ARTICLE

https://doi.org/10.1038/s41586-019-1258-4

Surface erosion events controlled the evolution of plate tectonics on Earth

Stephan V. Sobolev[1,2]* & Michael Brown[3]

Plate tectonics is among the most important geological processes on Earth, but its emergence and evolution remain unclear. Here we extrapolate models of present-day plate tectonics to the past and propose that since about three billion years ago the rise of continents and the accumulation of sediments at continental edges and in trenches has provided lubrication for the stabilization of subduction and has been crucial in the development of plate tectonics on Earth. We conclude that the two largest surface erosion and subduction lubrication events occurred after the Palaeoproterozoic Huronian global glaciations (2.45 to 2.2 billion years ago), leading to the formation of the Columbia supercontinent, and after the Neoproterozoic 'snowball' Earth glaciations (0.75 to 0.63 billion years ago). The snowball Earth event followed the 'boring billion'—a period of reduced plate tectonic activity about 1.75 to 0.75 billion years ago that was probably caused by a shortfall of sediments in trenches—and it kick-started the modern episode of active plate tectonics.

Plate tectonics is a key geological process operating on Earth, shaping its surface, strongly influencing its deep structure and making it unique among the rocky planets in the Solar System. How plate tectonics emerged and which factors controlled its subsequent evolution are widely discussed questions, but the answers remain uncertain and the proposals controversial[1,2].

We acknowledge that secular cooling of the Earth is an important underlying factor[3–5]. However, in this study we challenge the popular view that secular cooling has been the only major control on the evolution of plate tectonics on Earth since about 3 billion years (Gyr) ago. We first discuss the main controls on variations in contemporary plate velocities and the stability of subduction. This allows us to formulate the hypothesis that the emergence and evolution of plate tectonics on Earth was related to the rise of the continents and the accumulation of sediments at continental edges, and subsequently in trenches, to lubricate and stabilize subduction since the mid-Mesoarchaean (about 3 Gyr ago). To test this hypothesis, we use geological and geochemical datasets. We focus in particular on two major surface erosion events following the Palaeoproterozoic Huronian (2.45–2.2 Gyr ago) and Neoproterozoic snowball Earth (0.75–0.63 Gyr ago) glaciations. We argue that these events provided an increased sediment supply to lubricate subduction, which had a crucial role in reactivating plate tectonics after a period of reduced subduction in each case, first after the Siderian tectono-magmatic lull[6] and then after the boring billion[4].

Controls on contemporary plate tectonics

Since the 1960s it has been widely accepted that mantle convection has an important role in driving lithospheric plates on Earth. However, it took some time to realize that plate tectonics was a special type of convection that was possible only because the plate boundaries are mechanically very weak. A key factor in the first successful model of present-day plate motions driven by mantle convection[7], which was not explicitly discussed by the authors, was an assumption of null shear stress at plate boundaries. The necessity of low strength at plate boundaries and particularly along the plate interfaces in subduction, hereafter called subduction channels, has been discussed from different points of view in many subsequent studies[8–13].

In contrast to mid-ocean ridges, subduction channels are relatively cold because surface rocks enter the channel and depress the isotherms. At low temperatures, the shear strength of the plate boundary is mainly controlled by brittle deformation (that is, effective friction; see Methods) rather than by ductile deformation (that is, viscous flow), which is the reason why great earthquakes rupture subduction channels to depths of several tens of kilometres[14]. The friction-controlled part of the subduction channel extends deeper than the base of the seismogenic zone, which is typically located at a depth of about 40–50 km in most subduction channels[14], and that is where the highest interface stress is achieved (see Methods). Using a numerical model of plate motions on present-day Earth (see Methods for a description of the model), we obtain results from a series of experiments with different effective friction coefficients at the convergent plate boundaries (Fig. 1a, b). A reasonable fit of the model results to observed velocities is achieved if the average friction coefficient at convergent plate boundaries is about 0.03, whereas for a friction coefficient of 0.1 the plate velocities appear to be too low (Fig. 1b). We note that a friction coefficient of 0.03 is much lower than the typical friction coefficient of dry rocks measured in the laboratory[15] (0.6–0.8). The conventional and very natural explanation of this discrepancy is that friction in subduction channels is lowered by the high pressure of pore fluid that results from the compaction of fluid-rich materials in the subduction channel.

Effects of sediments and mantle temperature

The lubricating effect of sediments in subduction channels was recognized more than thirty years ago[16]. Subsequently, Lamb and Davis[17] used an analysis of the present-day mechanical equilibrium between the South America and Nazca plates to argue that in the Southern Andes unconsolidated sediments filling the trench and subduction channel led to a much lower friction coefficient (<0.03) than in the Central Andes, where sheared rocks cannibalized from the overriding plate by the subducting slab increased the friction coefficient to about 0.1. These authors attributed the difference in surface topography and deformation of the upper plate between the Southern and Central Andes to this difference in interface friction. Geodynamic modelling of the Cenozoic evolution of the Central and Southern Andes[18] confirmed this

[1]GFZ German Research Centre for Geosciences, Section of Geodynamic Modeling, Potsdam, Germany. [2]Institute of Geosciences, University of Potsdam, Potsdam, Germany. [3]Laboratory for Crustal Petrology, Department of Geology, University of Maryland, College Park, MD, USA. *e-mail: stephan@gfz-potsdam.de

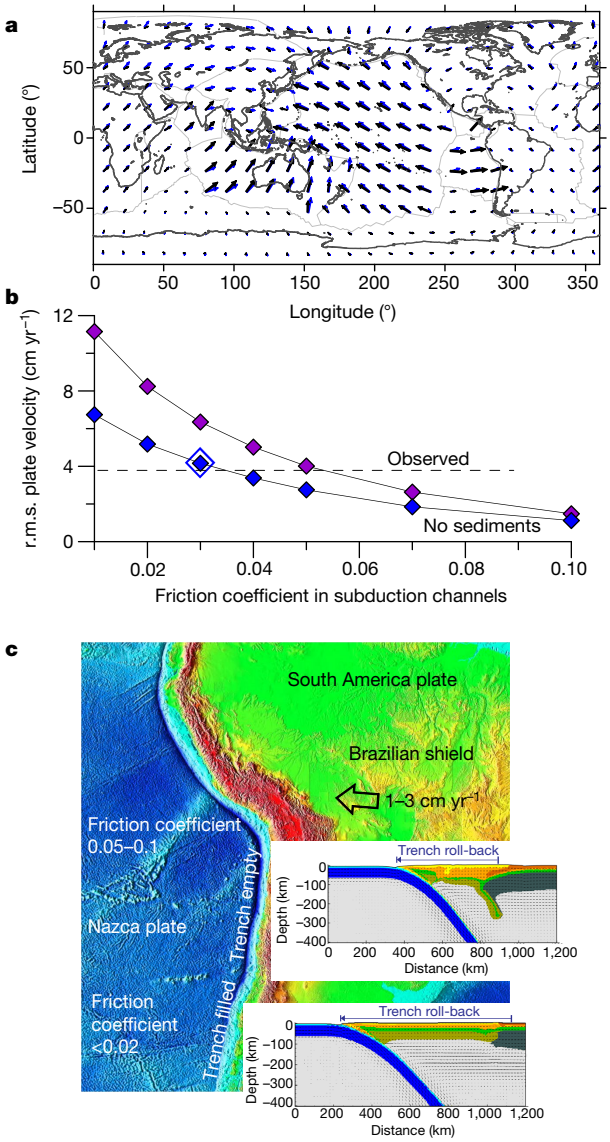


Fig. 1 | Global and regional models showing the effect of sediments on contemporary subduction. a, Observed plate velocities from the NUVEL 1A model in a no-net-rotation reference frame (black arrows) versus computed velocities (blue arrows) for the global model (Methods) with a friction coefficient of 0.03 at convergent boundaries and of 0.1 at divergent and transform boundaries (large blue diamond in **b**). **b**, Root-mean-square (r.m.s.) values of computed plate velocities in the global model versus friction coefficients at convergent plate boundaries. Blue (and magenta) diamonds correspond to model results without (and with) consideration of slab pull at depths less than 300 km. The expected true values are closer to the blue diamonds (see Methods). The horizontal dashed line shows the r.m.s. value of observed plate velocities. **c**, Models of subduction-related orogenesis in the Central and Southern Andes, with a higher friction coefficient in the subduction channel lacking sediments in the Central Andes and a lower friction coefficient in the subduction channel filled with sediments in the Southern Andes. The westward drift of the South America plate during the last 30 Myr was almost entirely compensated by roll-back of the Nazca plate trench in the southern Andes (trench filled with sediments, lower friction) and was partially compensated by deformation of the South America plate in the central Andes (no sediments in the trench, higher friction).

