# **The Long Precursory Phase of Most Large Interplate Earthquakes**

# **Supplementary Information**

# **Supplementary Methods**

## **1. Data and Classification**

We select the earthquakes studied using the worldwide USGS National Earthquake Information Center (NEIC) catalog and consider all the  $M > 6.5$  events in the catalog which occurred in the two selected zones (Fig. 1) [20°N to 46°N, 118°E to 148°E] and [20°N to 65°N, 180◦W to 109◦W] between 01/01/1999 and 01/01/2011 and are located above 50km depth. This yields a set of 72 events (Supplementary Tables S1-S3). Of these we remove 10 events which are early aftershocks of large prior events and occur less than one week after. In the case of the very large Tokachi-Oki M 8.3 earthquake which is followed by many large aftershocks, we double this period of exclusion to 2 weeks (Supplementary Table S3). This provides a set of 62 events, which divides equally into 31 interplate and 31 intraplate earthquakes.

The identification of an earthquake as an interplate or an intraplate event is generally straightforward. We follow here the classification used by the USGS which is the leading agency for reporting earthquakes worldwide, the term interplate meaning an event which, to the best of our knowledge, occurs on the plate interface. In contrast the term intraplate is applied to events which represent the internal deformation of a plate. For some of the largest events which have occurred since 2002, referred to by (1) in Supplementary Tables S1 & S2, this identification is provided by the USGS in its Tectonic Summary of Significant Earthquakes. In this case, the terms used in the Tectonic Summary to characterize the event are cited in Supplementary Tables S1 & S2. The Alaska Earthquake Information Center (AEIC) also provides a tectonic description of most of the Alaska-Aleutian events and identifies them as interplate or intraplate events. These events are denoted by (3) in Supplementary Tables S1 & S2 and the terms of the AEIC description are cited. Some events have also been the subject of scientific publications and their classification is taken from these published studies referred to in Supplementary Tables S1 & S2 and the terms used to characterize them are repeated. For most of the remainning events (referred to as (2) in Supplementary Tables S1 & S2), the USGS provides (since 2002) detailed technical information, maps and cross-sections which allow a generally clear and straightforward classification. Out of the 62 events of our dataset, 53 have at least one of the types of information or identification described above. For the 9 remaining events, we use the International Seismological Center (ISC) location together with the Centroid Moment Tensor (CMT) of Harvard (HRVD) or GCMT to classify the event. Following the USGS definition of interplate and intraplate earthquakes, in subduction zones all the events not located on the subducting interface are considered as intraplate events. On transform faults, in the few cases where the plate boundary is diffuse or geometrically complex, events located near the boundary which display the expected boundary slip mechanism are classed in the interplate group. This is the case for instance of the 10/16/1999 Hector Mine earthquake which occurred within the eastern California shear zone, known to accommodate about 24 % of the relative Pacific-North American plate motion<sup>25</sup>.

For each selected event, we obtain the pre-event seismicity in the first zone studied (Fig. 1 left) from the bulletin of the International Seismological Centre (ISC) which combines catalogs from different agencies and regional networks. In this catalog, we exclusively use the locations and magnitudes reported by the TAP/CWB (Taiwan Central Weather Bureau) and JMA (Japan Meteorological Agency) agencies, as they best cover the region. In the second zone studied (Fig. 1 right), we use the AEIC (Alaska Earthquake Information Center) catalog for Alaska, the Natural Resources of Canada catalog for Canada and southern coastal Alaska, the USGS NEIC catalog together with the SCEC (Southern California Earthquake Center) catalog for California and Baja California. The analysis of the resulting dataset (Supplementary Fig. S1) shows that the magnitude of completeness is about 2.5 for both interplate and intraplate sequences. We shall use this value of completeness when modeling or comparing sequences while we will use all the catalog events when investigating individual sequences to keep the maximum of information.

