Dear referee,

We are glad to receive the review report and would like to express our sincere thanks to you and the editors. Without the constructive comments from you, the quality of this manuscript cannot be significantly improved. All the comments and suggestions have been carefully considered to revise the manuscript. Detailed reply to all comments and the associate manuscript modifications are given below.

Reply on RC1

1. The language of the manuscript needs to be fully checked and revised by a professional editing service or a native speaker

Reply: *We appreciate your advice and have revised the manuscript carefully to improve the language, with the help of all the co-authors.*

2. In this study, the relation between spreading rate and age of oceanic lithosphere is ignored. Usually, higher spreading rates create younger lithospheres at a constant distance. In this study, the authors assumed that the lithospheric age is constant near the side boundaries (50 Myr). As a result, by imposing some higher velocities near the side boundaries to simulate higher spreading rates, the lithosphere becomes under extension and since the ridge is the weakest point in the system the width of ridge changes (it is clearly seen in e.g., Figs 3e,4e and 5a-b); higher rates lead to wider ridges. Could the authors explain to what extend is this assumption realistic? I think the formation of cracks in the lithosphere is the consequence of this assumption

Reply: *We appreciate this comment which triggers us a lot of thinking. Indeed, the initial temperature distribution of the oceanic plate consists of the half-space cooling model and thermal equilibrium part. The half-space cooling model is used to describe the oceanic plate younger than 50Myr, and the thermal equilibrium structure is used to describe older oceanic parts. The thermal equilibrium thickness of the older lithosphere is constant (i.e., ~100 km; corresponding to a thermal age of 50 Ma). We further checked our model results, especially the stress state in the whole lithosphere. The result* *shows that the lithosphere seems under extension owing to the high internal velocities (Figs. S1, S4). Actually, the "tension cracks" used in the main text may be inaccurate. There is no weakening or plastic deformation in the lithosphere, and there are no normal faults near the surface in the models. We describe the stress distribution in the lithosphere to highlight the stress localization, which occurs in the lithosphere where plume flows to the ridge. As a result, we revised and simplified the description of the "cracks" in the main text (lines 212-213 and lines 249-250).*

