of friction at the ambient sliding velocity, $\Delta f_0 = f_0 - f_\mathrm{ss}(V_0)$. The fault is embedded in an infinite elastic medium, and its rate-and-state dependent resistance to sliding[22] is governed by three parameters: the direct effect coefficient a, the state evolution coefficient b, and the characteristic slip distance L over which the state of the sliding surface is renewed (not to be confused with the apparent slip-weakening distance of the rupture, δ_c). Evolution of the state is often modeled in the framework of either aging[3] or slip[47] laws, which manifest in different apparent weakening behaviors of strength with incurred slip in a transient (e.g., behind a propagating rupture front). Specifically, a transient governed by the *aging*-law corresponds to a rate of weakening with slip ($W = -d\tau/d\delta$) which is approximately independent of the strength of the transient[9,48], $W \approx b\sigma/L$, while for the *slip*-law, the weakening rate increases with the transient strength, $W \sim \Delta f_\mathrm{p}\sigma/L$, where Δf_p is the maximum departure of friction from the steady-state during the transient (aka 'friction breakdown'). Conversely, the apparent slip-weakening distance δ_c required to 'break' the strength by amount $\Delta f_\mathrm{p}\sigma$ from the peak to the steady-state 'residual' value at the above values of the weakening rate W is an increasing function of Δf_p, $\delta_\mathrm{c} = (1/2)(\Delta f_\mathrm{p}/b)L$, for the *aging*-law, while it is invariant, $\delta_\mathrm{c} \sim L$. for the *slip*-law. Considering the apparent invariance of the weakening rate W in our experiments (Fig. 2), we adopt the aging law for the remainder of the analysis. However, previous studies have indicated that the slip law may be more appropriate for natural fault material, conversely to PMMA. Note that the model presents a qualitatively-similar behaviour using the slip law. For consistency and simplicity, we extend our results to natural earthquakes using the aging law rather than the slip law (Fig. 4d)

EoM Framework

The EoM considers the balance between the energy release rate $G = \bar{K}^2/(2\bar{\mu})$ into the propagating crack tip with the fracture energy of the frictional breakdown process G_c. Here $\bar{K} = k(v_\mathrm{r})K$ is the dynamic stress intensity factor obtained from the static value K multiplied by the wave-mediated dynamic prefactor $k(v_\mathrm{r})$; and $\bar{\mu} = g(v_\mathrm{r})\mu'$ is the apparent dynamic shear modulus formed from the static value $\mu' = \mu$ and dynamic prefactor $g(v_\mathrm{r})$. These dynamic factors are mostly inconsequential during the nucleation phase characterized by $v_\mathrm{r} \ll c_\mathrm{s}$, with $k \approx g \approx 1$, and full expressions[9] omitted here for simplicity. The fracture energy of a rate-and-state fault can be defined as the breakdown work $G_\mathrm{c} = \int \Delta f\,\sigma\,d\delta$ expended near the rupture front in dissipating the strength-excess over the steady-state value, $\Delta f = f - f_\mathrm{ss}(V)$, from its the peak value Δf_p, attained at the front, to zero some distance behind the front over some slip distance δ relatable to the state evolution slip distance L. Since the peak friction $\sim a\ln v_r$ and the peak friction breakdown $\Delta f_\mathrm{p} \sim b\ln v_\mathrm{r}$ are increasing with increasing rupture velocity by the direct and state-evolution effects, respectively, so is the fracture energy G_c. Contrary to the energy release rate G which materially depends on v_r only in the seismic range, the dependence of the fracture energy G_c on v_r is persistent over the entire range of rupture speeds, from quasi-static to dynamic. For aging law used here, Garagash[9] gives

$$G_\mathrm{c}(v_\mathrm{r}) = -\mathrm{Li}_2\left(1 - \exp\left(\frac{\Delta f_\mathrm{p}(v_\mathrm{r})}{b}\right)\right) b\sigma L, \tag{7}$$

where Li_2 is the dilog function. This expression further simplifies to $G_c \approx (1/2)\left(\frac{\Delta f_\mathrm{p}}{b}\right)^2 b\sigma L$ for transients with $\Delta f_\mathrm{p}/b \gg 1$. The dependence on rupture speed stems from that of the peak friction breakdown $\Delta f_\mathrm{p}(v_\mathrm{r})$, given implicitly by Equation (2.17) in Garagash[9]. Here we use an *explicit* expression,

$$\frac{\Delta f_\mathrm{p}(v_\mathrm{r})}{b} \approx -\frac{2}{3} W\left(-\frac{3}{2}\left(\frac{1}{\kappa_0\, g(v_r)}\frac{v_r}{\bar{v}_0}\right)^{-3/2}\right) \tag{8}$$

where $\bar{v}_0 = \exp\left(-\frac{\Delta f_0}{b}\right)\frac{\mu'}{b\sigma}V_0$ is the characteristic rupture velocity embodying initial fault conditions (ambient sliding velocity V_0 and initial overstress $\Delta f_0 = f_0 - f_\mathrm{ss}(V_0)$), $\kappa_0 \approx 0.838$, and $W(x)$ is the branch (W_{-1}) of the Lambert productlog function providing the real-valued solution of $n\,e^n = m$ for $-1/e < m < 0$. The above explicit inversion of Garagash's Eq. (2.17) for Δf_p is exact for the slip-law case and is an excellent approximation for the aging law case considered here. (Indeed,[9] observes that $\Delta f_\mathrm{p}(v_\mathrm{r})$ are essentially indistinguishable for the two laws when plotted on his Fig 4a).

A reasonable (zeroth-order) approximations, $\Delta f_\mathrm{p} \approx b\ln(v_r/\bar{v}_o)$ and $G_\mathrm{c} = (1/2)\,(\ln(v_\mathrm{r}/\bar{v}_0))^2\, b\sigma L$, allow one to readily glance at a simplified (logarithm-based) relation of the peak friction breakdown and the associated fracture energy on the rupture speed across the entire range of slip transients, from slow-to-fast.