#### **2. Time evolution graphs**

The graphs showing the time evolution of the cumulative number of seismic events (Fig. 2, Supplementary Figs S2-S9,S11) and of the cumulative seismic moment (Fig. 3a) for individual sequences include all the events listed in the catalogs and located within 50km from the mainshock epicenter, regardless of depth. The choice of this radius is based on the observation that the sequences which show the clearest increase of activity in the hours preceding the main shock (the ones displayed in Supplementary Fig. S9) have more than 98% of their pre-earthquake events in this range. As shown in Supplementary Fig. S10, this choice is not critical to the study. The curves of Supplementary Figs S2-S9 show the evolution of the cumulative numbers of events prior to the interplate earthquakes in different time windows (each sequence is presented at least once).

## **3. Stacks**

Stacking of the pre-earthquake sequences is done by giving the same weight to each one, so that sequences with many or large events do not dominate over those with fewer or smaller events.

The stacks over 6 months (Fig. 3b, Fig. S10) are done over 28 pre-earthquake sequences. Three sequences are not included, n<sup>∘</sup>8,14,17 because of large prior earthquakes occurring nearby within this period.

The stacks over 5 days (Fig. 3c, Fig. S10) are done over 26 pre-earthquake sequences. Five sequences are not included, either because they shortly follow another large earthquake (n°14) or because no event occurred in this period.

The stacks over 1 day (Fig. 3d, Fig. S10) are done over 22 pre-earthquake sequences. All sequences with events that day are included.

The stacks of the cumulative numbers of events over 150 days (Fig. 4a,b, Supplementary Fig. S15) do not include interplate sequences n<sup>∘8,14,17</sup> because of large prior earthquakes occurring nearby within this period (see Supplementary Table S1) and intraplate sequence n<sup>∘</sup>p15 which is associated with magmatic activity (see Supplementary Table S2). Adding these sequences, however, would not significantly change the stacks, because each sequence carries the same weight. All the events of magnitude  $> 2.5$ , the magnitude of completeness of the dataset, are included.

### **4. Statistical analysis of interplate and intraplate sequences**

In order to try to quantify the observed seismicity increase and to evaluate how significant it is, we design a simple and straightforward statistical test: A pre-mainshock time period T is chosen and the number of earthquakes in the first half of T is compared to the later half. If the second half has more events the rate has increased. In this case, the process continues and we look now if seismicity accelerates also in this second half. To do this, we simply replace T by T/2 (the pre-mainshock time period is now half of what it was previously) and we compare again the number of events in the first and second half of this new time period. If the second half has more events the rate has increased again and we keep repeating this process, dividing each time the current pre-mainshock time period by 2. The process stops when the second half of the current pre-mainshock time period has a number of events equal or less than its first half. The acceleration index, denoted by n, is the number of times the process can be repeated until it stops. Thus, the larger n is, the higher is the acceleration of seismicity of the sequence. To insure that the results are independent of the choice of the starting value chosen for T, we systematically make the calculation for 6 different starting values T=6months, 3months, 1month, 10days, 5days, 1day. This yields for each sequence 6 values of n, of which we select the largest one. We compare this number to the one obtained for synthetic sequences having the same number of events and obeying Poisson statistics. From this comparison, we obtain a probability that the observed value of n is due or not to chance. This algorithm is as follows:

(1) Begin with  $T=6$ months and n=0

(2) Divide the time duration T of the sequence in two:  $-T < t < -T/2$  (N1 events) and  $-T/2 < t < 0$  (N2 events).

There is acceleration if  $N2 > N1$  (more events occur in the second half of the time window):

 $-$  if  $N1 > N2$ , the calculation stops and the current value of n is kept.

 $-$  if  $N1 < N2$ , then n=n+1, T=T/2, and step (2) begins again.

At the end an index n is obtained for the 6-months time window.

(3) Steps (1) and (2) are performed for 5 other time windows: T=3months, 1 month, 10days, 5days, 1day as well. This yields 6 values of n. The largest one  $n$  (observed)=max $[n(T)]$  is kept.

(4) For each original 6-months long sequence of Ne events, 1000 synthetic sequences of Ne events randomly distributed in time over 6 months (Poisson statistics) are generated.

(5) The optimal value of n (denoted n(synthetic)) is calculated for each random sequence following steps (1) to (3) above and using the same 6 time windows.

