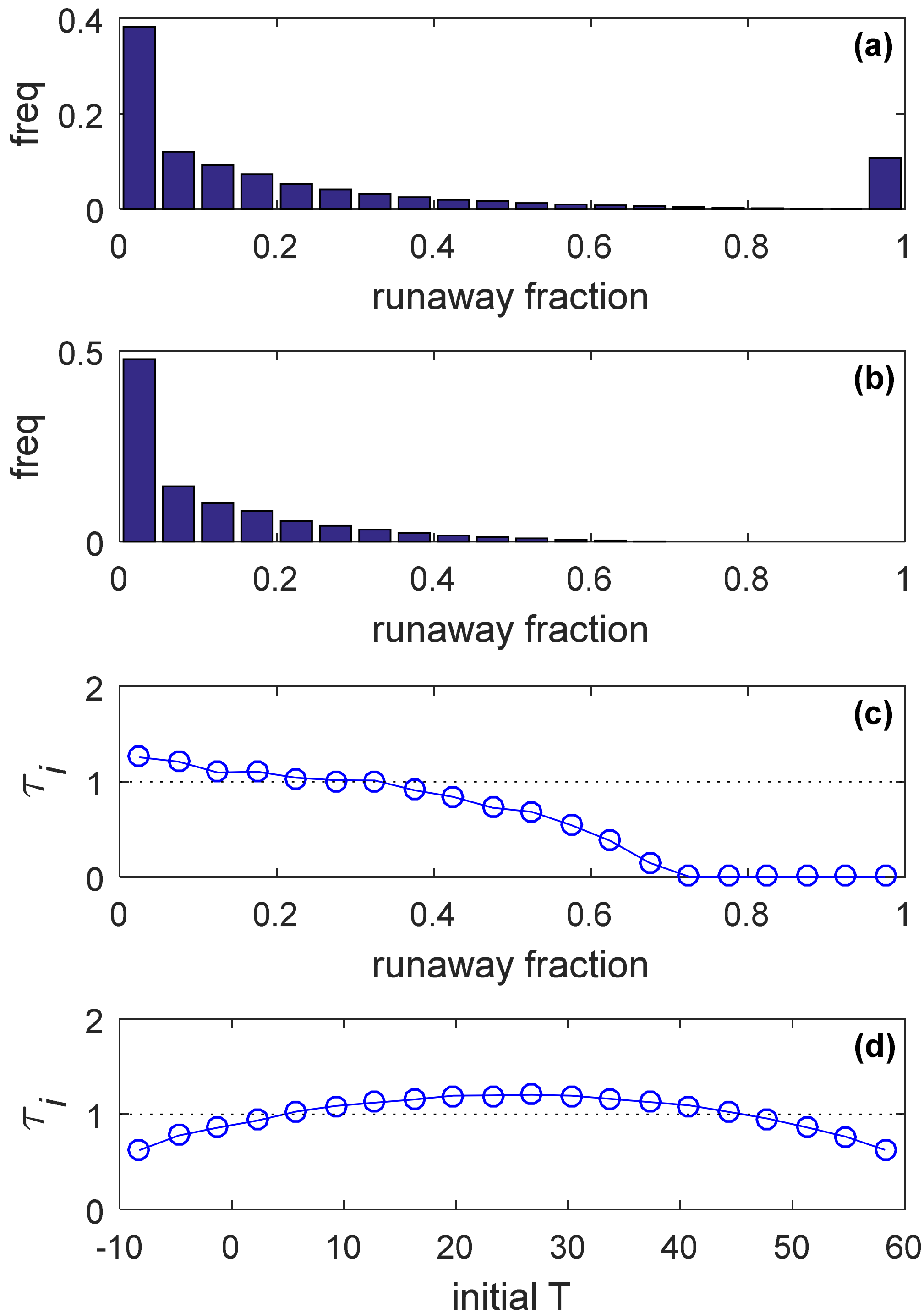
Supplementary Information

Accompanying the manuscript "Chance played a role in determining whether Earth stayed habitable" by Toby Tyrrell