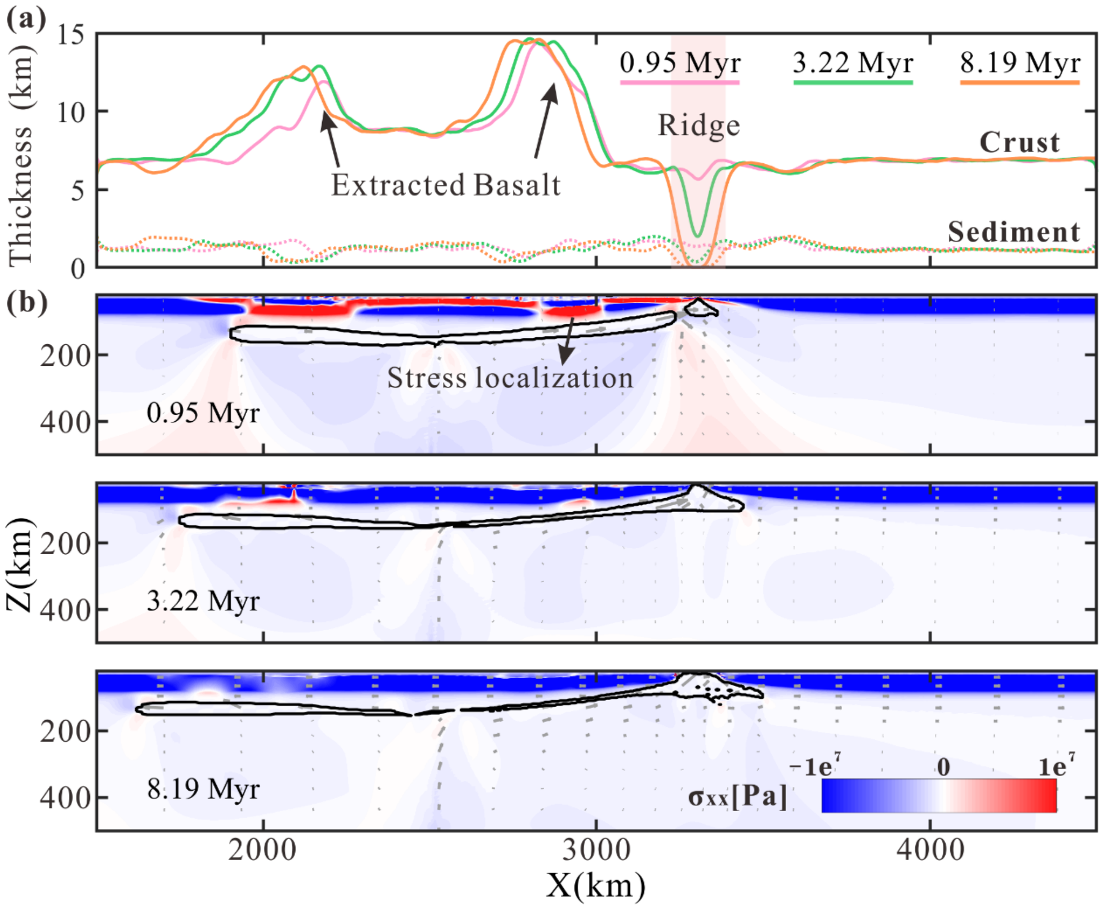

Figure S1. Reference model (M12, see Table S1, same as Fig. 3) evolution of ridge-ward plume flow shown by (a) crust and sediment thickness, (b) normal stress. The mantle plume weakens the overlying oceanic plate and changes the stress state of the overlying oceanic plate. Molten plume material beneath the lithosphere is extracted to the crust.

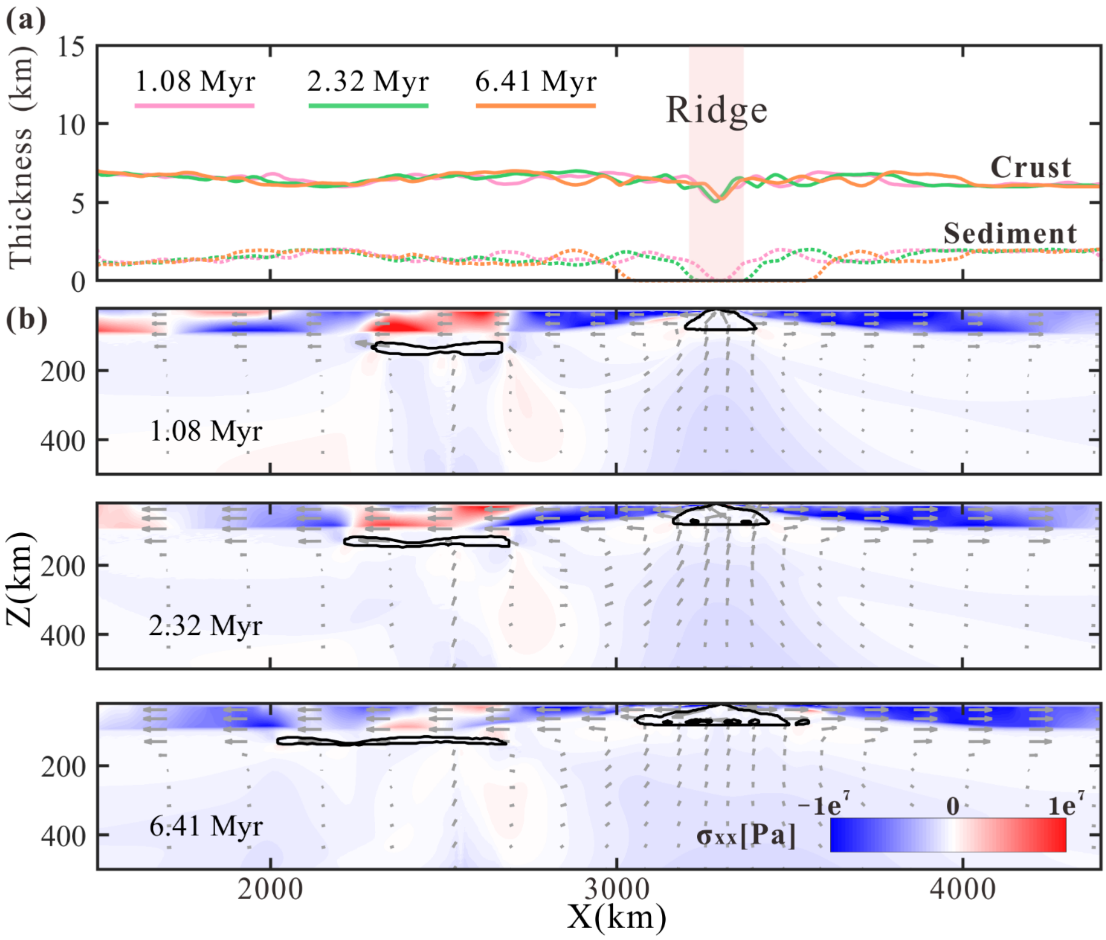

Figure S4. Reference model (M77, see Table S1, same as Fig. 4) evolution of plate-drag plume flow shown by (a) crust and sediment thickness, (b) normal stress. The mantle plume weakens the overlying oceanic plate and changes the stress state of the overlying oceanic plate. Molten plume material beneath the lithosphere is extracted to the crust.