(6) The comparison between the value of n observed and the 1000 values of n calculated gives the probability that the level of n observed is due or not due to chance: Probability(*chance*) = [Number of random sequences for which  $n(synthetic) \geq n(observed)$  / 1000; Probability(*not chance*) = 1 -Probability(*chance*).

The results are presented in Fig. 4c-d.

To analyse if the acceleration present in interplate sequences is really related to the occurrence of the large earthquakes, we perform the same test on the interplate sequences of the previous 6 months (that is on the sequences which begin 1 year before the  $M > 6.5$  earthquakes and end 6 months before they occur). The results are presented in Supplementary Fig. S12.

## **5. Statistical analysis of the effects of background seismicity and clustering on the observed accelerations**

It has been shown  $21$  that the clustering of earthquakes causes an acceleration of seismicity before a mainshock. Earthquakes trigger aftershocks, the more so as the magnitude of the trigger is large. The probability to pick an event is thus greater just following a large shock, when the seismicity rate is high. Immediately before the picked event, there are thus a greater number of earthquakes than normal, causing an apparent acceleration of seismicity that has nothing to do with anomalous precursory activity like pre-slip.

To estimate how much statistical (rather than mechanical) acceleration contributes to our observations, we perform two different tests. First, we select the event in each sequence, which takes place anywhere between 6 months and 3 months before the  $M > 6.5$  earthquake and is located the closest to it in this period. These limits are set so that the selected event samples the same type of background seismicity as the  $M > 6.5$  earthquake and is not affected by the pre-earthquake acceleration. We then stack the seismicity of the 31 interplate sequences relatively to the occurrence time and location of these selected events. The resulting graph, shown in Supplementary Fig. S13, shows that the acceleration due to clustering is much smaller than the acceleration observed.

In a second test, we perform Monte-Carlo simulations of an ETAS model<sup>22</sup>. In this model, each earthquake of magnitude m triggers aftershocks according to a K  $e^{\alpha m}$  (t+c)<sup>-p</sup> density, where t is the time after the triggering earthquake, and K,  $\alpha$ , c and p are parameters. A constant background rate  $\theta$  that models tectonic loading is added to this triggering. We separately consider the two sets of interplate and intraplate earthquakes: (1) since the model is linear, we simply superpose all 31 sequences together for each set; (2) a completness magnitude equal to 2.5 is found by inspection of the frequency-magnitude relationship for both sets (Supplementary Fig. S1); (3) the best parameters  $(K, \alpha, p, c)$  are estimated for the time interval extending from -365 days to -60 days before the mainshock, in order to avoid the estimates to be affected by the acceleration, see Supplementary Table S4; (4) the background rate  $\theta$  is then optimized individually for each 62 sequences, so to reproduce similar numbers of foreshocks in our simulations as with the real sequences; (5) 100 independent realisations of this model are run for each 62 sequences over 3 years. We only keep the first 100 simulations in which the largest earthquake (the 'mainshock') has magnitude  $m > 6.5$ , occurs at least 6 months after the start of the simulations, and then select the foreshocks as all m  $>$  2.5 earthquakes that occur within 6 months prior to the mainshock.

Supplementary Fig. S14 shows that the numbers of foreshocks over 6 months typically reproduce the real numbers (computed for  $-365 \le t \le -182.5$  days prior to the mainshock). The stacked cumulative number of foreshocks clearly exhibits an acceleration (Supplementary Fig. S15c,d), that is however much less pronounced as with real data. Moreover, since the ETAS parameters (hence the clustering properties) of interplate and intraplate differ, the resulting accelerations also differ, but the strongest one is found for the intraplate earthquakes. This is in contradiction with our observation. Finally, the probability P of obtaining an acceleration at least as strong in the case of a purely random time series is 0.496 and 0.359 (median values) for the interplate and intraplate populations, respectively. This is well above the 0.123 median value obtained for the interplate earthquakes (Fig. 4c). We therefore conclude that, although the clustering of earthquakes causes a purely statistical acceleration of seismicity before the mainshocks, it contributes little to the actual acceleration.