The crack energy balance $G = G_{\mathrm{c}}$ can be further recast as $K = K_{\mathrm{c}}$, where static stress intensity factor K matches the effective fault toughness $K_{\mathrm{c}} = \left(\sqrt{g}/k\right)\sqrt{2\mu' G_{\mathrm{c}}}$ with the prefactor $\left(\sqrt{g}/k\right)(v_{\mathrm{r}})$ encapsulating the wave-mediated effects and $G_{\mathrm{c}}(v_{\mathrm{r}})$ the quasi-static-to-dynamic range of the fracture energy.

The SIF $K = K_{\mathrm{bg}} + K_{\mathrm{fs}}$, is comprised of the contributions from the background stress, K_{bg}, and the foreshock, K_{fs}, which can be approximated by the solutions for a bi-wing crack of half-length ℓ loaded by an effective uniformly-distributed stress drop $\Delta\tau_{\mathrm{eff}}$ and a localized, hypocentral point-force ΔT, respectively,

$$K_{\mathrm{bg}} = \Delta\tau_{\mathrm{eff}}\sqrt{\pi\ell}, \qquad K_{\mathrm{fs}} = \frac{\Delta T}{\sqrt{\pi\ell}}. \tag{9}$$

We note that an edge-crack geometry and corresponding slightly modified expressions for the SIF components can be alternatively used to describe the experimental ruptures, which start proximally to one of the experimental fault ends. However, such a modification is not expected to drastically improve the prediction of an already significantly-simplified model. To close-form the EoM, we now inform the SIFs with expressions for the effective stress drop $\Delta\tau_{\mathrm{eff}}$ and hypocentral force ΔT. The hypocentral (foreshock) force ΔT, by a dimensional argument, is

$$\Delta T = C\mu'\delta_{\mathrm{a}} \tag{10}$$

where δ_{a} is the foreshock 'asperity' slip and $C \sim 0.3$ is a numerical coefficient, which will be used as a fitting parameter. The effective (uniform) stress drop $\Delta\tau_{\mathrm{eff}}$ is the energetically equivalent measure of the actual, spatially-varying stress drop along the crack defined as

$$\Delta\tau(V) = (f_0 - f_{\mathrm{ss}}(V))\,\sigma \tag{11}$$

the drop from the initial stress $f_0\sigma$ to the 'residual' stress $f_{\mathrm{ss}}(V)\,\sigma$ set by the frictional steady-state

$$f_{\mathrm{ss}}(V) = (f_0 - \Delta f_0) + (a-b)\,\ln\left(\frac{V}{V_0}\right), \tag{12}$$

as parametrized here in terms of the fault initial state: f_0, V_0, and overstress $\Delta f_0 = f_0 - f_{\mathrm{ss}}(V_0)$. The spatial distribution of sliding velocity $V = V(x)$ along the crack then defines the spatial distribution of the stress drop. The energetically equivalent $\Delta\tau_{\mathrm{eff}}$ can be found from matching its SIF contribution (i.e., K_{bg}) to that of $\Delta\tau(V(x))$, i.e.,

$$K_{\mathrm{bg}} = \frac{\sqrt{\pi\ell}}{\pi}\int_{-\ell}^{\ell}\frac{\Delta\tau\left(V(x)\right)}{\sqrt{\ell^2 - x^2}}dx = \Delta\tau_{\mathrm{eff}}\sqrt{\pi\ell}. \tag{13}$$

Alternatively, we rewrite

$$\Delta\tau_{\mathrm{eff}} = \Delta\tau(V_{\mathrm{eff}}) = \frac{1}{\pi}\int_{-\ell}^{\ell}\frac{\Delta\tau\left(V(x)\right)}{\sqrt{\ell^2 - x^2}}dx, \tag{14}$$

where the effective sliding velocity V_{eff} is implicitly defined by the 2nd equality when the sliding velocity $V(x)$ is known. Approximating the latter from the solution for the slip on the propagating crack loaded by $\Delta\tau_{\mathrm{eff}}$ and ΔT (Section 3(b) of Garagash[9]), and solving for V_{eff} ultimately results in,

$$V_{\mathrm{eff}} = \frac{4K_{\mathrm{c}}}{(g/k)\,\mu'\sqrt{\pi\ell}}\,v_{\mathrm{r}}. \tag{15}$$

Finally, the EoM

$$\Delta\tau(V_{\mathrm{eff}})\sqrt{\pi\ell} + \frac{\Delta T}{\sqrt{\pi\ell}} = K_{\mathrm{c}} \tag{16}$$

together with above expressions for $K_{\mathrm{c}}(v_{\mathrm{r}})$ and $V_{\mathrm{eff}}(v_{\mathrm{r}}, \ell)$ can be solved for v_{r} as a function of crack length ℓ. The relation between these two parameters can be integrated to solve for time. Evolution of the fracture toughness, fracture energy, and sliding velocity also result.

In the main text, we simplify Equation 16 by introducing $K_{\mathrm{bg}} = \Delta\tau(V_{\mathrm{eff}})\sqrt{\pi\ell}$. K_{bg} is increasing with crack half length ℓ and scales with the rate-dependent transient stress drop $\Delta\tau$ - a function of the initial fault overstress parameter Δf_0 defined as the distance between the background stress ratio f_0 from the steady-state friction value $f_{\mathrm{ss}}(V_0)$ at the ambient sliding velocity V_0, and transient rupture velocity v_{r}. In our experiments, the fault is likely neutrally or mildly understressed ($\Delta f_0 \lesssim 0$; Fig. (3c, S6)). Additionally, we introduce $K_{\mathrm{fs}} = \Delta T/\sqrt{\pi\ell}$, which denotes the SIF contribution from the foreshock. It decreases with the crack half length ℓ and scales with a localized hypocentral Coulomb force of magnitude $\Delta T \propto \mu'\delta_{\mathrm{a}}$, generated by the impulse associated with the foreshock slip, δ_{a}, that initiates the nucleation process.

Illustrative examples of EoM solution for the sliding velocity histories resembling the experimentally observed, foreshock-mediated rupture nucleation regimes are shown in Fig. S9 and Fig. 3b using the parametrization explained in the following.