explanation by showing that the difference in the friction coefficient of the subduction channel from the Central Andes (>0.05 and probably up to 0.1; see Methods) to the Southern Andes (<0.02) indeed might have been one reason for the different shortening rates and deformation styles in these regions (Fig. 1c). Based on the results of this modelling, we infer that the effective friction coefficient in a subduction channel without unconsolidated sediments is around 0.07–0.1, whereas the friction coefficient in a subduction channel fully lubricated by sediments is less than 0.02. The likely reason for such a low effective friction coefficient in the subduction channel is the efficient compaction of fine-grained wet sediments that preserve highly pressurized pore fluids. With increasing depth in the subduction channel, the fluid is first structurally bound in hydrous minerals, which generally have low shear strength, and then is partially released by successive dehydration reactions that further reduce material strength. Importantly, the modelling shows that such a low effective friction coefficient and low material strength in the subduction channel do not preclude the occurrence of great subduction earthquakes[19], such as the Great Chile Earthquake of 1960. We note that the viscosity of metasedimentary rocks that reach the deeper, ductile part of the subduction channel is low and consequently such rocks also act as a lubricant in the deeper part of the channel[20]. Thus, sediments and metasedimentary rocks lubricate the subduction channel along its full length.

The effect of sediments in subduction channels on plate velocities can be inferred from the results presented in Fig. 1b. The root-mean-square value of present-day plate velocities is about 4 cm yr^{-1}. If all subduction channels were free of sediments, the velocities would be less than half of this value. Behr and Becker[20] obtained a similar result by assuming that deformation in the subduction channel is controlled by ductile flow and considering that the viscosity of metasedimentary rocks is much lower than the viscosity of metabasic rocks. If such low plate velocities operated for about 100 million years (Myr), the density heterogeneity of the mantle, which is mostly controlled by the amount of material returned by subduction, would be reduced. In turn, this would reduce the driving force of plate tectonics and further lower plate velocities (see Methods for details).

In addition, lubrication by sediments could enable both formation of new subduction zones and extension of existing subduction zones. This is consistent with successful numerical models of subduction initiation, which always assume very low (often null) yield strength at an initial fault and very low friction of material dragged into the developing plate boundary[21,22].

A number of modelling studies have demonstrated that subduction should have been less stable at the higher mantle temperatures predicted for early Earth compared to the present[13,23] (see Methods). In one particular study[13], subduction is shown to be stable and continuous at mantle temperatures up to 100 K higher than the present if shear stress at the subduction interfaces is negligible (equivalent to a fully lubricated subduction channel), whereas at temperatures 200–300 K higher than the present (typical of the Archaean) subduction becomes less stable, with frequent slab break-offs, especially at the higher end of this temperature range. If shear stress at the interface is 30 MPa (corresponding to a reduced amount of sediment in the subduction channel), subduction is stable only at present-day mantle temperatures[13]. Therefore, we conclude that for early-Earth conditions, subduction of weaker slabs would have been more sensitive to the lubricating effects of sediments than contemporary subduction (see Methods for more details).

Testing the hypothesis

On the basis of the above discussion we infer that continental sediments in subduction channels act as a lubricant for subduction and that the presence of these sediments in trenches is a necessary condition for the stable operation of plate tectonics, particularly earlier in Earth's evolution, when the mantle was warmer and slabs were relatively weak. Once plate tectonics had emerged on Earth, this hypothesis predicts that a fully linked global network of plate boundaries and periods of stable plate tectonics should follow widespread surface erosion events, whereas times of diminished surface erosion would lower the volume of sediment in trenches and should be associated with reduced subduction and possibly intermittent plate tectonics. We test these predictions using geological proxies believed to identify plate tectonic activity

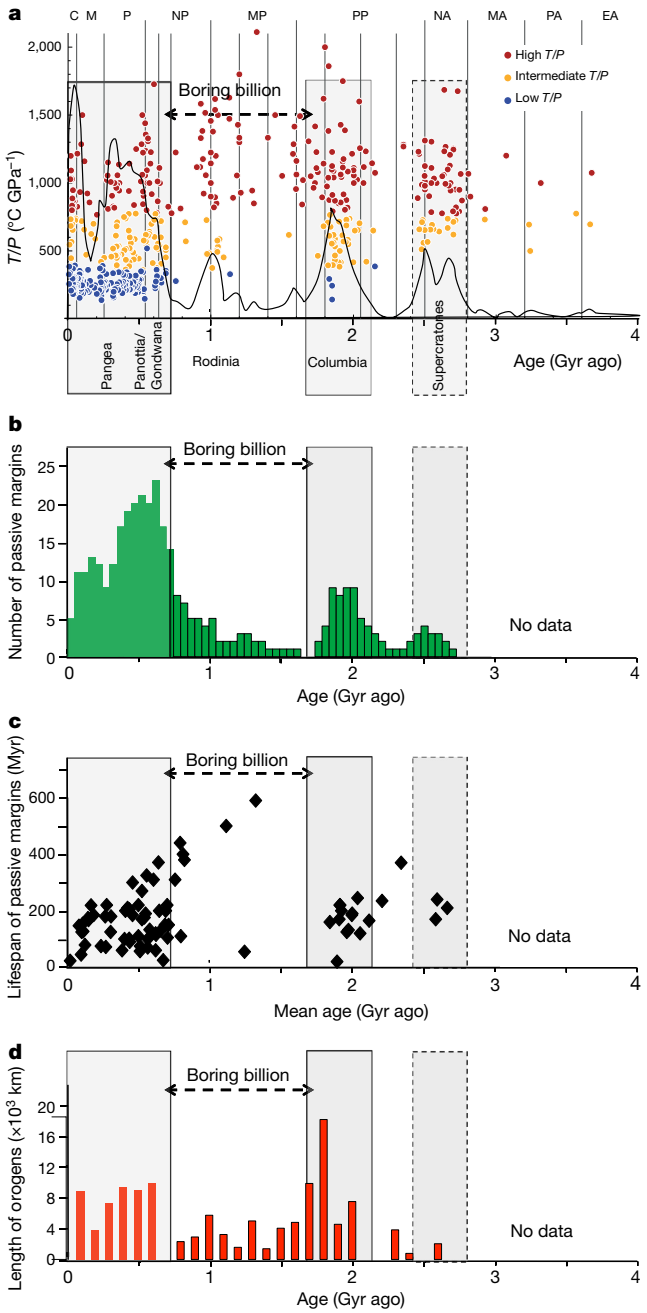


Fig. 2 | Geological proxies for subduction and plate tectonic activity. a, Probability of occurrence (black curve) and thermobaric ratios (metamorphic temperature (*T*) to pressure (*P*); solid circles) of metamorphic rocks[5]. Grey rectangles show periods of enhanced subduction and increased plate tectonic activity probably corresponding to supercontinent assembly[4,5] (solid outlines), and the period of relatively less prominent activity forming supercratons in the Archaean (dashed outline). The abbreviations at the top refer to the geological eras (C, Cenozoic; M, Mesozoic; P, Palaeozoic; NP, Neoproterozoic; MP, Mesoproterozoic; PP, Paleoproterozoic; NA, Neoarchaean; MA, Mesoarchaean; PA, Paleoarchaean; EA, Eoarchaean). **b**, **c**, Cumulative number of ancient passive margins in 50-Myr bins (**b**) and the lifespan of passive margins versus mean age[25] (**c**). **d**, Cumulative length of orogens[26,27]. Also indicated is the long period of decreased plate tectonic activity (the boring billion[4]).