3. In the abstract it is written "plume migration driven by plate drag is promoted by fast-ridge spreading rate." This is true only if the plume radii are small. For large plumes the rate of spreading is irrelevant (Fig. 6). This should be mentioned here and also in discussion and conclusions.

Reply: *We agree with you and rephrased this sentence (lines 19-20). Indeed, plate dragging is most significant when the plume buoyancy is relatively small. When the plume is buoyant enough, plate drag plays a minor role than the plume self-spreading* *on plume-ridge interaction. We have revised this part and the discussion section in the manuscript.*

4. Usually decomposition melting of plume head causes the formation of a plateau above the plume head. Where do plateaus form in the models? I suggest that the authors add information about where plateaus form and how thick the crust is to the manuscript. The temporal evolution of plateaus is also interesting to be investigated.

Reply: *We appreciate this suggestion and add figures to present the temporal evolution of the extracted melt, displaying the crust thickness in the model over time (see figure S1, S4 in the supplementary material). Mantle plumes melt beneath the lithosphere and are then extracted into the oceanic crust, converting into basalt to form a thickened oceanic crust (Specific mechanism described in method section). In the model, oceanic plateaus (thickened crust) are formed directly above the spreading plume head. To describe plateau formation in our models, we amended the text in lines 213-215 and 251-252.*

5. Lines 20-22: "Our results highlight fast-spreading ridges exert strong plate dragging force, rather than suction on plume motion, which sheds new light on the natural observations of plume absence along the fast-spreading ridges, such as the East Pacific Rises." As I indicated above this is true only if plume radii are smaller than 250 km (based on Fig. 6). This conclusion implies that plumes in the Pacific are smaller than those in Atlantic. Are there any observations supporting this? I'm interested in a discussion about this issue in the paper.

Reply: *Thank you for your comments. We revised this sentence (lines 24-25). Observations show approximate plume buoyancy flux distributions in the Pacific and Atlantic (Figure 1 below). There is no obvious correlation between the distribution of buoyancy of mantle plumes in different oceans with their spreading rate. We conclude that plume size is not the deciding factor to explain the difference between the Pacific and Atlantic in terms of plume-ridge interaction mode. We add a figure and discussion about this in the discussion section (lines 470-481). Please also see our reply to* *comment 6.*

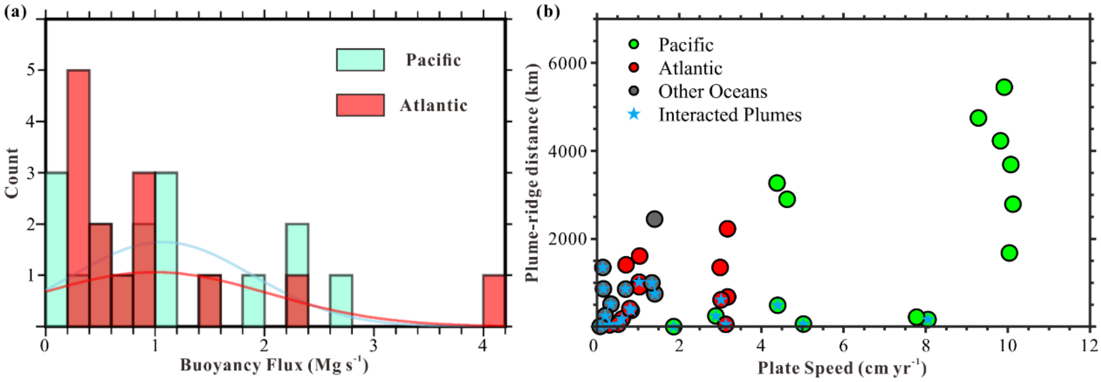

Figure 1. Buoyancy flux, plate speed and plume-ridge distance of mantle plumes in different oceans. (a) Histogram of plume buoyancy flux distributed in the Pacific and Atlantic. The guassian distribution curves are shown in light blue and red lines, respectively. (b) The plot of plate speed at each plume and their off-axis distance. Blue stars mark the plumes shown to be interacted with the nearby ridges.

6. Looking at distribution of plumes and their sizes in Fig.9, one cannot see any correlation between plume size with plate drag and ridge suction. Can authors comment on that? Besides, in conclusion it is written: “The plume size, that is, the plume buoyancy flux, may play a critical role in controlling the connection between the two units, compared with distance and spreading rate.” Why does plume size play important role compared to the two other factors? This is not discussed in the main text

Reply: *Thanks for your comments. We suggest that all three factors play an important role (Fig. 8). The predicted effects of plate velocity and plume-ridge distance are fully consistent with observations (Fig. 11b). The effects of plume buoyancy flux are less obvious when compared to observations (Fig. 11a). For discussion, see lines 470-481. We also reconsidered the importance of different influence factors and rephrased the sentence in the conclusion.*