EoM Solution in Parametric Space

We use a suite of multiple simulations of the EoM to probe the dependence of foreshock-mediated nucleation process: foreshock - transient slow slip - acceleration to mainshock, as encapsulated by the nucleation time Δt_c, nucleation length ℓ_c, and the transient minimum sliding velocity $V_\mathrm{eff,min}$, on the foreshock force ΔT, initial fault conditions (Δf_0 and V_0), and a set of fault elasto-frictional constitutive parameters. We use a normalized EoM formulation which allows for the reduction of the parametric space to a non-dimensional foreshock force, $\frac{\Delta T}{\mu' L}$, initial overstress, $\frac{\Delta f_0}{b}$, and ambient sliding velocity, $\frac{V_0}{(b\sigma/\mu')c_\mathrm{s}}$, and a single constitute parameter, the ratio of the direct to state coefficients, a/b. We present results for the PMMA-like value $a/b \approx 0.55$[49], while also performing a limited suite of solutions for the rock-like value $a/b \approx 0.75$[37] to further validate the emerging correlations discussed below.

Nucleation Length, Nucleation Time, Transient Minimum Velocity, and Inversion of Fault State-Evolution Distance

The EoM suite of solutions for the normalized nucleation length vs. nucleation time is shown on Fig. S2 for a wide variety of the non-dimensional loading parameters, foreshock force $\frac{\Delta T}{\mu' L}$ (implicit) and initial overstress $\frac{\Delta f_0}{b}$, and non-dimensional ambient sliding velocity $\frac{V_0}{(b\sigma/\mu')c_\mathrm{s}}$. Length and time to nucleation decrease with both increasing foreshock impulse and increasing fault initial overstress. The theoretical maximum nucleation length ℓ_∞[6,50] serves as the upper bound which is approached asymptotically on understressed ($\Delta f_0 < 0$) faults when the foreshock force decreases to the Δf_0-dependent threshold value for the transient arrest (and as time to nucleation diverges).

Performed EoM simulations show a collapse of time-to-nucleation Δt_c versus minimum effective sliding velocity, $V_\mathrm{eff,min}$, for a wide variety of the non-dimensional loading parameters, $\frac{\Delta T}{\mu' L}$ and $\frac{\Delta f_0}{b}$, and ambient sliding velocity $\frac{V_0}{(b\sigma/\mu')c_\mathrm{s}}$ (Fig. S5a). This collapse manifests two asymptotic behaviours where $\Delta t_\mathrm{c} \sim L/V_0$ is a constant for transients with very small minimum sliding velocity, comparable to the initial fault value, i.e. when $V_\mathrm{eff,min} \sim V_0$, and $\Delta t_c \sim (L/V_\mathrm{eff,min}) \ln(V_\mathrm{eff,min}/V_0)$ when $V_\mathrm{eff,min} \gg V_0$. Essentially, to the first order Δt_c is inverse in $V_\mathrm{eff,min}$ with the low-velocity cut-off at $\sim V_0$. Small departure from the collapsed curve on Fig. S5a corresponds to the solutions with appreciable understress $\Delta f_0 < 0$. This second-order behavior is more apparent when we remove the dominant (inverse in $V_\mathrm{eff,min}$) behaviour when plotting the product $\Delta t_c V_\mathrm{eff,min}$ in Fig. S5b.

The overall relation between Δt_c and $V_\mathrm{eff,min}$ (Fig. S5a) can be well modeled by,

$$\Delta t_\mathrm{c} = \frac{1}{\pi(1-a/b)} \frac{L}{V_\mathrm{eff,min}} \times \mathscr{D}\left(\frac{V_\mathrm{eff,min}}{V_0}\right), \tag{17}$$

with the transition function

$$\mathscr{D}(\mathscr{V}) = \left((C_0 \mathscr{V})^{-\alpha} + \left(C_1 \ln \frac{\mathscr{V}}{\mathscr{V}_1} \right)^{-\alpha} \right)^{-1/\alpha}, \tag{18}$$

between the small-velocity, $\mathscr{D} \sim C_0 \mathscr{V}$, and large-velocity, $\mathscr{D} \sim C_1 \ln(\mathscr{V}/\mathscr{V}_1)$, asymptotes, and fitting constants C_0, C_1, $\mathscr{V}_1$, and α found as 0.64174, 1.33245, 0.01332, and 2.15680, respectively.