Supplementary Figure 1: Frequencies of occurrence as a function of planet properties, from a simulation of 100,000 randomly different planets. Checks were made to verify that the model exhibited sensible behaviour in terms of how likelihood of success is influenced by various factors. The method of analysis is shown here, together with results for two factors. (a) histogram of the relative frequencies with which the automated planet-generation process produces planets with different values of fr_{runaway}, the fraction of the habitable range within which temperatures run away to uninhabitable values; (b) histogram of the relative frequencies of successful planets (those staying habitable at least once) with different values of fr_{runaway}; (c) variation of the τ_i statistic (measure of success) against the value of fr_{runaway} for the planet; (d) variation of the τ_i statistic against the starting temperature. In (c) and (d), the τ_i statistic for each bin is the proportion of successful planets from that bin divided by the proportion of all planets from that bin:

$$\tau_i = \frac{\left[N_i^s / \left(\sum_{i=1}^{i=nbins} N_i^s\right)\right]}{\left[N_i^{\mathrm{all}} / \left(\sum_{i=1}^{i=nbins} N_i^{\mathrm{all}}\right)\right]} \qquad (10)$$

where i is the bin number, $nbins$ is the total number of bins (20), N_i^{all} is the number of planets for which the value of the planet property is in the ith bin's range, and N_i^s is the number of successful planets for which the planet property is in the ith bin's range. As expected, (c) shows that increasing extent of runaway feedback zones leads to diminishing probability of planet success; (d) shows that the starting temperature has only a modest impact (runs starting very near to the edges of the habitable range are somewhat less likely on average to stay habitable, presumably because there is a greater probability near the edges of starting within a runaway feedback zone or of being immediately nudged outside the limits by a very early perturbation).

Supplementary Methods (Sensitivity Analyses)

Sensitivity analyses were conducted to see if the main conclusion of this paper is critically dependent on particular details of the model or whether it is more robust. The procedure for evaluating the effect of each model detail was to: (1) make a change to that detail in the model (for instance, change a parameter value); (2) execute a parallel but smaller simulation (1,000 planets each rerun 50 times rather than 100,000 planets each rerun 100 times for the standard simulation); then (3) examine the results to see if the major conclusion is again borne out.

How sensitivity was assessed

Results were examined in a few different ways to see if the main conclusion held in each altered version of the model.

First, success rates were examined to see if there was a spectrum of different probabilities of success (i.e. if some planets were successful between 1 and 49 times of out of 50, as opposed to all planets being either always successful – 50 out of 50 – or never successful – 0 out of 50).

Second, plots equivalent to Figure 3b were produced for each sensitivity analysis (not shown). In these plots, the distribution of planet success rates was again compared graphically to the distribution of success rates expected according to H1 and H2. If the blue line was again completely separate and distinct from the red and black lines, and positioned somewhere between the two of them, then this was taken as supporting the main conclusion (that both chance and mechanism played a role in whether planets stayed habitable).

Third, the number of sometimes habitable (at least 1 time out of 50) planets for each simulation was compared to the number consistent with destiny (H2) or chance (H1). This check consisted of a quantitative as opposed to a qualitative procedure. If planets are either always successful or always unsuccessful (i.e. H2), with no intermediate possibilities, then the fraction of planets that are habitable must be the same as the fraction of runs that stay habitable, or in other words:

$$N_{\mathrm{occ}}^{\mathrm{p}} = N_{\mathrm{total}}^{\mathrm{p}} * \left(N_{\mathrm{successful}}^{\mathrm{r}} / N_{\mathrm{total}}^{\mathrm{r}}\right) \tag{11}$$

On the other hand, if each run of each planet has exactly the same chance of staying habitable regardless of its feedbacks (i.e. H1), then the probability of any one run being successful (P^{r}) must be exactly the same as the fraction of all runs that are successful ($N_{\mathrm{successful}}^{\mathrm{r}} / N_{\mathrm{total}}^{\mathrm{r}}$). The probability ($P^{\mathrm{p}}$) of any one planet being successful at least once (1 or more times out of 50 reruns) would then be given by :

$$P^{\mathrm{p}} = 1.0 - (1.0 - P^{\mathrm{r}})^{50} \tag{12}$$

$$= 1.0 - \left[1.0 - \left(\frac{N_{\mathrm{successful}}^{\mathrm{r}}}{N_{\mathrm{total}}^{\mathrm{r}}}\right)\right]^{50} \tag{13}$$