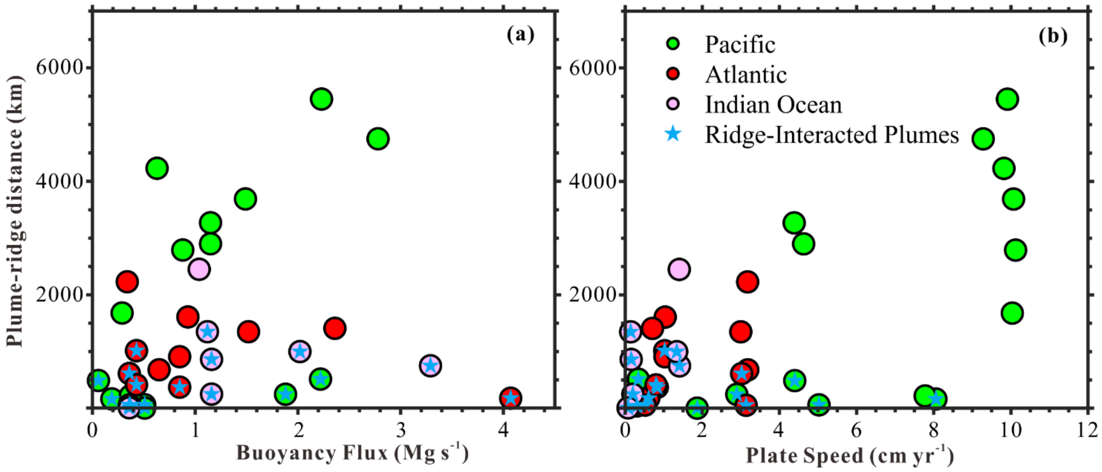

Figure 11. Buoyancy flux, plate speed and plume-ridge distance of mantle plumes in different oceans. Mantle plumes in the Pacific, Atlantic and Indian Ocean are shown in green, red and pink circles, respectively. Blue stars marked the ridge-interacted plumes according to Ito et al. (2003). **(a)** Plot of plume-ridge distance and plume buoyancy flux. Data are from Hoggard et al. (2020). **(b)** Plot of plume-ridge distance and plate speed at the location of plumes. Plume-ridge distance come from GPlates (Müller et al., 2016; Whittaker et al., 2015), and plate speed data come from Becker et al. (2015)

7. Line 413- 415: "Based on a series of numerical modeling as well as geological and geophysical observations, we predict that mantle plumes in the Pacific Ocean are more likely to be dragged away by the spreading ridge." The authors emphasize on the fast spreading rate of Pacific ocean as a main factor for dragging plumes away from the ridges. I think the plume-ridge distance may be a main factor in this case; most of plume tails (shown as blue dots in Fig. 9a) in the Pacific Ocean are located away from the ridges.

Reply: *We appreciate this comment and agree with the reviewer. We reorganized the manuscript discussion (mostly in the final paragraph of section 4) accordingly.*

8. Line 145: Temperature of 2513 K is very high for temperature at 660 km. Considering adiabatic temperature gradient of 0.5 K km-1 and temperature of 1573 K at the base of lithosphere, temperature at the bottom of model should be ~1873 K.

Reply: *The model sizes in our study are set as 6600(width) and 1200(depth). The temperature of 2513K refers to the temperature at the bottom of model, that is, the temperature at 1200 km. Based on adiabatic temperature gradient of* 0.5 K km^{-1}, *the temperature at 660 km in all models is set initial to* 1573 K + (660-120)km*0.5K/km *=1843K.*

9. Fig. 5: I expect the heat flux and melt are initially maximum in the area above the plume head. Then due to underplating of plume and its flowing towards ridge, the location of maximum heat flux and melt changes in time. That would be worth to show the evolution of heat flux and melt in time (similar to what is shown for surface topography in Fig. 3 and 4). For (5e-f): I suggest to show the results of plate drag model from ridge to some distances away from it, similar to what is shown in (c) and (d). I suspect that in plate drag due to imposing higher extension rate, the whole lithosphere is experiencing cracks and becomes extremely weak. Figure 5 shows the results in the early stage of deformation. Can authors provide a figure showing results at later stages?

Reply: *We appreciate this suggestion and add figures to present the temporal evolution of the extracted melt (see figure S1, S4 in the supplementary material). The maximum extracted melt above the plume head, shown as crust thickness in Figure S1, S4, change with flowing of the plume, which is consistent with the topography evolution in Figure 3a and Figure 4a. Besides, considering that the heat flux is not that important to our model result (suggested by another referee), we removed the description of heat flux evolution but reserve the change of extracted melt over time.*

Secondly, we also present the lithosphere stress in a wider perspective (Figures. S1, S4). We agree with you. The whole lithosphere is under extension sightly because of the imposed extension rate. The presence of buoyant mantle plume leads to the stress localization within the above lithosphere. The normal stress in blue represents *extension, while the compression region is plotted in red. Actually, since plastic deformation does not occur in a widespread manner in our models, it indeed does not seem to be appropriate to describe such localized stress areas as "tension cracks". Please also see reply to comment 2. As a result, we removed the description of "cracks" in the main text (lines 212-213 and lines 249-250).*

Other comments

Line 36: What is Amsterdam?