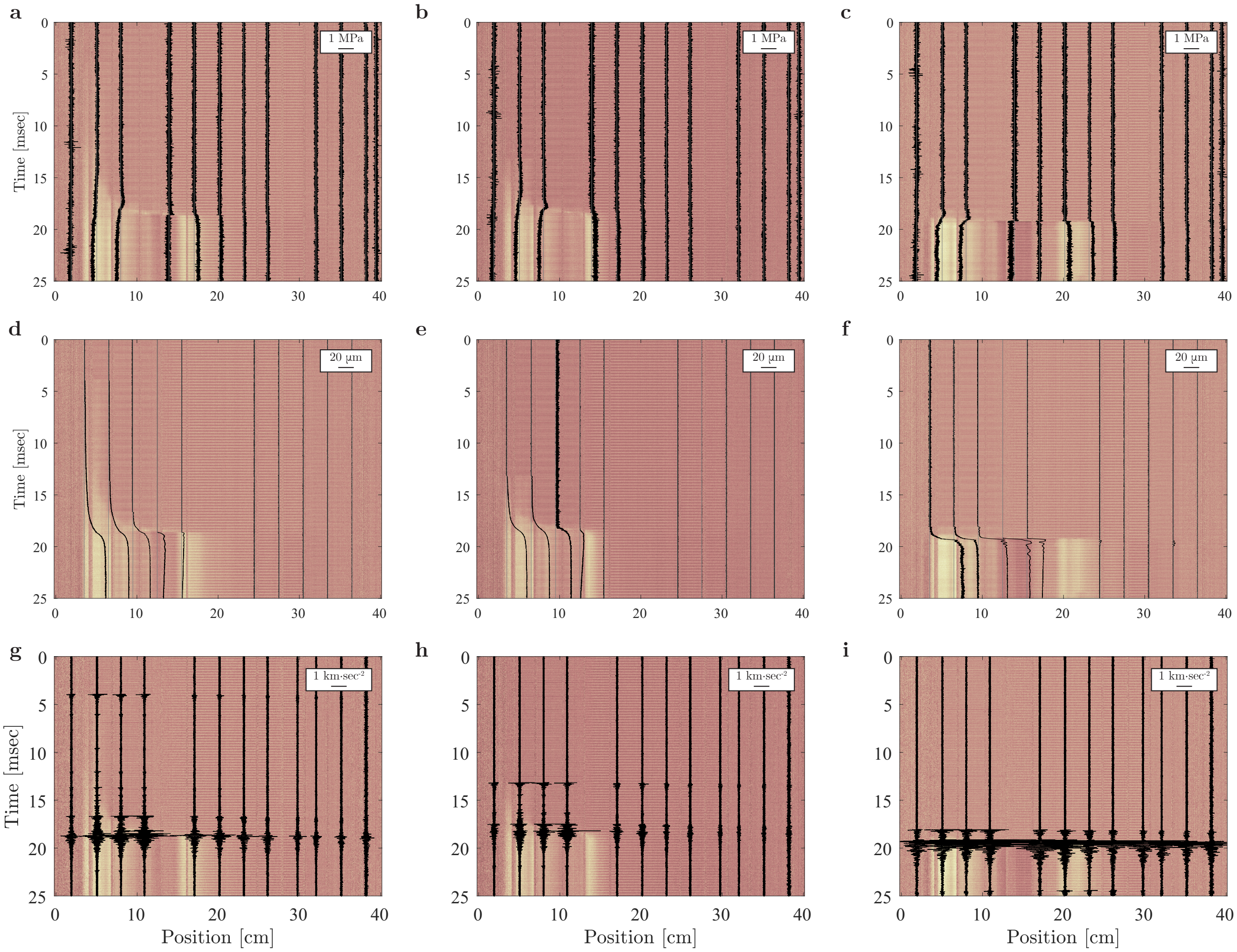
To perform the fit of the EoM empirical relation of the time-to-nucleation Δt_c and the average minimum sliding velocity $V_\mathrm{eff,min}$, Equation (17), to the correlation of experimental events Δt_c with the minimum value of the sliding velocity V_min measured at a fixed point on the fault, we need to relate the point-wise value V_min to the crack-average one $V_\mathrm{eff,min}$. To accomplish this, we use the EoM approximate solution for the sliding velocity distribution[9] $V(x) = (V_\mathrm{eff}/2)\left(1-(x/\ell)^2\right)^{-1/2}$ and model the sliding velocity measured at a fixed location on the fault proximal to the foreshock as the hypocentral value of the EoM distribution. In other words, we use the model's $V_\mathrm{eff,min}/2$ as the proxy for the experimental V_min.

Taking $a = 0.016$ and $b = 0.029$ for PMMA laboratory fault[49], Equation (17) was fit to the laboratory data, yielding an fault ambient sliding velocity $V_0 = 1.35 \cdot 10^{-6}$ m/sec and the state evolution slip distance of the interface $L = 0.19\,\mathrm{m}$ comparable to $\sim 0.4\,\mathrm{m}$ inferred from velocity-stepping experiments on a PMMA fault[49].

Similarly, in Fig. 4d, Equation (17) was used to prepare possible fits for the natural earthquake data assuming granite friction[37] with $a = 0.014$ and $b = 0.019$ and natural fault ambient sliding velocity $V_0 = 10^{-12}$ m/sec. We note that the transient (minimum) sliding velocity values inferred for the natural earthquakes in our dataset (Fig. 4d) are of the order of or exceed the typical plate convergence rate, i.e. $V_\mathrm{min} \gtrsim 10^{-9}$ m/s. Thus, any suitable choice of the ambient sliding velocity assuming an initially 'locked' fault, i.e. $V_0 \ll 10^{-9}$ m/s, will not have significant affect on the EoM fit for the natural events.

Transient Minimum Velocity vs. Foreshock Force vs. Foreshock Slip

The EoM empirically predicts a collapse of $V_\mathrm{eff,min}$ and ΔT only in the case of large overstress (Fig. S6). In those cases, we find $\frac{V_\mathrm{eff,min}}{V_0} \approx 10 \cdot e^{\frac{\Delta T}{0.51\mu' L}}$, where $\Delta T = C\mu'\delta_\mathrm{a}$. We use this relation together with the fitting constant $C = 0.3$ to produce the upper-bound dashed curves in Fig. 3c for different values of L bracketing the best-fit value of 0.19 μm. Indeed, the EoM does



Supplementary Figure S1. The videograms of the events in Fig. 1 overlain with variations in **a-c**, shear stress, **d-f**, slip displacement, and **g-i**, acceleration. In all sub-figures, the reading is zeroed to its position on the experimental fault and deviation from this position corresponds to a change in the measured parameter in accordance with the scale given in the top-right of each videogram.

not indicate that there is a one-to-one relationship between $V_{\mathrm{eff,min}}$ and the hypocentral force, ΔT. Instead, $V_{\mathrm{eff,min}}$ also depends on the overstress. Rather, $V_{\mathrm{eff,min}}$ and its experimental proxy $V_{\min}$ encapsulate the loading conditions on the fault from both prestress and the foreshock (asperity slip), and can not be reduced to one or the other.

EoM modeling of experimental sliding velocity history

The model parameters used to produce the experiment-reflecting results for transient sliding velocity evolution of Fig. 3b are the same as used above for PMMA with the addition of the shear wave speed taken as 1345 m/sec, the apparent shear modulus taken as 1.24 GPa, and the normal stress taken as 0.8333 MPa (the average of the normal stresses given in Fig. 2). The asperity slip δ_a is converted into an equivalent hypocentral Coulomb force ΔT using a single asperity equivalent point force, $\Delta T = C\mu'\delta_a$ with $C = 0.3$.

Supplementary Table S1. Laboratory and natural earthquake data from literature.