The most likely number of successful planets out of 1,000 can then be calculated from the binomial probability distribution. When $((1{,}000 + 1) * P^{\mathrm{p}})$ is not an integer then $N_{\mathrm{occ}}^{\mathrm{p}}$ is predicted to be given by:

$$N_{\rm occc}^{\rm p} = \lfloor (1{,}000 + 1) * P^{\rm p} \rfloor \quad (14)$$

where ⌊ ⌋ is the floor function which rounds a decimal number to the nearest integer below.

The predicted values of $N_{\rm occ}^{\rm p}$ according to H2 (equation 12) and H1 (equation 15) were then compared to the observed value of $N_{\rm occ}^{\rm p}$ from the simulation for that sensitivity analysis.

As an additional check, the matlab command 'runstest' was used to check for random order of results[1]. When the run durations from all the runs in a sensitivity analysis are placed in a single vector with the first 50 being those from planet 1, the second 50 being those from planet 2, etc then random order would only be expected if H1 is correct. The 'runstest' command returns a value of 1 if the null hypothesis of random order is rejected (refutation of H1), or 0 if it cannot be rejected.

As noted in the main text, the frequency of long-term habitability varies greatly (e.g. orders of magnitude variation in $N^{\rm r}$ – see tables below) depending on which assumptions are used. The simulation results are not, therefore, a reliable guide to the frequency of long-term planetary habitability in the universe.

Results (sensitivity to parameter values)

The parameter changes investigated are shown on the left of Table S1 below. After every different parameter change it was found that there was always a spectrum of different probabilities of success. Likewise, in the equivalent of Figure 3b the blue line was always completely separate and distinct from the red and black lines, and was situated somewhere between the two of them. As shown in Table S1, the observed value of $N_{\rm occ}^{\rm p}$ was always much greater than (always more than twice as large as) the value according to H2 and always much less than (always less than half as large as) the values according to H1. Finally, the 'runstest' results were always 1 (p « 0.001), confirming rejection of H1 for all SAs.

Supplementary Table 1: Results of sensitivity analyses to examine the effect on the main conclusion of changing the values of model parameters.

Parameter [default value]	New value	Number of Successes			Number of Habitable Planets			1<2<3?
		Planets*		Runs†				
		$N^{\mathrm{p}}_{\mathrm{occ}}$	$N^{\mathrm{p}}_{\mathrm{always}}$	N^{r}	1.Destiny‡	2.Simulation	3.Chance§	
standard model	-	*86*	*0*	*923*	*18*	*86*	*606*	✓
$\boldsymbol{T_{\min}}$ [-10]	0	49	0	376	8	49	314	✓
$\boldsymbol{T_{\min}}$ [-10]	-50	201	0	2,905	58	201	950	✓
$\boldsymbol{T_{\max}}$ [60]	50	27	0	143	3	27	133	✓
$\boldsymbol{T_{\max}}$ [60]	100	199	2	2,847	57	199	947	✓
$\boldsymbol{N_{\mathrm{N}}^{\max}}$ [20]	40	48	0	384	8	48	320	✓
$\boldsymbol{N_{\mathrm{N}}^{\max}}$ [20]	10	123	0	1,151	23	123	688	✓
$\boldsymbol{\sigma_{\mathrm{f}}}$ [100]	50	41	0	309	6	41	266	✓
$\boldsymbol{\sigma_{\mathrm{f}}}$ [100]	200	124	0	1,163	23	124	692	✓
$\boldsymbol{\sigma_{\emptyset}}$ [50]	25	130	0	1,260	25	130	721	✓
$\boldsymbol{\sigma_{\emptyset}}$ [50]	100	41	0	396	8	41	328	✓
$\boldsymbol{N_{\mathrm{S}}}$ [10^4] $\boldsymbol{N_{\mathrm{M}}}$ [200] $\boldsymbol{N_{\mathrm{L}}}$ [4] ‡	10x more**	52	0	379	8	52	372	✓
$\boldsymbol{N_{\mathrm{S}}}$ [10^4] $\boldsymbol{N_{\mathrm{M}}}$ [200] $\boldsymbol{N_{\mathrm{L}}}$ [4] ‡	10x fewer	302	2	4,405	88	302	991	✓
$\boldsymbol{P_{\mathrm{S}}}$ [2,1] $\boldsymbol{P_{\mathrm{M}}}$ [5,2.5] $\boldsymbol{P_{\mathrm{L}}}$ [10,5]¶	all 2x larger**	22	0	47	1	22	46	✓
$\boldsymbol{P_{\mathrm{S}}}$ [2,1] $\boldsymbol{P_{\mathrm{M}}}$ [5,2.5] $\boldsymbol{P_{\mathrm{L}}}$ [10,5]¶	all 2x smaller	315	3	5,267	105	315	997	✓
$\boldsymbol{t_{\max}}$ [3By]	2By	121	0	1,122	22	121	679	✓
$\boldsymbol{t_{\max}}$ [3By]	4By	73	0	689	14	73	500	✓