Reply: *The Amsterdam here means the Amsterdam-Saint Paul mantle plume. We have made revision in the main text (line 41).*

*L*ines 136-138: It is not clear what this sentence mean. Please modify this sentence.

Reply: *We rephrased this sentence (lines 170-172) as "To reproduce the oceanic lithosphere, we choose a typical layered model, where the crust is composed of a water level (2 km), a sediment layer (1.5 km), a basalt layer (7.5 km)."*

Lines 147-148: This is not consistence with cooling half space. The temperature of the oceanic lithosphere tends to change linearly with depth when lithosphere is very old (older than ~80 Myr).

Reply: *Thanks for your comments. Indeed, the initial temperature distribution of the oceanic plate is prescribed by the half-space cooling model and thermal equilibrium structure. The half-space cooling model is used to describe the oceanic plate younger than 50Myr, and the thermal equilibrium structure is used to describe older oceanic parts. The thickness of the half-space cooling part is defined by the thermal structure. We set up the models in this way because we consider that the theoretical half-space cooling model has a good match with some geophysical observations when the plate is young, but the fit becomes poor when the age is greater than 60/70 Ma (Turcotte and Schubert, 2014; Stein and Stein, 1994). Therefore, we set a half-space cooling model with a maximum age at 50Ma, and the thermal equilibrium thickness of the older* *lithosphere is constant (i.e., ~100 km; corresponding to a thermal age of 50 Ma). We have made further description of the model initial setup in the main text.*

Line 153: “An additional velocity is imposed on both sides of the ridge to represent the half spreading rate. “Are they internal boundaries? Please explain more about it here; where are they and until which depth they extend.

Reply: *Yes. The velocity boundaries are internal boundaries, which are imposed on 500 km from each side of boundaries in the lithosphere (i.e., from 20 km to 120 km in depth). We made revision in the main text (lines 97-98; lines 187-189).*

Line 181: “The mantle flow vertical velocity profiles” It is a bit confusing. The profiles shown in Fig. 3f are the horizontal component of mantle velocities along two vertical profiles. Please rephrase this part and also explain the depths which were selected for these profiles. Are they from the surface till ~250 km depth?

Reply: *Thanks for your suggestions. The profiles shown in Fig. 3f and Fig. 4f are the horizontal velocities for two vertical profiles (i.e., from 80 km in the lithosphere to 240 km in the asthenosphere). The profiles are located 100 km away from the plume stem. We rephrase these sentences.*

Lines 353-354: “However, without plume further supplies, the overpressure difference from the plume head to the spreading center decreases slowly with time (Fig. 355 8d).” What does it mean?

Reply: *Thanks for your comments. We reworded this paragraph in section 3.4.3 (lines 411-417)*

Lines 186-187: “The overriding plate moves slower than the ponding plume, and hence actually slows down the spreading plume branches.” It is not clear what the message of this sentence is. According to model setup, since plume is located on the left side of MOR, the overriding plate motion speeds up the plume flow towards left (since the plume flow and plate motion have the same direction) and slows down the flow in the right plume branch.

Reply: *We disagree with the reviewer. The left plume branch spreads in the same direction but faster than the overriding plate. Hence it is slowed down by plate drag. The right branch is moving in the opposite direction, and hence is also slowed down by plate drag. We improved the description of these processes (lines 235-239).*

Lines 187-188: "Without suction effect from the spreading center, the left plume branch flows out much slower than the right branch." Similar to what I mentioned in my previous comment, I expect faster flow towards left.

***Reply*:** *We improved the sentences (lines 235-239). Please also see our reply to the previous comment. We also refer the reviewer to our added discussion of driving forces for plume spreading in section 3.3.*

Fig. 7: It is a very complicated figure. What do the upper panels of Fig. 7a-c stand for? They show the results at different times. Are the results shown in the lower panels of Fig. 7a-c showing the results at similar times as those shown in the upper panels? The scale of Fig. 7a-c is very small and one can hardly distinguish all the curves shown in the Figure. Please make the figure bigger. I suggest to move the legend of Fig. 7a into the right side of Figure because Fig. 7b-c also shows the results of models with different plume-ridge distances. The colors of curves for different plume-ridge distances are very similar and hard to differentiate them from each other. I suggest to change the colors. What do "plume head stage- positive spreading out" and "plume tail stage- passive flow driven by plate" mean? How are buoyancy fluxes calculated?