Earthquake	Year	Δt_c [sec]	Quasi-static length [m]	Quasi-static slip [m]	V_{min} [m/s]
Tohoku-Oki[a]	2011	1.99e6	75000	0.2	1.6e-8
Illapel[b]	2015	1.21e7	50000	0.113	3.4e-9
Iquíque[c]	2014	2.33e7	110000	0.09	8.9e-10
Papanoa[d]	2014	5.26e6	250000	0.05	6.5e-10
Ibari-Oki[e]	2008	3.33e5	44600	0.017	2.5e-8
Izmit[f]	1999	2.64e3	300	0.0124	2.7e-7
Izmit[g]	1999	2.64e3	300	0.160	1.3e-5
Valparaíso[h]	2017	3.46e5	90000	0.09	-
Kumamoto[i]	2016	1e5	20000	0.255	-
Silivri[j]	2019	3.72e5	7000	-	-
Duvernay[k]	2016	1.9e4	2000	-	-
Laboratory[l]	-	9	0.24	1.5e-5	3.6e-7
Laboratory[m]	-	4.5	0.76	5.8e-5	1e-7
Laboratory[n]	-	4.85e-2	1.88	-	1e-5
Laboratory[o]	-	1.22e-1	1.27	-	-
Laboratory[p]	-	1.2e-2	0.027	-	-
Laboratory[q]	-	3.3e-1	1.1	-	-

[a] Δt_c from Figure 3. Quasi-static length from Figure 2. V_{min} from Figure 3Bc[12].

[b] Quasi-static length from Figure 3b. Slip and Δt_c from page 3918. V_{min} from Figure 4c evaluated at -140 days[31].

[c] Δt_c from abstract. Slip from Figure 3b. Quasi-static length from Figure 4. V_{min} from Figure 3b[14].

[d] Quasi-static length from Figure 2. V_{min} estimated from Figure 2c assuming mainshock at slip of 5 cm and constant patch size[28].

[e] Quasi-static length and Δt_c from Figure 5a. V_{min} from Figure 5d[30].

[f] Slip and V_{min} from Table S1 assuming shear modulus of 30 GPa. Patch size from page 879[18].

[g] Slip and V_{min} from Table S2[29]. Patch size assumed the same as [f].

[h] Quasi-static length from Figure 5. Δt_c from text page 10293. Slip from Figure 3[15].

[i] Quasi-static length and Δt_c from Figure 3. Slip from page 8951[27].

[j] Quasi-static length from abstract. Δt_c from Figure 5[34].

[k] Δt_c from Figure 5d. Quasi-static length from Figure 3[33].

[l] All data from Figure 20 and 21[25].

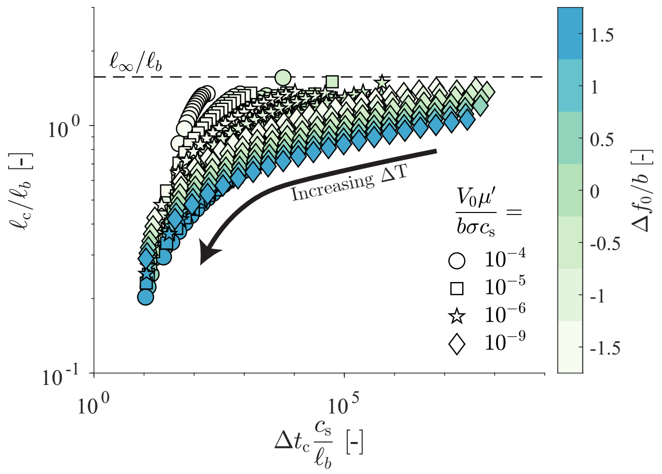
[m] Δt_c and quasi-static length from Figure 2c. Slip and V_{min} from Figure 2b[26].

[n] Quasi-static length from Figure 14a. Δt_c, slip, and V_{min} from Figure 14b[26].

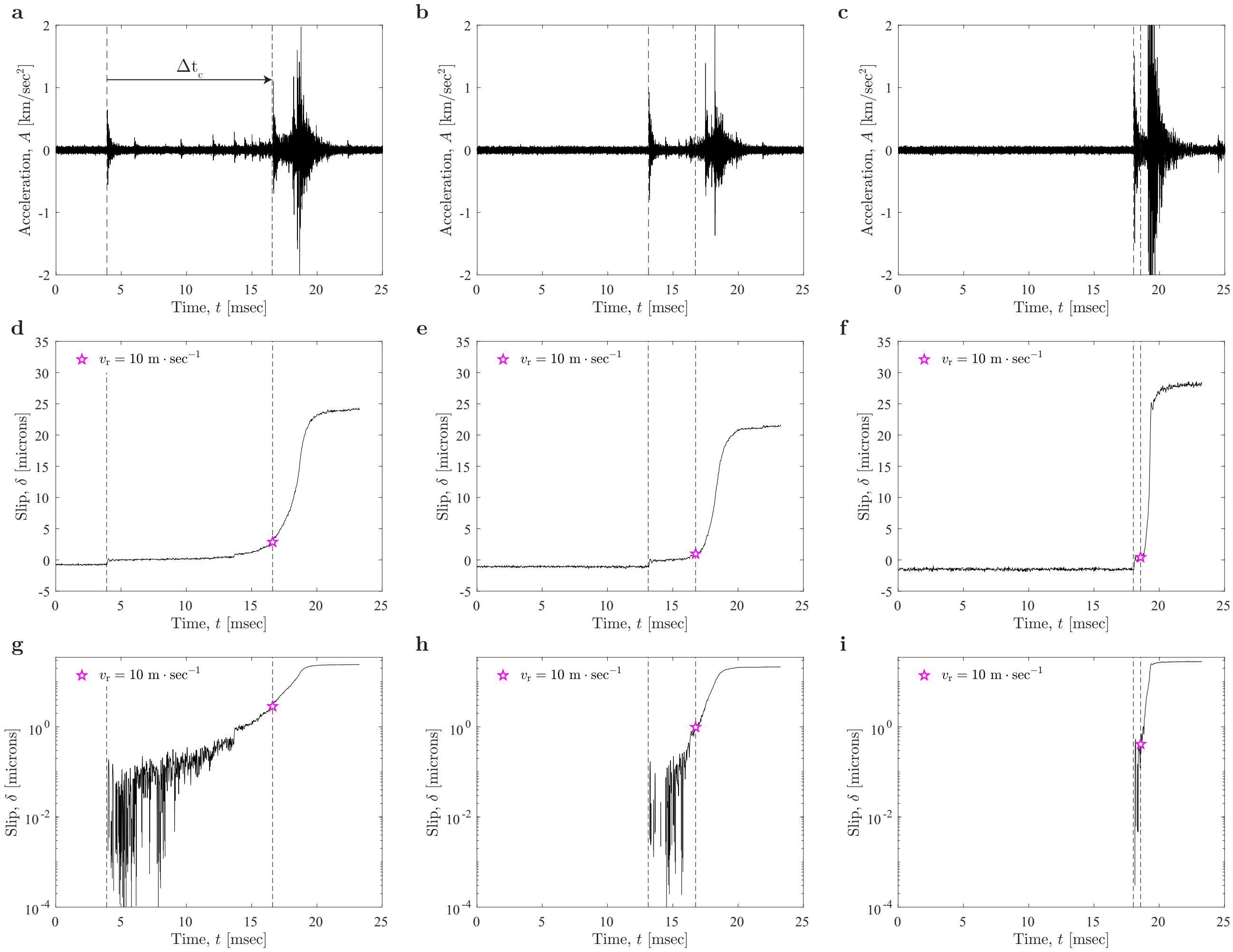
[o] Quasi-static length and Δt_c from Figure 2d. Figure 2c considered dynamic[36].