(Fig. 2) and geochemical proxies that trace the influence of the continental crust on the composition of seawater and continental sediments in the source of subduction-related magmas (Fig. 3a, b).

The global appearance of paired, intermediate- and high-thermal-gradient metamorphism[5] (yellow and red symbols in Fig. 2a) and evidence of a 1,000-km-scale relative motion of cratons[24] in the Neoarchaean indicate that plate tectonics had emerged on Earth by this time. Since the Mesoarchaean there have been three principal periods of amplified subduction activity, corresponding to the times of the assembly of supercratons in the Neoarchaean (relatively less pronounced) and the supercontinents Columbia and Pannotia/Gondwana in the Palaeoproterozoic and Neoproterozoic, respectively. These three periods are identified by the clustered occurrence of metamorphic rocks[5] (Fig. 2a), the notable rise in the number of passive margins and their decreasing duration[25] (Fig. 2b, c) and by the higher cumulative length of orogens[26,27] (Fig. 2d). The clustered occurrence of metamorphic rocks indicates collisional orogenesis in these periods[5], which requires extensive motion of continental plates that cannot happen without stable subduction. The increase in the number of passive margins is closely related to subduction activity, which creates passive margins along the trailing edges of continents[25]. The same is true for the origin of the passive margins associated with back-arc basins, which are related to subduction retreat, and possibly also for continental break-up in general[28]. Therefore, the increasing number of passive margins and their decreasing lifespan (Fig. 2b, c) probably indicate enhanced subduction activity, whereas the smaller number of long-lived passive margins in the intervening periods signifies lowered subduction activity. The higher cumulative length of orogens also points to more frequent collisions of continents and higher forces available for plate tectonics, again probably associated with increased subduction activity.

The period of highly active subduction that culminated in the formation of Columbia (Fig. 2a) was followed by the boring billion (1.75–0.75 Gyr ago)[4]. The boring billion was a time of reduced subduction[4,5] and limited development of passive margins[25]. We note that the peak in the occurrence of metamorphic rocks for the transition from Columbia to Rodinia about 1 Gyr ago relates mostly to Grenvillian orogenic events at the margin of Columbia and limited internal basin opening and closing, rather than a major break-up and reorganization of the supercontinent[5,27]. Therefore, we do not consider this peak to be an indication of high global subduction activity, in agreement with other studies (for example, ref. [4]). The boring billion was followed by the period of modern plate tectonics (Fig. 2a).

In Fig. 3 we show proxies for the crustal and mantle influence on ocean chemistry and magma composition, in particular variations with time in the strontium (Sr) isotope composition of seawater[29] (blue curve in Fig. 3a, b), and the hafnium (Hf)[30] (red curve in Fig. 3a, b) and O (magenta curve) isotope composition[31] of detrital zircons. These variations record differences in the amount of continental weathering that contributed to the composition of seawater and the availability of continental sediments in the source of magmas that make up the continental crust.

There are three time intervals in which Sr and O consistently show an increase and Hf a decrease, as shown by the transparent green rectangles in Fig. 3a. We relate these three periods to increased erosion of the continents and a larger volume of sediments in trenches. At these times, the normalized $^{87}Sr/^{86}Sr$ in seawater rises while εHf decreases and $\delta^{18}O$ increases in detrital zircons (Fig. 3a; see figure legend and Methods for definition of parameters). In relation to the hypothesis proffered above, these three periods correspond to the timing of lubrication events in pre-existing and/or newly initiated subduction channels that aid the smooth operation of the plate tectonic machine. Interestingly, the first (about 2.8–2.7 Gyr ago) and second (about 2.3–2.1 Gyr ago) periods follow major glaciations, and the first period also follows the rise of continents above sea level[32–34]. The third period coincides with the Neoproterozoic snowball Earth glaciations (about 0.75–0.63 Gyr ago[35]). It is noteworthy that each of the proxies shown in Fig. 3a increases in magnitude with successive events; this suggests that the second and third events each had a larger impact than the one before, which may be a response to the secular decline in mantle temperature and increasing volume of continental crust. Furthermore, the snowball Earth glaciations are associated with a global crustal erosion and sediment subduction event of unprecedented scale, as recently demonstrated by a new modelling approach to the Hf and O isotope composition of detrital zircons[36].