* $N^{\mathrm{p}}_{\mathrm{occ}}$ is the number of planets out of 1,000 that were successful in at least 1 rerun, $N^{\mathrm{p}}_{\mathrm{always}}$ the number that were successful in all 50 reruns; † N^{r} is the number of runs out of 50,000 that were successful; ‡ (from equation 12); § (from equation 15); ¶ for this sensitivity analysis, the numbers or magnitudes of the three classes of perturbations were all increased or decreased together); ** from 10,000 planets x 50 reruns (10 times more than normal).

Structural sensitivity analyses

A second set of sensitivity analyses examined the effect on the simulations results of changing various structural assumptions. The rationale for each structural change is described below, followed by a description (in italics) of how the change was implemented.

1. The stock of asteroids and comets likely to intersect a planet's orbit may become depleted over time (less frequent impact events), planets may cool and stabilise internally over time (less frequent supervolcanoe eruptions), and the population of massive nearby stars (those prone to gravitational collapse to produce supernovae) is likely to become depleted over time. *{Perturbations were made to decline in frequency over time instead of being equally frequent at all times. Each perturbation time* t_{p} *was calculated with:* $t_{\mathrm{p}} = t' * (1+k) * ((R+1)^{-8} - k)$ *where* $k = 2^{-8}$ *and* $R \sim U(0,1)$ *}*

2. Complex life is, on the whole, less tolerant of extreme environmental conditions (prokaryotes can tolerate a wider range of conditions than can eukaryotes, for instance[2]). {*The habitable envelope was made to become narrower over time, instead of remaining of constant width. The habitable range was made to start off 20ºC wider than in the default simulation (100ºC rather than 80ºC) and become progressively narrower, ending up 60ºC wide after 3 By}*

3. The feedbacks may possibly vary more smoothly as a function of temperature. {*The value of the feedback f between nodes was calculated using a spline (cubic) rather than linear interpolation*}

4. Alternatively, feedbacks could be quite constant across large temperature intervals but then change quite rapidly across narrow temperature intervals. {*Nodes were given random positions within the habitable range, instead of being positioned at regular intervals across it*}

5. The default model contains no bias either for or against stabilisation but planets as whole may in reality be biased towards stability because of the Stefan-Boltzmann Law (planets with hotter surface temperatures radiate heat more rapidly to space, unless subject to different greenhouse effects)[3]. {*A bias towards stabilisation was implemented by addition of an amount* ϑ_i *to all* f_i*, where* $\vartheta_i = (\sigma_f/2) - ([\sigma_f * (T_i - T_{\min})/(T_{\max} - T_{\min})])$ }

6. Warming or cooling forcings are equally likely in the default model but warming is arguably more likely because stars on the main sequence always become more luminous over time[4]. {*A positive bias was imparted by adding a fixed positive amount (equal to* $\sigma_\emptyset$*) to the randomly determined forcing in order to make positive values more likely than negative ones, changing equation (7) to:* $\emptyset = \{\sim N(0, \sigma_\emptyset) + \sigma_\emptyset\}$}