[p] Data from Figure 3[11].

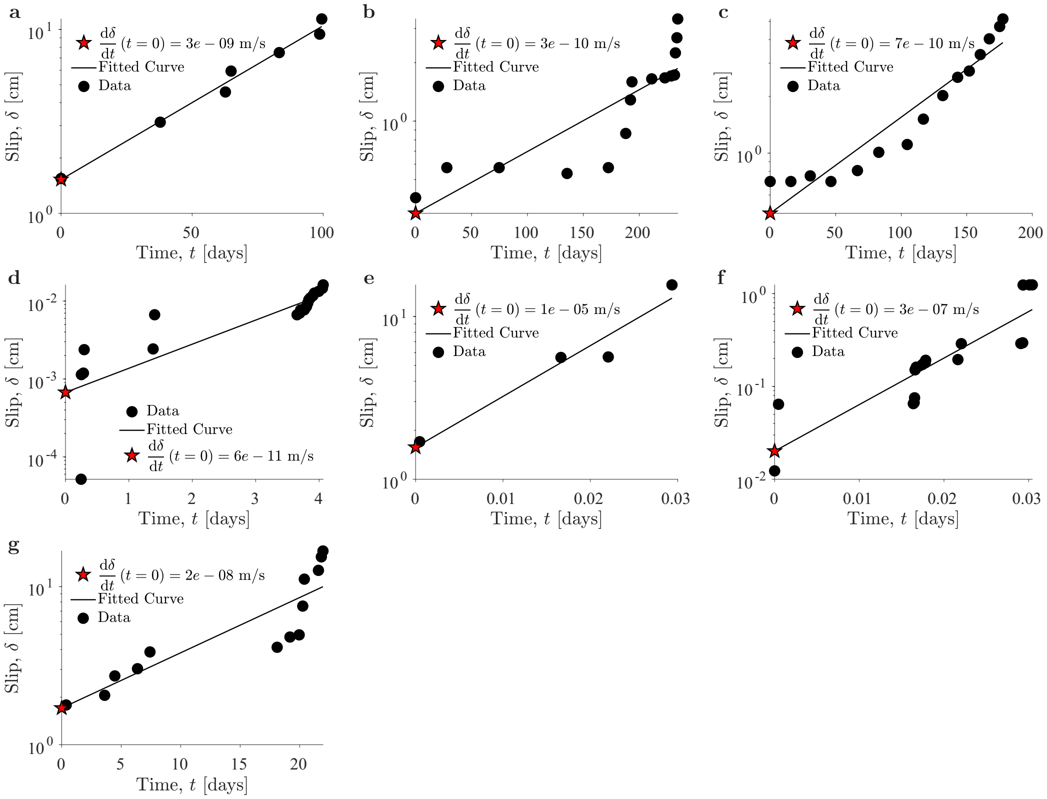
[q] Quasi-static length and Δt_c from Figure 2a[35].