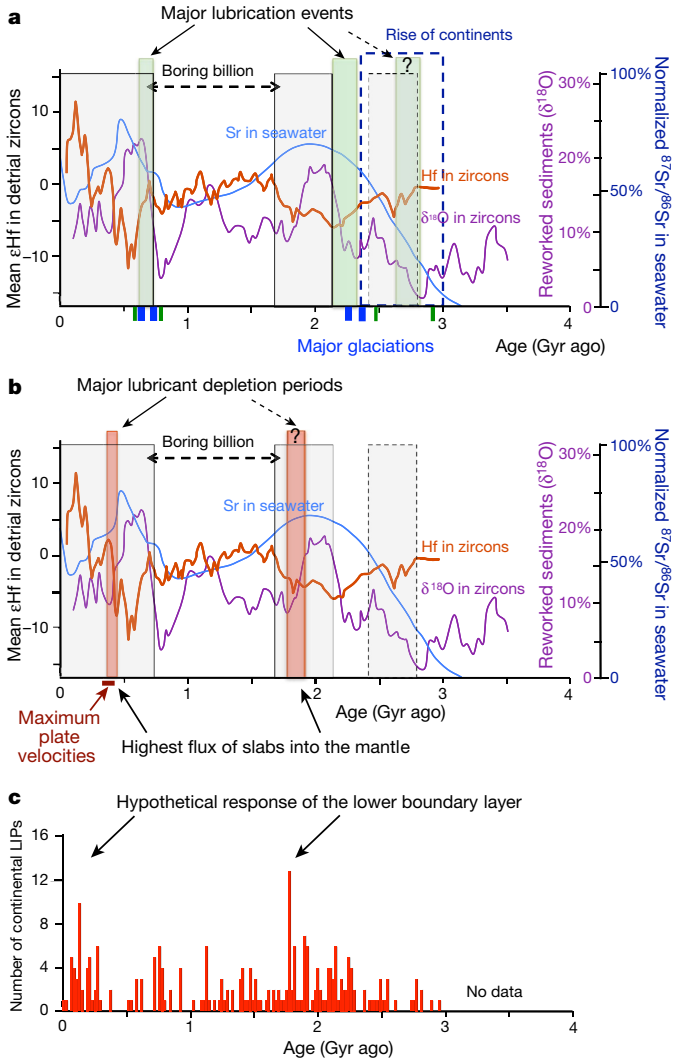


Fig. 3 | Geodynamic interpretation of geochemical proxies for recycling of sediments. a, **b**, Proxies for the effects of continental erosion on seawater composition[29] (normalized $^{87}Sr/^{86}Sr$ in seawater, based on binary mixing of river runoff versus mantle influence[29]; blue curves) and continental sediments in the source of subduction-related magmas (εHf, where $\varepsilon Hf = [(^{176}Hf/^{177}Hf)_t/(^{176}Hf/^{177}Hf)_{chondrites} - 1] \times 10^4$ and t is the age of the sample; red curves) and $\delta^{18}O$ (where $\delta^{18}O = [(^{18}O/^{16}O)_t/(^{18}O/^{16}O)_{seawater} - 1] \times 10^3$); magenta curves) in detrital zircons[30,31]). Grey rectangles show times of plate tectonic activity probably corresponding to supercontinent assembly[4,5] (solid outlines) and the formation of supercratons in the Archaean (dashed outlines). The blue dashed rectangle in **a** shows the most likely time of the rise of the continents[32–34]. The blue and green solid rectangles along the age axis in **a** show the times of major global (blue) and regional (green) glaciations[35]. Transparent green rectangles in **a** show the time intervals in which all three proxies consistently indicate increasing crustal influence, here interpreted as major lubrication events related to increased continental erosion. Transparent red rectangles in **b** show the time intervals in which all three proxies consistently indicate decreasing crustal influence, here interpreted as periods of extensive sediment subduction and lubricant depletion. The highest flux of slabs into the mantle occurs during the periods of lubricant depletion, indicated by the two black arrows. **c**, Number of LIPs versus age (in 20-Myr bins; data from a previous work[27]). The hypothetical response of the lower boundary layer to the highest flux of slabs into the mantle is indicated by the two black arrows.

We emphasize that the first of these three periods of enhanced surface erosion and increased lubrication directly preceded the first period of amplified subduction activity, which created several Neoarchaean supercratons[37]. The two more recent periods each occurred at the beginning of a new cycle of amplified subduction activity leading to the formation of a supercontinent (Columbia and Pannotia/Gondwana, respectively) after a period of reduced subduction (first in the Siderian[6] and then during the boring billion[4]). This is also in agreement with the prediction of the hypothesis proffered above.

These three geochemical proxies indicate a reduction in the influence of continental crust or sediments on the composition of the oceans or in magma sources during two periods, one in the Palaeoproterozoic and another in the Cambrian–Ordovician (red transparent rectangles in Fig. 3b). We interpret these as periods of lubricant depletion, that is, times when the amount of continental erosion and the volume of sediments in trenches were generally decreasing owing to the reduced elevation of mature orogenic belts and the effect of accelerated subduction depleting the supply of continental sediments faster than they could be replaced. In the Palaeozoic, this interpretation is confirmed by the well constrained period of maximal plate velocities about 0.3–0.4 Gyr ago[26,38,39], that is, at the time close to the second period of lubricant depletion. This maximum was followed by a reduction in plate velocities, a decrease in the frequency of metamorphic rocks (Fig. 2a) and a reduction in the cumulative length of orogens (Fig. 2d). In the Proterozoic, the boring billion was also a time of reduced subduction that followed the first period of lubricant depletion. The boring billion lasted until the snowball Earth glaciations and the ensuing major surface erosion event provided an increased supply of sediments to lubricate subduction.

The periods of lubricant depletion should also represent periods of maximal flow of subducted slabs into the mantle (Fig. 3b), which should produce a response from the mantle, albeit after a time lag. One hypothesis[40], supported by several numerical modelling studies[41,42], suggests that some large igneous provinces (LIPs) were related to plumes that originated from the edges of deep lower-mantle heterogeneities because of the interaction of descending slabs with the lower boundary layer of mantle convection. Interestingly, a compilation of LIP ages (used in a previous work[27]) does show an increased frequency of LIPs 100 Myr after the Palaeoproterozoic lubricant depletion period and about 300 Myr after the Palaeozoic lubricant depletion period (Fig. 3c). The longer delay of the response in the late Neoproterozoic–early Palaeozoic compared to the Palaeoproterozoic is to be expected because of the increasing viscosity of the mantle with cooling.

As discussed above, the subduction lubrication events that followed the global glaciations probably had a profound effect on plate tectonic activity and possibly also on deep-mantle convection. We note that the two periods when Earth had global glaciations are separated by a 1.5-Gyr gap in which evidence of glaciation at any latitude is lacking[35]. This observation suggests that it is the absence of global glaciations between the Huronian (2.45–2.2 Gyr ago) and the snowball Earth (0.75–0.63 Gyr ago) events that was responsible for the 1,500-Myr variation in Earth's geodynamic evolution recently discussed[27]. Lastly, we suggest that an increased volume of sediments available for the lubrication of subduction channels not only stabilized subduction on Earth and subsequently stimulated periods of increased activity, but also could have contributed to the emergence of plate tectonics in the first place.

The tectonic regime that preceded plate tectonics

Important evidence about the transition to a plate tectonic regime comes from systematic variations in Hf and O isotope compositions of zircons of different ages, which reveal the relative proportions of reworked and juvenile crust through time[43]. On the basis of these data, it has been argued that before 3.0 Gyr ago, the production rate of continental crust was higher and the recycling rate of continental crust was lower than after 3.0 Gyr ago[43]. Because subduction is believed to be the main mechanism of recycling continental crust[44], the lower recycling rate before 3.0 Gyr ago is interpreted as evidence of the absence of large-scale subduction[43].

This earlier geodynamic regime could have involved intense deformation of mechanically weak lithosphere caused by dripping of

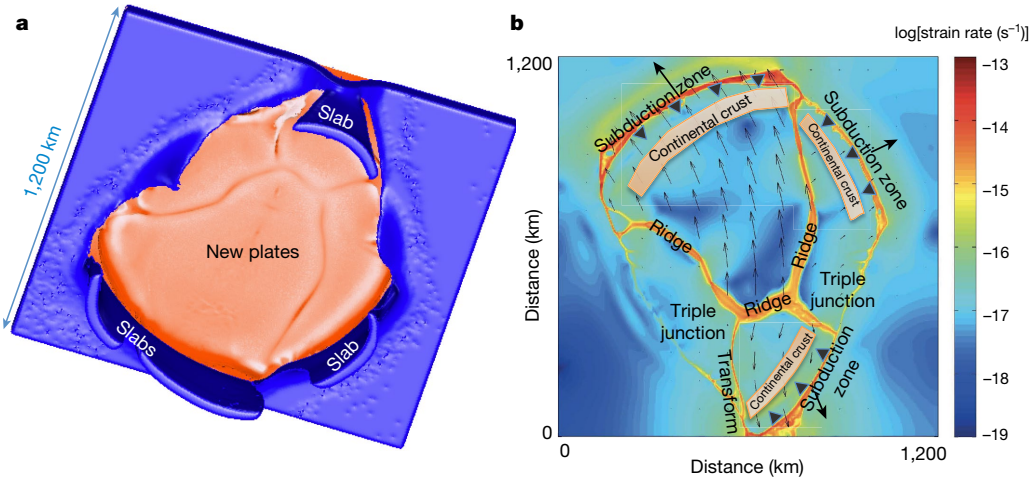


Fig. 4 | Plume-induced retreating subduction generating a regional 'plate tectonics' cell[48]. **a**, View from the deep mantle of the subducting slabs and the opening 'plate tectonics' cell. **b**, View of the 'plate tectonics' cell from above, showing that it includes all types of plate boundaries, even though all subducting slabs in the cell are retreating. Collisions of such cells could have produced protocontinents and led to the emergence of global plate tectonics. The blue arrows represent vectors of plate velocities and the black arrows show the directions and magnitudes of the retreat of the slabs.

eclogitized lower crust and delamination of the underlying mantle lithosphere[23,45,46]. One alternative for this convection regime is a type of episodic lid overturn and resurfacing[12] due to retreating large-scale subduction triggered by mantle plumes[47,48] or meteoritic impacts[49]. These retreating subduction zones would have formed regional cells of plate-like behaviour (Fig. 4), within which all types of plate boundaries are present, even though all subducting slabs around the cell are retreating. Retreating subduction would not lead to much erosion of the overriding plates, in agreement with the low rate of crustal recycling in the geochemical model[43]. Simultaneously, retreating slabs would bring water into the upwelling, hot asthenospheric mantle via the sinking hydrated upper crust of the slabs, enabling a large volume of magma to be generated. This is an efficient way to produce the early, more mafic continental crust. Multiple collisions of such cells and included arcs could have formed protocontinents, again in agreement with the model of crustal evolution derived from the zircon record[43].