Reply: *We appreciate your suggestions and replot the Figure 7 (Fig.6 in the revised manuscript).*

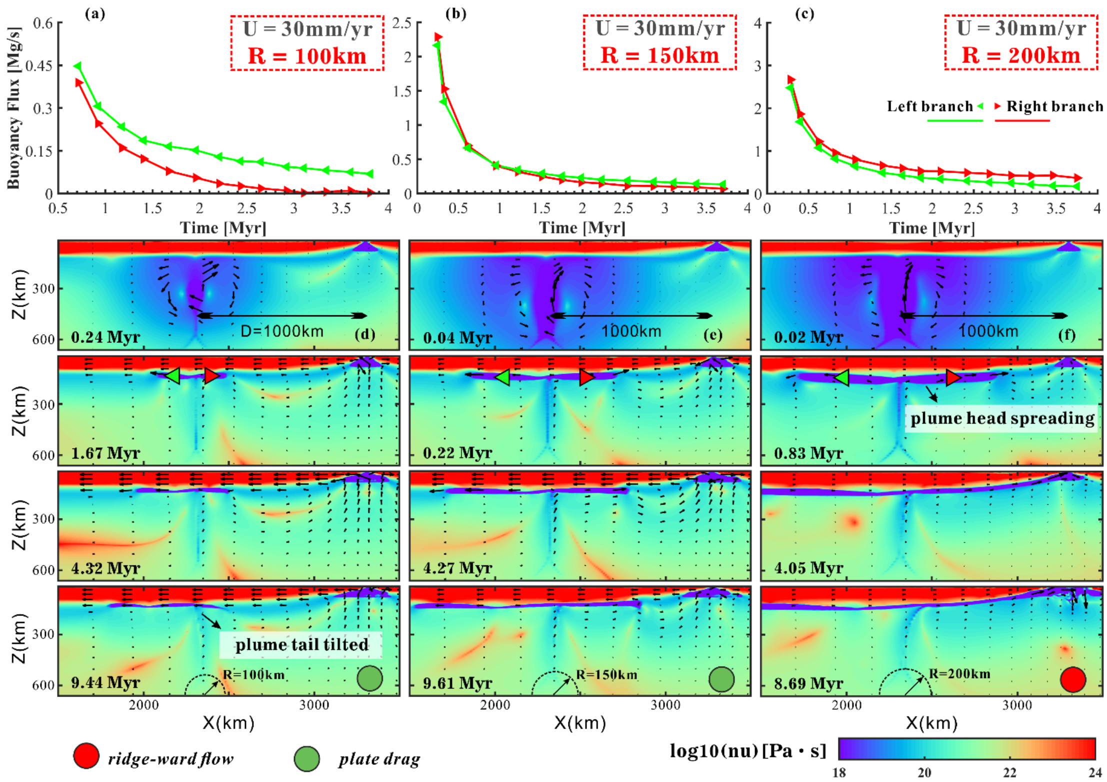

Figure 6 Models varying initial plume head radii (model M53, M58, and M63, Table S1 in supplementary material) shown by buoyancy flux and viscosity. **(a-c)** Buoyancy flux in spreading plume branches over time. Green and red triangles are markers used for buoyancy flux calculation. **(d-f)** Viscosity snapshots of models with different plume head radii. Models with green circle represent plate-drag flow and ridge-ward flow in red.

Line 330: How does Fig. 8a indicate that fast-spreading ridge promotes plume dragging. In this figure, from three models with fast spreading rates two are representing ridge suction mode.

Reply: *We appreciate this comment. We replot the Figure 8 (Fig.9 in the revised manuscript) to demonstrate the effects of spreading rates, choosing different typical models. In fast spreading ridge models, more plume material is dragged away, which make it difficult for plume to interact with the ridge. Therefore, we suggested that fast-*

spreading ridge promotes plume dragging.