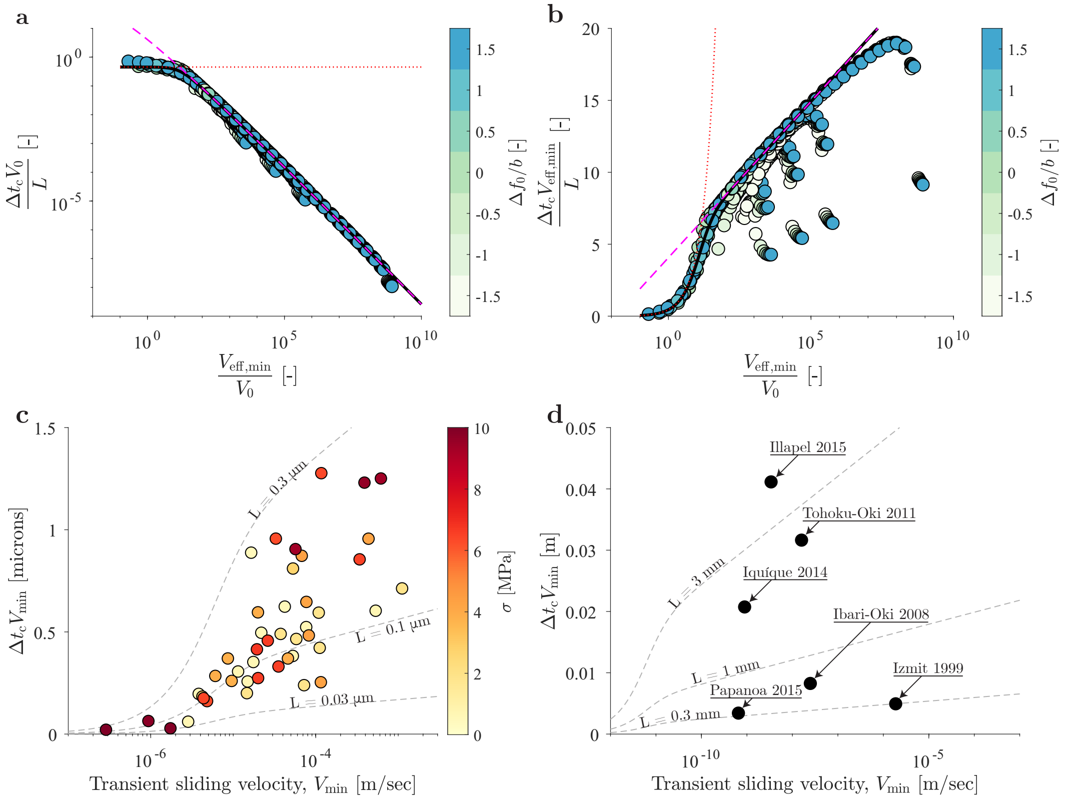
Supplementary Figure S2. EoM solution for the nucleation length ℓ_c vs. time Δt_c for varying foreshock force ΔT (implicit in this plot), various values of the overstress Δf_0 given by the color bar and ambient sliding velocity V_0 shown by symbols. Both overstress and foreshock impulse reduce the nucleation length ℓ_c up to an order of magnitude below the theoretical maximum length ℓ_∞ indicated by dashed line.



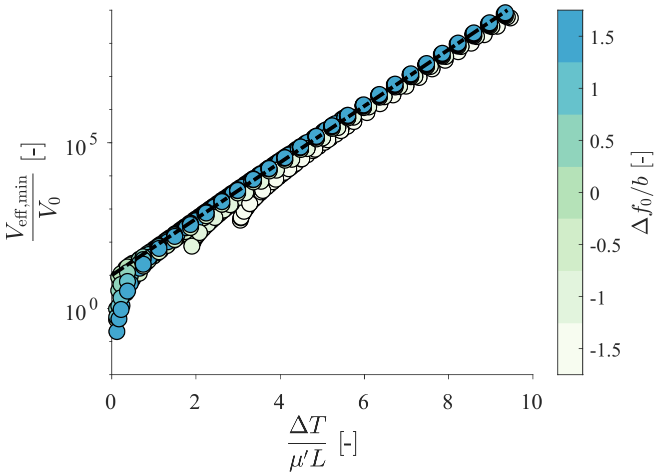
Supplementary Figure S3. **a-c**, The accelerations recorded by the three closest accelerometers to the foreshock plotted one top of the other for the events shown in Fig. 1. **d-f**, The local slip recorded by the closest displacement sensor to the nucleating patch. The moment at which the rupture front passes 10 meters per second is indicated with a star. The nucleation time, Δt_c, is the time between the foreshock and the rupture velocity reaching 10 meters per second. Both of these moments are indicated with dotted vertical lines. **g-i**, The same figure as **d-f** but with slip shown on a log scale to illustrate that the slip is occurring.



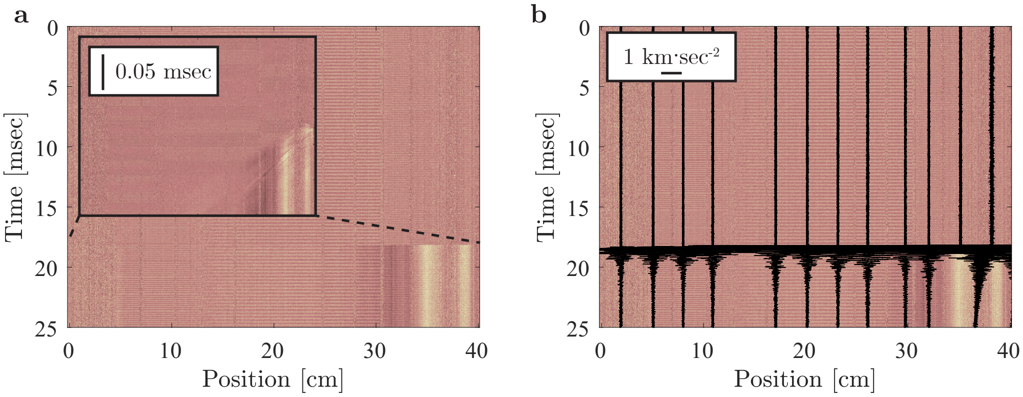
Supplementary Figure S4. Demonstration of how the transient sliding velocities of the **a**, Illapel 2015[31], **b**, Iquíque 2014[14], **c**, Papanoa 2014[28], **d**, Ibari-Oki 2008[30], **e**[29], **f**[18], Izmit 1999[18,29], **g**, Tohoku-Oki 2011[12] earthquakes were estimated. The data (Tab. S1) were fit with an exponential function. The derivative of this exponential was then evaluated at the start of the sequence, yielding an estimate of transient sliding velocity. The geometric average of **e** and **f** was taken for Izmit 1999 with the two values providing a crude error estimate.