#### **6. Spatial distribution of foreshocks**

Fig. 4e shows the projected location on the plate interface of the last shock of the 22 interplate sequences with last day events (Fig. 2) relatively to the main shock hypocenter. Error bars are not

drawn because they are difficult to estimate. For the JMA located events, which constitute the majority of the subduction earthquakes, a value of the location error in latitude, longitude and depth is provided in the catalog. Once projected on the plate interface, the largest error for the 10 JMA located events of Fig. 4e is about 3km (the error in depth, which is the largest reported error does not produce a large projected error because the dip of the plate interface is generally small). However, these reported errors are based on RMS residuals at the stations, so they are only a lower bound to the real errors which can be much larger.

The fault plane orientation, used to locally define the plate interface, is taken from the published centroid moment tensor solution (HRVD or GCMT, www.globalcmt.org).

25. Sauber, J., Thatcher, W., Solomon, S.C. & Lisowski, M. Geodetic slip rate for the eastern California shear zone and the recurrence time of Mojave Desert earthquakes. *Nature* **367**, 264-266 (1994).



Supplementary Figure S1: Determination of the magnitude of completeness of the interplate and intraplate seismicity datasets. Blue circles show the number of occurrences of events of a given magnitude in each set. Data come from different catalogs and the reported magnitudes may slightly differ from the NEIC-USGS reported values listed in Supplementary Tables S1-S3. The red line indicates the inferred magnitude of completeness. The black line with slope b shows the corresponding Gutenberg-Richter distribution.



1 year

Supplementary Figure S2: Evolution of the cumulative number of seismic events in the year preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.





Supplementary Figure S3: Evolution of the cumulative number of seismic events in the 6 months preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.





Supplementary Figure S4: Evolution of the cumulative number of seismic events in the 3 months preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.



Supplementary Figure S5: Evolution of the cumulative number of seismic events in the 2 months preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.



Supplementary Figure S6: Evolution of the cumulative number of seismic events in the 30 days preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.





10 days

Supplementary Figure S7: Evolution of the cumulative number of seismic events in the 10 days preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.



5 days

Supplementary Figure S8: Evolution of the cumulative number of seismic events in the 5 days preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.



Supplementary Figure S9: Evolution of the cumulative number of seismic events in the day preceding some interplate earthquakes. Numbers identify earthquakes in Supplementary Table S1. The amplitudes and lengths of all the traces are the same but curves may be shifted for easier reading. Each curve ends just prior to the earthquake.





Supplementary Figure S10: Comparison of the normalized stacks of the cumulative numbers of events of the interplate sequences for 3 different values of the radius of the zone in which the events are counted. Each sequence is given the same weight.

**Intraplates** 



Supplementary Figure S11: Evolution of seismic activity (cumulative number of events) over the 4 days prior to all the intraplate earthquakes of the dataset. All the curves with events are normalized to their total number. Red circles indicate the largest event of the period. Trace numbers identify earthquakes in Supplementary Table S2.



Supplementary Figure S12: Probability that the acceleration of seismicity observed is not due to chance. **a**, For the 6 months period immediately preceding the interplate earthquakes. **b**, For the 6 months period beginning 1 year and ending 6 months before the interplate earthquakes.



Supplementary Figure S13: Stack of the cumulative numbers of seismic events in the 5 days preceding a randomly chosen event in each interplate sequence. Each seismic sequence is given the same weight.



Supplementary Figure S14: Real and median (over 100 simulations) numbers of  $m \geq 2.5$  preearthquake events over a 6 month period for the 62 sequences.



Supplementary Figure S15: Normalized stacks of the cumulative numbers of events prior to the interplate and intraplate earthquakes of the dataset (red curves), compared to the corresponding ETAS simulation (blue curves). **a,b**, Evolution over 150 days. **c,d**, Zoom over the last 10 days.