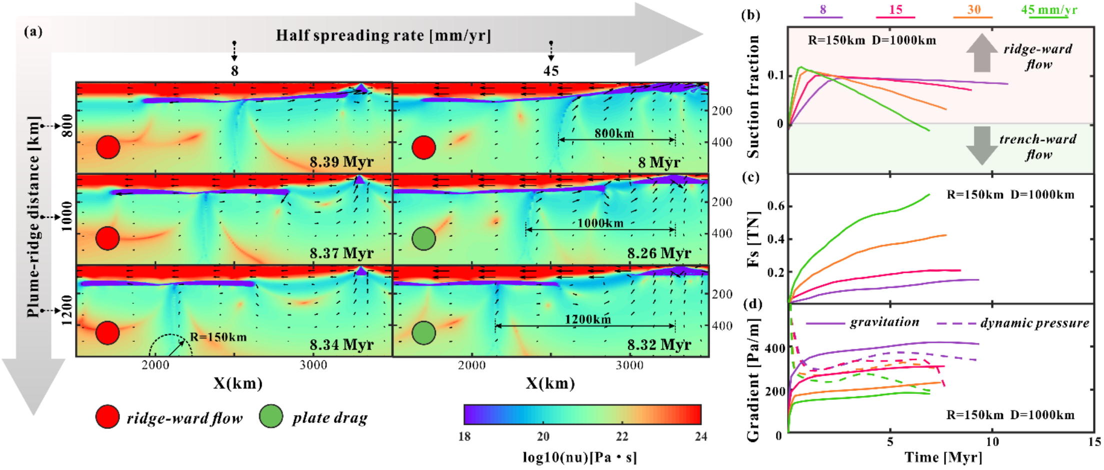

Figure 9. Model results influenced by different half spreading rates. **(a)** Effect of spreading rate on ridge-ward flow verse plate-drag flow. Viscosity snapshots are shown (model M7-M9, M82-M84, Table S1 in supplementary material). Fast-spreading ridge promotes plume material dragged. Models with green circle represent plate-drag flow and ridge-ward flow in red. **(b)** Dynamic evolutions of ridge-ward and plate-drag plume flow, revealed by defined ridge spreading fraction (eq.14). **(c)** Shear force (*Fs*) between moving plate and plume material under different spreading rates. **(d)** Pressure gradient between plume head and ridge center in different half spreading rate models. The solid and dash lines are the plume gravitation and dynamic pressure gradient, respectively.

Lines 340-347: It is not clear how shear force and pressure difference were calculated. Please re-write this part. Was the shear force calculated for the grids in the upper part of plume head or the whole plume head? The box of 50*50 km^2 in Fig. 8a is shown only for the plume head (and not for ridge center).

Reply: *Thanks for your advice. The shear force* F_s *is calculated by integrating the shear stress* σ_{xz} *of the uppermost plume head grids, not the whole plume head. Moreover, we replace the pressure difference with pressure gradient to clarify the* *mechanism of ridge suction. We calculate the plume gravitation and dynamic pressure gradient by tracing the plume markers which record the pressure, density, etc. The method how we compute the shear force and pressure gradient have revised and described in detail (lines 301-308; lines 401-407).*

Fig. 8: What is the distance of plume-ridge in models shown in Fig. 8b-d? What do the dashed color curves in Fig. 8b stand for? Please explain them in the caption. The scale of figures are small. What does "plume head spreading" in Figure 8d mean? What is the effect of plume size on shear force and overpressure difference?

Reply: *We appreciate your suggestions. We replotted the Figure 8(now Fig. 9 in the manuscript) and revised the caption. First, the "plume head spreading" in both figure and main text means the plume head spreads lateral under the plate. We explain it in the revised manuscript. Second, we added figures of shear force and pressure gradient of different plume size models in the supplementary material. The results shows that bigger size plumes are subjected to bigger shear force (Fig. S7). Meanwhile, bigger plume size means a more buoyant strong plume, which creates a larger pressure gradient between the plume and the ridge (Fig. S8). We have discussed this in the section 3.4.1.*