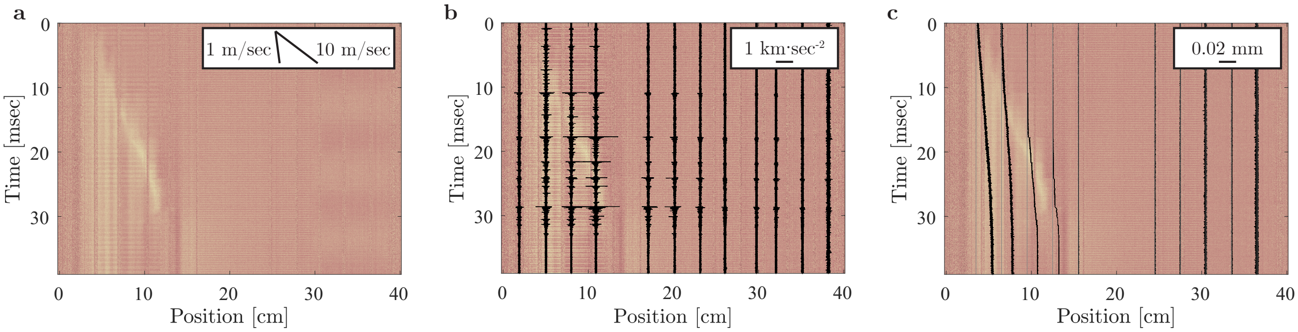
Supplementary Figure S5. **a**, EoM simulations upon which Equation (17) is based. The data show a collapse, regardless of overstress, onto the primary $\Delta t_c \sim \frac{L}{V_{\text{eff,min}}}$ dependence with the low-velocity cut-off at $V_{\text{eff,min}} \sim V_0$. **b**, By removing the primary $\Delta t_c \sim \frac{L}{V_{\text{eff,min}}}$ dependence, the second order trend is visible. In **a** and **b**, the overstress is given by the color bar. The first order trend captured by $\mathscr{T}_1$ is shown with the magenta dashed line. The second order trend captured by $\mathscr{T}_2$ is shown with the red dashed line. The combined trend is shown with the black line. **c**, Fig. 4c detrended in order to show the second order dependence of Δt_c on V_{min}. The collapse (for varying ΔT and Δf_0) for $L = 0.03$, 0.1, and 0.3 microns are shown with $\frac{a}{b} = 0.55$, $V_0 = 10^{-6}$ m/s. **d**, The equivalent to **c** for natural earthquakes and $\frac{a}{b} = 0.75$, $V_0 = 10^{-12}$ m/s, and $L = 0.3$, 1, and 3 mm.



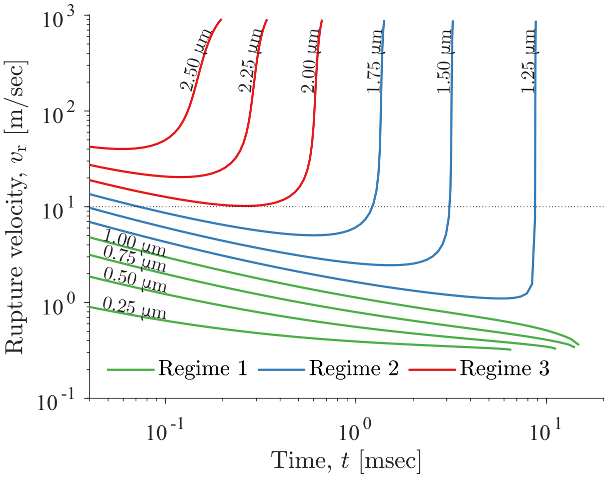
Supplementary Figure S6. EoM simulations showing the relationship between foreshock size and initial minimum sliding velocity. The EoM only predicts collapse for large overstress cases. $V_{\text{eff,min}}$ (and V_{min} by association) depends not only on foreshock size but also overstress. The overstress is given by the color bar. The collapsed portion of the data is fit with $\frac{V_{\text{eff,min}}}{V_0} = 10 \cdot e^{\frac{\Delta T}{0.51\mu' L}}$; this trend is shown with a dashed black line.



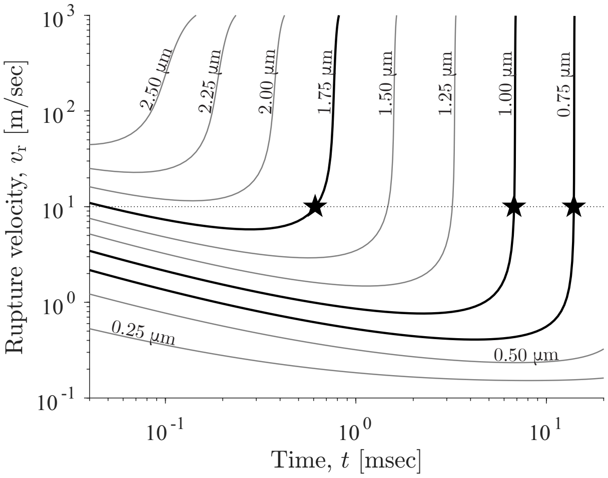
Supplementary Figure S7. An example of an event with no discernible nucleation phase. **a**, The videogram of the event with an inset showing a zoom on the start of the dynamic event. **b**, The same videogram overlain with the accelerometer data, showing no discernible foreshock prior to the event.



Supplementary Figure S8. An example of an event with an extended quasi-static slip phase but without the nucleation of dynamic rupture (maximum rupture velocity approximately 4 m/sec). **a**, The videogram of the event. **b**, The same videogram overlain with the accelerometer data, showing the initiation of the quasi-static slip phase with a foreshock. **c**, The videogram overlain with the local slip displacement data. Note that, while this event can be interpreted to not have had an appropriate hypocentral force and overstress to develop into dynamic rupture as predicted by the EoM, an alternative interpretation could be that the crack tip propagated into an understressed region (barrier) resulting in its arrest.



Supplementary Figure S9. EoM simulations for an understressed case, $\frac{\Delta f_0}{b} = -1$. The asperity slip, δ_a, during the foreshock for each simulation is written. The other parameters are the same as those used in Fig. 3b. The three regimes discussed in the main text are present and given by the color. Regime 1: transient deceleration followed by arrest without the nucleation of dynamic rupture. Regime 2: transient deceleration to a minimum sliding velocity, V_{min}, following by acceleration slip and rupture velocity leading to dynamic rupture. Regime 3: near-immediate acceleration into dynamic rupture, without a quasi-static phase.



Supplementary Figure S10. The rupture velocity corresponding to the simulations shown in Fig. 3b. The overstress is $\Delta f_0 = 0$ and the asperity slip, δ_a, is shown on each simulation's curve. Dark lines correspond to the cases most closely resembling the asperity slip of the experiments shown in Fig. 3a. Stars indicate the nucleation of dynamic rupture.