1) USGS National Earthquake Information Center, Significant Earthquakes, Tectonic Summary: <http://earthquake.usgs.gov/earthquakes/eqinthenews/>

- 2) USGS National Earthquake Information Center, Significant Earthquakes: [http://neic.usgs.gov/eq\\_depot/](http://neic.usgs.gov/eq_depot/)
- 3) Alaska Earthquake Information Center:<http://www.aeic.alaska.edu/quakes/>
- 4) International Seismological Center (ISC):<http://www.isc.ac.uk/>
- 5) Harvard Centroid Moment Tensor Catalog
- 6) Global Centroid Moment Tensor Catalog (GCMT): http://www.globalcmt.org/
- 7) Vasudevan, K., Eaton, D.W. & Iverson, A. Did the November 17, 2009 Queen Charlotte Island earthquake fill a predicted seismic gap? *Am. Geophys. Union, Fall Meeting*, S43A-2036 (2010).
- 8) Hauksson, E, Jones, L.M. & Hutton, K. The 1999 Mw 7.1 Hector Mine, California, earthquake sequence : Complex conjugate strike-slip faulting. *Bull. Seism. Soc. Am*. **92**, 1154-1170 (2002).
- 9) Hauksson, E. *et al*. The 2010 Mw 7.2 El Mayor-Cucapah earthquake sequence, Baja California, Mexico and southernmost California, USA: Active seismotectonics along the Mexican Pacific margin. *Pure Appl. Geophys.* **168***,* 1255 (2011).

**Supplementary Table S1** : List of the 31 interplate earthquakes considered from the USGS National Earthquake Information Center catalog. Hypocentral latitude, longitude and depth and moment magnitude are given for each event. m is the largest foreshock magnitude inferred for the 25 sequences of Fig. 3a. The location of the events is shown in Fig. 1. Events are numbered according to longitude. Earthquakes 1 to 22 are on subduction boundaries while those from 23 to 31 occur on transform boundaries.





- 1) USGS National Earthquake Information Center, Significant Earthquakes, Tectonic Summary: <http://earthquake.usgs.gov/earthquakes/eqinthenews/>
- 2) USGS National Earthquake Information Center, Significant Earthquakes: [http://neic.usgs.gov/eq\\_depot/](http://neic.usgs.gov/eq_depot/)
- 3) Alaska Earthquake Information Center:<http://www.aeic.alaska.edu/quakes/>
- 4) International Seismological Center (ISC):<http://www.isc.ac.uk/>
- 5) Harvard Centroid Moment Tensor Catalog
- 6) Global Centroid Moment Tensor Catalog (GCMT): http://www.globalcmt.org/
- 7) Kao, H. & Chen, W.P. The Chi-Chi earthquake sequence : Active out-of-sequence thrust faulting in Taiwan. *Science* **288**, 2346-2349 (2000).
- 8) Angelier, J., Chu, H.T. & Lee J.C. Shear concentration in a collision zone: Kinematics of the Chihshang fault as revealed by outcrop-scale quantification of active faulting, Longitudinal Valley, eastern Taiwan. *Tectonophysics* **274**, 117-143 (1997).
- 9) Ohmi, S. et al. The 2000 Western Tottori earthquake : Seismic activity revealed by the regional seismic networks. *Earth Planets Space* **54**, 819-830 (2002).
- 10) Toda, S, Stein, R.S. & Sagiya, T. Evidence from the 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nature* **419**, 58-61 (2002).
- 11) Ratchkovski, N.A. & Hansen R.A. Sequence of strong intraplate earthquakes in the Kodiak island region, Alaska in 1999-2001. *Geophys. Res. Lett*. **28**, 3729-3732 (2001).

**Supplementary Table S2** : List of the 31 intraplate earthquakes considered from the USGS National Earthquake Information Center catalog. Hypocentral latitude, longitude and depth and moment magnitude are given for each event. The location of the events is shown in Fig. 1. Events are numbered according to longitude.



**Supplementary Table S3** : List of the 10 M≥6.5 earthquakes not included in the study. They are all aftershocks of earthquakes considered in Supplementary Tables S1 or S2 and occur on the same day or following days. The i and p letters denote interplate and intraplate events respectively.



**Supplementary Table S4 :** Model parameters for the two sets of earthquakes, for m≥2.5