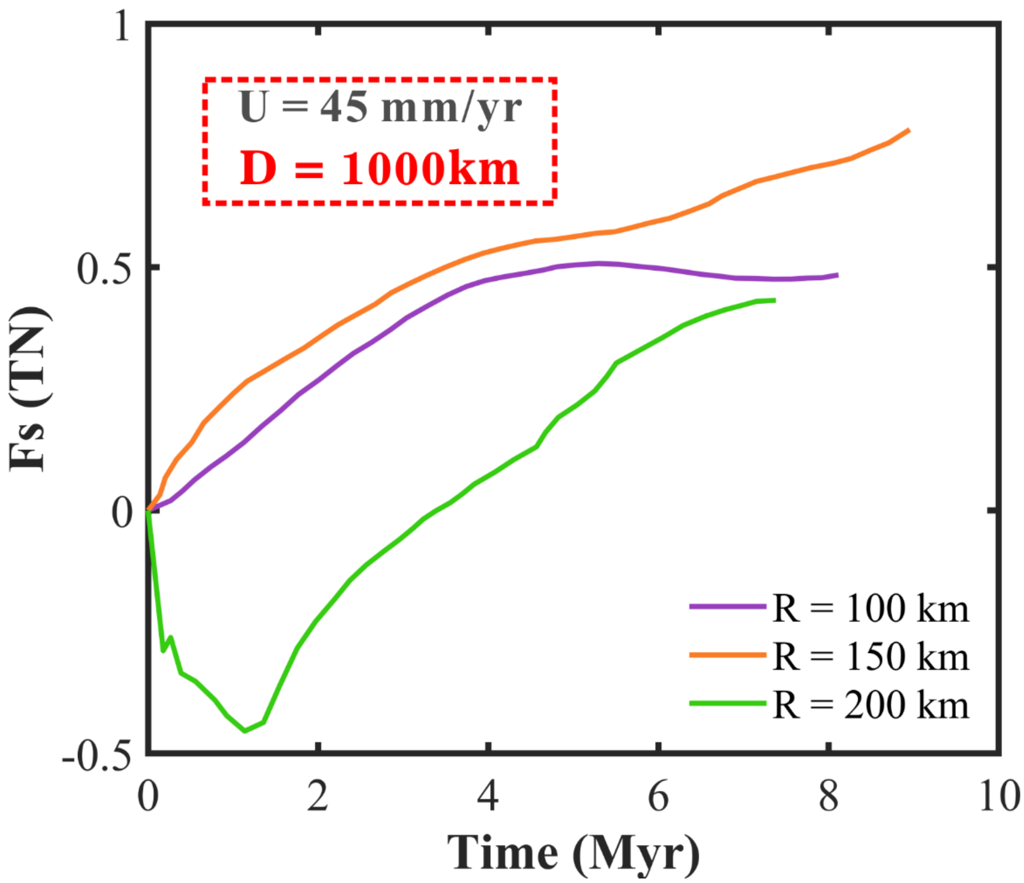

Figure S7. Shear force (*Fs*) between plate and plume in different plume head size models. The shear force imposed on the plume increases with plume size. The negative shear force indicates stronger friction imposed on the ridge-ward flowing (right) plume branch than that on the plate-drag flowing (left) plume branch.

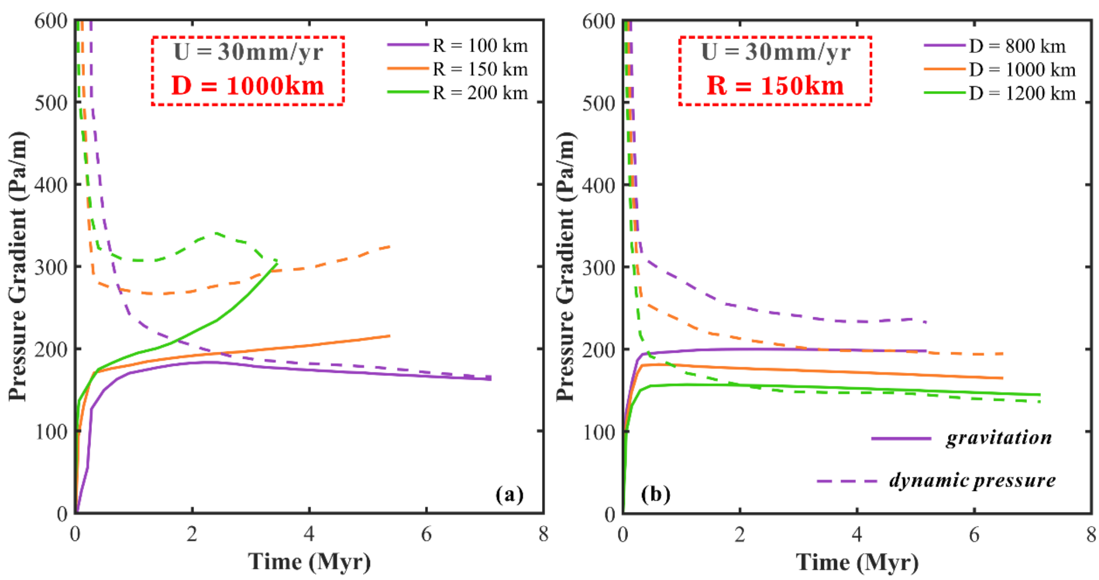

Figure S8. Pressure gradient between plume head and ridge center. The plume gravitation and dynamic pressure gradient of **(a)** different size plumes, **(b)** different plume-ridge distances are shown by solid and dash lines, respectively.

Lines 360-361: "while all models gradually switch from ridge suction in the plume head stage to dominant plate drag in the plume-tail stage" Is it valid for all models or only those representing plate drag regime?

Reply: *We appreciate this comment which triggers us a lot of thinking. The description may be incorrect here. Only those plate drag models shift from ridge suction in the plume head stage to dominant plate drag in the plume-tail stage. We removed this sentence to avoid semantic ambiguity.*

Figure 9: What do "MAR" and "EPR" stand for? Please explain them in the caption. How did the authors obtain the plume buoyancy flux (which indicate plume size) of hotspots shown in Fig. 9?

Reply: *The "MAR" and "EPR" indicate the Mid-Atlantic Ridge and the East Pacific Rise, respectively. The plume buoyancy flux data in Fig.9 come from Hoggard (2020) and are presented in different circle sizes. We modified the Figure 9 (Figure 10 in the revised manuscript) and revised its caption.*