Following a previous work[50], and assuming that the underlying mantle lithosphere was weak, we expect that the emergent protocontinental crust could have had sufficient gravitational potential energy to extend. Therefore, as sediments accumulated at the edges of the protocontinents and were subsequently overridden by the extending protocontinental margins, nascent subduction channels were formed along the edges of the protocontinents. Thus, a global plate tectonic regime could have evolved from a plume-induced retreating subduction regime, which already included the elements of sea-floor spreading and transform faults inside the 'plate tectonic' cells.

Summary of hypothesis

Before about 3 Gyr ago, when the mantle was probably at its hottest and not much sediment was available in the oceans, the tectonic regime could have been of the 'squishy lid' type[23,45,46] or of plume- (or impact-) induced 'retreating subduction' type[47,48] (Fig. 5). Between 3 Gyr ago and 2 Gyr ago this tectonic regime gradually evolved into a global plate tectonic regime enabled by the rise of the continents and increased surface erosion. The first period of major erosion that led to an increased volume of sediments available for lubrication of subduction probably followed the late Mesoarchaean glaciation; this enabled the formation of supercratons from protocontinents. After a period of reduced subduction during the Siderian, the assembly of Columbia was triggered by the second period of major surface erosion that again led to an increased volume of sediments available for lubrication, which followed the early Palaeoproterozoic (Huronian) glaciations. The diminished delivery of sediments in the trenches during the boring billion led to reduced subduction globally and a period of decreased plate tectonic activity. The continuing secular cooling of the mantle and the unprecedented scale of surface erosion following the snowball Earth glaciations brought the boring billion to an end, initiated the contemporary geodynamic cycle with continuous global plate tectonics and probably triggered the Cambrian explosion of life on Earth.

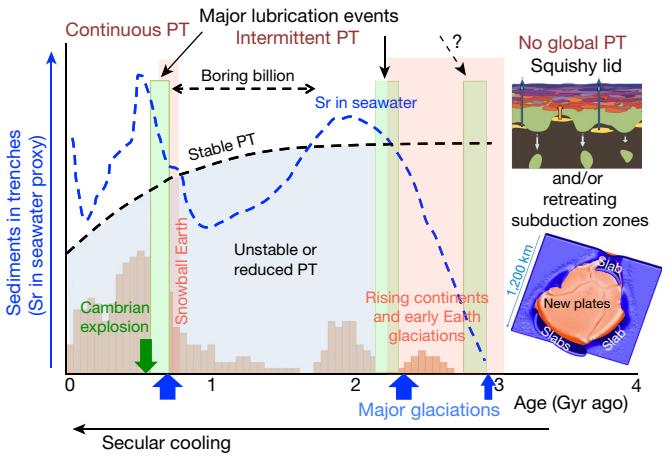


Fig. 5 | Summary of the factors that control the emergence and evolution of plate tectonics on Earth. The limited sediment supply and hotter mantle in the early Archaean, before 3 Gyr ago, did not allow global plate tectonics (PT). Instead, a 'squishy lid' regime or a regime of plume- or impact-induced retreating subduction (or a combination of the two) was dominant. Enhanced surface erosion due to the rise of the continents and major glaciations in the Mesoarchaean and Palaeoproterozoic, and particularly during the Neoproterozoic snowball Earth 0.75–0.63 Gyr ago, resulted in a large influx of sediments into trenches (lubrication events), which—together with the secular cooling of the Earth's mantle—stabilized subduction and plate tectonics. Lack of sediments in the oceans during the boring billion[4] between 1.75 and 0.75 Gyr ago caused unstable behaviour of plate tectonics on Earth. Blue arrows mark major glaciations[35]; the blue dashed curve shows the mantle-normalized seawater $^{87}Sr/^{86}Sr$ ratio[29], a proxy for the influence of continental erosion on seawater composition; and the black dashed curve separates hypothetical domains of stable and unstable plate tectonics. The reddish domain shows the number of passive margins[25], here used as a proxy for the intensity of plate tectonics. Semi-transparent red rectangles show the time range of snowball glaciations in the Neoproterozoic and the rise of the continents and glaciations in the Mesoarchaean and Paleoproterozoic.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1258-4.

Received: 17 December 2018; Accepted: 24 April 2019;
Published online 5 June 2019.

1. Hawkesworth, C. J. & Brown, M. Earth dynamics and the development of plate tectonics. *Phil. Trans. R. Soc. A* **376**, 20180228 (2018).
2. Stern, R. J. The evolution of plate tectonics. *Phil. Trans. R. Soc. A* **376**, 20170406 (2018).
3. Herzberg, C., Condie, K. & Korenaga, J. Thermal history of the Earth and its petrological expression. *Earth Planet. Sci. Lett.* **292**, 79–88 (2010).
4. Cawood, P. A. & Hawkesworth, C. J. Earth's middle age. *Geology* **42**, 503–506 (2014).
5. Brown, M. & Johnson, T. Secular change in metamorphism and the onset of global plate tectonics. *Am. Mineral.* **103**, 181–196 (2018).
6. Spencer, C. J. et al. A Palaeoproterozoic tectono-magmatic lull as a potential trigger for the supercontinent cycle. *Nat. Geosci.* **11**, 97–101 (2018).
7. Ricard, Y. & Vigny, C. Mantle dynamics with induced plate tectonics. *J. Geophys. Res.* **94**, 17543–17559 (1989).
8. Zhong, S. & Gurnis, M. Viscous flow model of a subduction zone with a faulted lithosphere: long and short wavelength topography, gravity and geoid. *Geophys. Res. Lett.* **19**, 1891–1894 (1992).
9. Tackley, P. J. Self-consistent generation of tectonic plates in three-dimensional mantle convection. *Earth Planet. Sci. Lett.* **157**, 9–22 (1998).
10. Moresi, L. & Solomatov, V. Mantle convection with a brittle lithosphere: thoughts on the global tectonic style of the Earth and Venus. *Geophys. J. Int.* **133**, 669–682 (1998).
11. Bercovici, D. The generation of plate tectonics from mantle convection. *Earth Planet. Sci. Lett.* **205**, 107–121 (2003).
12. O'Neill, C. et al. Episodic Precambrian subduction. *Earth Planet. Sci. Lett.* **262**, 552–562 (2007).
13. van Hunen, J. & van den Berg, A. P. Plate tectonics on the early Earth: limitations imposed by strength and buoyancy of subducted lithosphere. *Lithos* **103**, 217–235 (2008).
14. Scholz, C. H. Earthquakes and friction laws. *Nature* **391**, 37–42 (1998).
15. Byerlee, J. Friction of rocks. *Pure Appl. Geophys.* **116**, 615–626 (1978).
16. Shreve, R. T. & Cloos, M. Dynamics of sediment subduction, melange formation, and prism accretion. *J. Geophys. Res.* **91**, 10229–10245 (1986).
17. Lamb, S. & Davis, P. Cenozoic climate change as a possible cause for the rise of the Andes. *Nature* **425**, 792–797 (2003).
18. Sobolev, S. V. & Babeyko, A. Y. What drives orogeny in the Andes? *Geology* **33**, 617–620 (2005).
19. Sobolev, S. V. & Muldashev, I. Modelling seismic cycles of great megathrust earthquakes across the scales with focus at postseismic phase. *Geochem. Geophys. Geosyst.* **18**, 4387–4408 (2017).
20. Behr, W. M. & Becker, T. W. Sediment control on subduction plate speeds. *Earth Planet. Sci. Lett.* **502**, 166–173 (2018).
21. Gurnis, M., Hall, C. & Lavier, L. Evolving force balance during incipient subduction. *Geochem. Geophys. Geosyst.* **5**, Q07001 (2004).
22. Baes, M. & Sobolev, S. V. Mantle flow as a trigger for subduction initiation: a missing element of the Wilson Cycle concept. *Geochem. Geophys. Geosyst.* **18**, 4469–4486 (2017).
23. Sizova, E. et al. Generation of felsic crust in the Archean: a geodynamic modeling perspective. *Precambr. Res.* **271**, 198–224 (2015).
24. Cawood, P. A. et al. Geological archive of the onset of plate tectonics. *Phil. Trans. R. Soc. A* **376**, 20170405 (2018).
25. Bradley, D. C. Passive margins through earth history. *Earth Sci. Rev.* **91**, 1–26 (2008).
26. Condie, K. C. A planet in transition: the onset of plate tectonics on Earth between 3 and 2 Ga? *Geosci. Front.* **9**, 51–60 (2018).
27. Li, Z. X. et al. Decoding Earth's rhythms: modulation of supercontinent cycles by longer superocean episodes. *Precambr. Res.* **323**, 1–5 (2019).
28. Dal Zilio, L. et al. The role of deep subduction in supercontinent breakup. *Tectonophysics* **746**, 312–324 (2018).
29. Shields, G. A. A normalised seawater strontium isotope curve: possible implications for Neoproterozoic-Cambrian weathering rates and the further oxygenation of the Earth. *eEarth* **2**, 35–42 (2007).
30. Cawood, P. A., Hawkesworth, C. J. & Dhuime, B. The continental record and the generation of continental crust. *Geol. Soc. Am. Bull.* **125**, 14–32 (2013).
31. Spencer, C. J., Roberts, N. M. W. & Santosh, M. Growth, destruction, and preservation of Earth's continental crust. *Earth Sci. Rev.* **172**, 87–106 (2017).
32. Flament, N., Coltice, N. & Rey, P. F. A case for late-Archaean continental emergence from thermal evolution models and hypsometry. *Earth Planet. Sci. Lett.* **275**, 326–336 (2008).
33. Korenaga, J., Planavsky, N. J. & Evans, D. A. D. Global water cycle and the coevolution of the Earth's interior and surface environment. *Phil. Trans. R. Soc. A* **375**, 20150393 (2017).
34. Bindeman, I. N. et al. Rapid emergence of subaerial landmasses and onset of a modern hydrologic cycle 2.5 billion years ago. *Nature* **557**, 545–548 (2018).
35. Hoffman, P. F. & Schrag, D. P. The snowball Earth hypothesis: testing the limits of global change. *Terra Nova* **14**, 129–155 (2002).
36. Keller, C. B. et al. Neoproterozoic glacial origin of the Great Unconformity. *Proc. Natl Acad. Sci. USA* **116**, 1136–1145 (2019).
37. Bleeker, W. The late Archean record: a puzzle in ca. 35 pieces. *Lithos* **71**, 99–134 (2003).
38. Domeier, M. & Torsvik, T. H. Plate tectonics in the late Paleozoic. *Geosci. Front.* **5**, 303–350 (2014).
39. Matthews, K. J. et al. Global plate boundary evolution and kinematics since the late Paleozoic. *Global Planet. Change* **146**, 226–250 (2016).
40. Torsvik, T. H., Smethurst, M. A., Burke, K. & Steinberger, B. Large igneous provinces generated from the margins of the large low-velocity provinces in the deep mantle. *Geophys. J. Int.* **167**, 1447–1460 (2006).
41. Tan, E., Leng, W., Zhong, S. & Gurnis, M. On the location of plumes and lateral movement of thermochemical structures with high bulk modulus in the 3-D compressible mantle. *Geochem. Geophys. Geosyst.* **12**, Q07005 (2011).
42. Steinberger, B. & Torsvik, T. H. A geodynamic model of plumes from the margins of Large Low Shear Velocity Provinces. *Geochem. Geophys. Geosyst.* **13**, Q01W09 (2012).
43. Dhuime, B. et al. A change in the geodynamics of continental growth 3 billion years ago. *Science* **335**, 1334–1336 (2012).
44. Clift, P. & Vannucchi, P. Controls on tectonic accretion versus erosion in subduction zones: implications for the origin and recycling of the continental crust. *Rev. Geophys.* **42**, RG2001 (2004).
45. Johnson, T. E. et al. Delamination and recycling of Archaean crust caused by gravitational instabilities. *Nat. Geosci.* **7**, 47–52 (2014).
46. Rozel, A. B. et al. Continental crust formation on early Earth controlled by intrusive magmatism. *Nature* **545**, 332–335 (2017).
47. Arndt, N. & Davaille, A. Episodic Earth evolution. *Tectonophysics* **609**, 661–674 (2013).
48. Gerya, T. V. et al. Plate tectonics on the Earth triggered by plume-induced subduction initiation. *Nature* **527**, 221–225 (2015).
49. O'Neill, C. et al. Impact-driven subduction on the Hadean Earth. *Nat. Geosci.* **10**, 793–797 (2017).
50. Rey, P. F., Coltice, N. & Flament, N. Spreading continents kick-started plate tectonics. *Nature* **513**, 405–408 (2014).

Acknowledgements We acknowledge comments by N. Arndt, A. Sobolev and colleagues from the Geodynamic Modeling Section in GFZ: A. Babeyko, S. Brune and B. Steinberger. We thank W. Behr for comments that prompted us to strengthen the arguments presented in the Article. We are grateful to R. Ernst who provided the LIP dataset and to Z.-X. Li who provided the orogen dataset[27]. S.V.S. is grateful to R. Stern, who brought to his attention the problems of the origin and evolution of plate tectonics. Deep Carbon Observatory supported S.V.S.'s participation in the Workshop on the Origin and Evolution of Plate Tectonics, Locarno, Switzerland in 2016, where he first presented the hypothesis discussed in the paper.

Reviewer information *Nature* thanks Whitney Behr and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.V.S. conceived the study, suggested the hypothesis, designed and computed the models and produced the figures. S.V.S. and M.B. interpreted the data and wrote the paper.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1258-4.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to S.V.S.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Strength of rocks in the subduction channel. *Contemporary subduction zones.* The frictional deformation domain extends below the down-dip limit of the seismogenic zone[14]. The average down-dip depth for this limit in all subduction zones is 46 km (from the extended dataset of subduction zones[51]). Below this depth there is still a friction-controlled domain related to a rate-strengthening mechanism of friction, which extends well within the creeping part of the subduction channel, down to the point of the brittle–ductile transition. The shear stress reaches a maximum at this point, beyond which the shear stress decreases exponentially with depth in the ductile deformation region. According to this conventional point of view, deformation in the upper 50 km (by depth) or so of a typical subduction channel is controlled by friction and the mean shear stress is highest at this depth. This is illustrated in Extended Data Fig. 1 using the recently published cross-scale model (that is, over several timescales) of the subduction process that reproduces the seismic cycles of the great earthquake in southern Chile[19], where the largest recorded earthquake (magnitude $M = 9.5$) occurred in 1960. In this model a friction coefficient of 0.015 is assumed, corresponding to the case where the channel is fully lubricated with sediments. The figure shows the modelled distributions of stress and strain rate (Extended Data Fig. 1a, b) during the inter-seismic phase of the seismic cycle, 320 years after the great earthquake. It is seen that the maximum shear stress in the channel (that is, the brittle–ductile transition) is achieved at a depth of about 52 km, where the temperature is about 400 °C (Extended Data Fig. 1a). That is well within the creeping part of the channel (Extended Data Fig. 1b), the top of which is at a depth of about 35 km, where the temperature in the channel is close to 310 °C.

In the model discussed above and in previous models of Andean subduction[18], in addition to friction, a weak ductile rheology is used in the channel (wet quartz rheology in the model[19]), even if it is not filled with sediments. This procedure is based on the assumption that in the absence of sediments, the channel consists of sheared felsic rocks eroded from the upper plate crust and/or of serpentinites eroded from the upper plate mantle. If that is not the case (as in the model of Behr and Becker[20]), the difference in viscosity between metasedimentary rocks and metabasic rocks further increases the lubrication effect of the sediments.

Application to early-Earth conditions. According to the analytical solution of corner flow for a Newtonian viscous fluid[52], the slab interface temperature is scaled to ($T_{mantle} - T_{surface}$). Assuming $T_{mantle} = 1{,}550$ °C for early Earth, recent $T_{mantle} = 1{,}350$ °C and $T_{surface} = 0$ °C (in both cases), and holding other subduction parameters the same, we obtain that early-Earth subduction channels were only about 15% (1,550/1,350 = 1.15) hotter than present-day subduction channels. This difference is several times smaller than the temperature difference of the subduction interfaces of presently active subducting slabs[53]. Therefore, we do not expect any substantial change in the strength of subduction channels on early Earth compared to contemporary Earth.

Consequences for subduction instability in early Earth. Even if the strength of subduction channels on early Earth were similar to their present-day strength, the subduction might proceed differently because of the differences in the rheology and buoyancy of the slabs. On the basis of the discussion in the main text, we conclude that for early-Earth conditions, subduction of weaker slabs would have been more sensitive to the lubricating effects of sediments than contemporary subduction. A greater propensity to break-off would also be expected owing to the different buoyancy structure of the subducting slabs on early Earth. With the thicker oceanic crust expected for early-Earth conditions[3], above the gabbro–eclogite transformation (that is, in the upper 50–100 km, depending on the kinetics of the reaction) subducting slabs would have been less negatively buoyant than contemporary slabs, but more negatively buoyant at depths below this transformation[13]. In this case, slab pull from the part of the slab below 50–100 km would have been higher than for contemporary slabs, but resistance of the uppermost buoyant part of the slab would also have been higher. In such a configuration, slab break-off would have been easier if the shear resistance of the subduction channel was increased owing to a lack of sediments for lubrication.

For a typical contemporary subduction channel, a friction coefficient of 0.07–0.1 translates to a yield strength of about 50–70 MPa, assuming that the brittle part of the channel extends to a depth of 50 km. On the basis of the results of numerical modelling[12], the calculated yield strength of 50 MPa is above the critical value that separates mobile and stagnant lid styles of convection (20–30 MPa), especially at the higher mantle heat flow expected in early Earth. This result suggests a high probability of either a stagnant-lid or an episodic mobile-lid type of convection on Earth at times when trenches were deprived of unconsolidated continental sediments. A similar conclusion was reached in another numerical modelling study[54], in which it was shown that if the friction coefficient in the lithosphere at potential plate boundaries was higher than a critical value of 0.1–0.2, then subduction should not occur on Earth. In addition, that study demonstrated that the evolving mantle and core structures in the models were only consistent with observations for Earth for very low friction coefficients, below 0.04. These conclusions, based on numerical modelling studies, are consistent with the hypothesis that the rise of continents and the accumulation of sediments at continental edges and in trenches has lubricated subduction since the mid-Mesoarchaean.

Model of global plate tectonics. *Approach and model setup.* The modelling is based on the observation that the largest heterogeneities in the composition and physical properties of rocks occur in about the upper 300 km of Earth, which is also where rocks have the most complex nonlinear rheologies[55]. Therefore, in the upper 300 km of the spherical shell of the model we solve the full system of mass, momentum and energy conservation equations using nonlinear elasto-visco-plastic rheology, with the parameters of the rheological models constrained by laboratory experiments. To solve the equations we use the finite-element code SLIM3D[56]. The shell has a true free surface (no sticky air or other approximation), which allows precise modelling of the surface topography. At depths greater than 300 km in the model we use a simplified numerical technique and a radial distribution of the viscosity[57]. The two modelling codes are coupled at a depth of 300 km using an iterative scheme that guarantees continuity of tractions and velocities at a depth of 300 km.

The full details of the modelling technique have been published recently[58]; therefore, here we describe only the features of the model specific to this study, as follows. (1) In considering the rheology of the asthenosphere we take into account underestimation of the water content of olivine in the experiments (by a factor of 3)[59]; therefore, in the rheological model of the asthenospheric mantle we use parameters for olivine with 300 p.p.m. H/Si, that is, about 3 times lower than that expected in the asthenosphere[59]. (2) We treat convergent and other plate boundaries separately. For the other plate boundaries, a friction coefficient of 0.1 is assumed, but for the convergent plate boundaries the friction coefficient is a parameter that is varied from 0.01 to 0.1. (3) For the continents, we use a three-dimensional thermal structure in a 300-km shell with the thickness of the lithosphere based on a thermal model[60]; below the sub-continental lithosphere we assume a potential temperature (Tp) of 1,300 °C. For the oceans, we use a plate model relating thermal structure to ocean age[61], with Tp = 1,300 °C at the asthenosphere. (4) For the mantle deeper than 300 km we employ a three-dimensional density distribution based on the history of subduction[62]. In some models we also use density heterogeneities due to slabs within the 300-km shell. (5) All models have a lateral resolution of about 100 km and a vertical resolution that varies from 20 to 50 km in the 300-km shell, and 124 orders of spherical harmonics in the spectral code deeper than 300 km. (6) All models are calculated for 500,000 yr, which ensures that a quasi-steady state is achieved. In all other aspects, including predefined plate boundaries, the model is the same as in the recently published study[58].

Short-term effect of change in friction. We conducted two sets of experiments, and in each set we varied only the effective friction coefficient at the convergent plate boundaries. In the first set of experiments (blue diamonds in Fig. 1b) we did not consider the density heterogeneity related to slabs in the 300-km shell. Thus, in this set of experiments we ignored the pull of the slab segments in the upper 300 km. We note that owing to the low lateral resolution in all models we also ignored the subduction-resisting slab-bending force, which has a magnitude of about 1/3 of the pull force from the slab segments in the upper mantle and transition zone[63]. As the hanging slab segment in the 300-km shell is about 200–300 km in length—that is, about 1/3 of the slab segment is in the upper mantle and the transition zone—the related slab pull force has about the same magnitude as the slab bending force, but with the opposite sign, so these forces should mutually balance. This balance justifies the results of the first set of experiments, in which both forces were ignored. In Fig. 1a we compare velocities computed in the experiments with a friction coefficient of 0.03 at the convergent plate boundaries, with the no-net-rotation data from the NUVEL 1A model[64]. In the second set of experiments (magenta diamonds in Fig. 1b) the density heterogeneity related to slabs in the 300-km shell was included. In these experiments the resisting forces were underestimated, but the pulling forces were included; therefore, we consider that the second set of experiments predict the upper limit of plate velocities.

Long-term effect of change in friction. The results shown in Fig. 1b demonstrate the effect of friction at plate boundaries on the plate velocities when the mantle driving force is fixed. However, if plate velocities were reduced owing to an absence of lubricant in trenches and a corresponding increase in the friction coefficient, then less cold material would have been subducted into the mantle, which in turn would have reduced the slab pull and mantle suction forces[63]. With a typical subduction velocity of 2–3 cm yr^{-1}, the slab sinks through the entire mantle in about 100–150 Myr, so 100–150 Myr is the characteristic time required for the density heterogeneity in the mantle to be renewed. Therefore, the long-term effect of change in the friction coefficient in subduction channels may be much larger than the short-term effect illustrated in Fig. 1b.

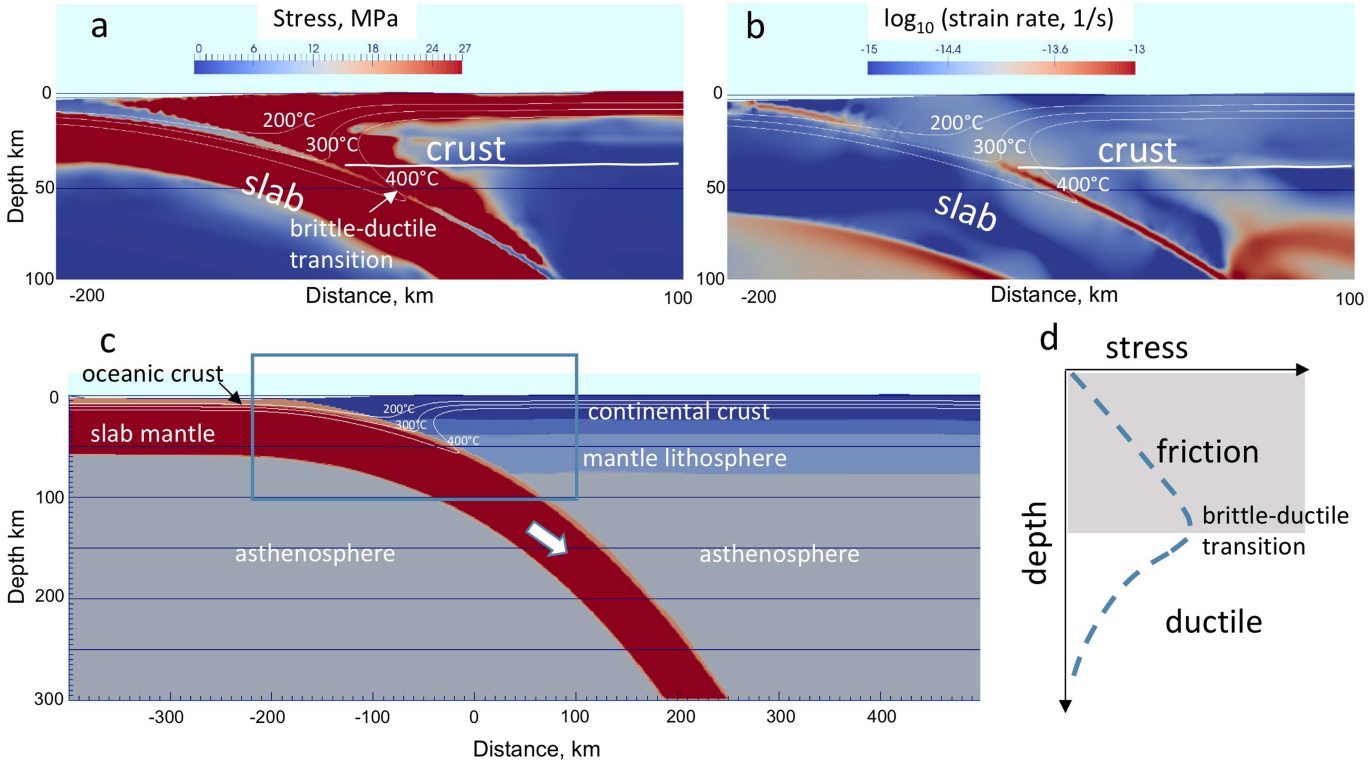
Interpretation of regional models of subduction orogenesis in the Andes. In the model of subduction orogenesis in the Andes[18], the present-day rate of advance of the South American plate towards the trench was assumed to be 3 cm yr^{-1}, which appears to be a substantial overestimation according to recent reconstructions[65]. Therefore, the model[18] underestimated the friction coefficient required to drive

shortening in the Central Andes. Taking this underestimation into account, we infer that the friction coefficient in the subduction channel without unconsolidated sediments should be considerably higher than the value of 0.05 estimated previously[18], perhaps closer to 0.07–0.1 according to the present-day estimate of force balance[17].

Data and code availability

The executable file, as well as all input and output files, used in the calculations of the global plate velocities are available on request.

51. Heuret, A., Lallemand, S., Funiciello, S., Piromallo, C. & Faccenna, C. Physical characteristics of subduction interface type seismogenic zones revisited. *Geochem. Geophys. Geosyst.* **12**, Q01004 (2011).
52. England, P. & Wilkins, C. A. simple analytical approximation to the temperature structure in subduction zones. *Geophys. J. Int.* **159**, 1138–1154 (2004).
53. Syracuse, E. M., van Keken, P. E. & Abers, G. A. The global range of subduction zone thermal models. *Phys. Earth Planet. Inter.* **183**, 73–90 (2010).
54. Nakagawa, T. & Tackley, P. J. Influence of plate tectonic mode on the coupled thermochemical evolution of Earth's mantle and core. *Geochem. Geophys. Geosys.* **16**, 3400–3413 (2015).
55. Bürgmann, R. & Dresen, G. Rheology of the lower crust and upper mantle: evidence from rock mechanics, geodesy, and field observations. *Annu. Rev. Earth Planet. Sci.* **36**, 531–567 (2008).
56. Popov, A. A. & Sobolev, S. V. SLIM3D: a tool for three-dimensional thermo mechanical modeling of lithospheric deformation with elasto-visco-plastic rheology. *Phys. Earth Planet. Inter.* **171**, 55–75 (2008).
57. Steinberger, B. & Calderwood, A. Models of large-scale viscous flow in the Earth's mantle with constraints from mineral physics and surface observations. *Geophys. J. Int.* **167**, 1461–1481 (2006).
58. Osei Tutu, A. et al. Evaluating the influence of plate boundary friction and mantle viscosity on plate velocities. *Geochem. Geophys. Geosyst.* **19**, 642–666 (2018).
59. Hirth, G. & Kohlstedt, D. L. in *Inside the Subduction Factory* (ed. Eiler, J.) 83–105 (American Geophysical Union, 2004).
60. Artemieva, I. Global thermal model TC1 for the continental lithosphere: implications for lithosphere secular evolution. *Tectonophysics* **416**, 245–277 (2006).
61. Parsons, B. & Sclater, J. G. An analysis of the variation of ocean floor bathymetry and heat flow with age. *J. Geophys. Res.* **82**, 803–827 (1977).
62. Steinberger, B. Slabs in the lower mantle – results of dynamic modelling compared with tomographic images and the geoid. *Phys. Earth Planet. Inter.* **118**, 241–257 (2000).
63. Wu, B. et al. Reconciling strong slab pull and weak plate bending: the plate motion constraint on the strength of mantle slabs. *Earth Planet. Sci. Lett.* **272**, 412–421 (2008).
64. DeMets, C., Gordon, R. G. & Argus, D. F. Geologically current plate motions. *Geophys. J. Int.* **181**, 1–80 (2010).
65. Oncken, O., Boutelier, D., Dresen, G. & Schemmann, K. Strain accumulation controls failure of a plate boundary zone: linking deformation of the Central Andes and lithosphere mechanics. *Geochem. Geophys. Geosyst.* **13**, Q12007 (2013).



Extended Data Fig. 1 | Strength in the subduction channel. a–c, The subduction channel of the seismic cycle model for the southern Andes[19]. Shown are the stress distribution (**a**) and the strain rate distribution (**b**) during the inter-seismic phase of the seismic cycle, 320 years after the great earthquake, as well as a magnified image of the subduction channel (**c**). Temperature isolines are shown in degrees Celsius, and the location of the maximum shear stress in the channel is given as a proxy for the brittle–ductile transition. **d**, Sketch of the stress distribution inside the subduction channel, showing the friction-controlled (brittle) and ductile deformation domains.