7. It is unclear whether feedbacks values should be uniformly or normally distributed. {*Instead of being chosen from a normal distribution centred at zero, values of f are chosen instead from a uniform distribution* [$f \sim U(-\alpha, +\alpha)$]*, where α is chosen so as not to alter the average absolute magnitude of f*}

8. Evolution is a process whose rate appears to vary, progressing more slowly when the environment is more static (for instance the 'boring billion' on Earth)[5] and more quickly during and following periods of environmental change. The time that a planet needs to remain habitable in order for intelligence to evolve may therefore vary greatly between planets rather than always being more or less the same[6]. {*Instead of being fixed at 3 By,* t' *for each planet is set to a random number between 0.5 and 10 By*}

9. A planet could lose habitability early on, only to recover and remain habitable for 3 By later. *{The range a planet's temperature can vary within is extended to (-200ºC to +500ºC), while the habitable range remains the same. The maximum number of nodes (*$N_{\mathrm{N}}^{\max}$*) is increased proportionately. Each run is simulated for 10 By, regardless of any exits / re-entries to the habitable range during that time. At the*

end of each run of 10 By, the longest continuous duration of habitability, whenever in the run it occurred, is compared to t'}

10. Because of its simplicity, the model cannot reproduce absolutely all of the behaviours of more complex models. This extreme simplicity can lead to a shortcoming through the lack of intervening climate system variables, such as albedo and greenhouse gas concentrations. Because the value of *f* for each value of *T* is fixed at the outset for each planet, the standard version of the simulation is not able to simulate feedbacks that change slowly at a constant planetary temperature. This is not a significant problem for most such delayed responses that take time to build up (e.g. the waxing or waning of ice sheets following a change in Earth surface temperature), because their timescales of readjustment to the prevailing temperature are still very short (e.g. millennia) compared to the lifespan of a planet. It is, however, more problematic for slower readjustments to temperature. For instance, the standard model cannot capture the rather slow (over 4 to 30 My [7]) accumulation of volcanic CO_2 in the atmosphere postulated to have brought an end to Proterozoic Snowball Earth events. {*To enable the possibility of accumulation or depletion of climate-affecting substances at a constant temperature, 3 additional climate-relevant variables (X_1, X_2 and X_3) were added and their behaviour controlled by three additional ODEs of the form $dX_i/dt = f_i(X_i) + \emptyset_i . t + c_i . T$, where f_i are randomly assigned feedbacks as for temperature in the standard model and $\emptyset_i$ are randomly assigned temporal forcings. The ODE for temperature was modified to incorporate the effects of these other factors: $dT/dt = f(T) + \emptyset . t + k_1 . X_1 + k_2 . X_2 + k_3 . X_3$; c_1, c_2 and c_3 are random coefficients $\sim U(-1, +1)$ and k_1, k_2 and k_3 are random coefficients $\sim U(-2, +2)$ }*

11. In the standard model, the feedbacks are fixed at the beginning and not subsequently altered. It is possible, however, for significant changes to planetary climate systems to occur later on in their histories[8], such as occurred on Earth ca. 380 Mya when trees evolved and the first forests spread out across the land[2]. {*A random number $\sim U_i(0,8)$ of changes were made to f at random times. Each change was of one of four types: (i) addition of the same amount ($\sim N(0, \sigma_f)$) to all f_i; (ii) a ramped change with a minimum addition at one edge of the habitable range and a maximum at the other, changing the overall slope of $\delta T/\delta t$ vs T; (iii) addition according to a peaked (triangular) profile; and (iv) addition according to a Gaussian profile with centre randomly placed on the T axis*}

12. As well as being subjected to random instantaneous perturbations, the Earth also experiences periodic fluctuations in the amount and seasonality of solar heat it receives (Milankovitch cycles), and most likely is affected also by other longer period cyclic forcings, such as supercontinent assembly/break-up due to plate tectonics. {*A random number $\sim U_i(1,4)$ of both short period [$\tau \sim U(0, 0.001\mathrm{By})$] and long period [$\tau \sim U(0, 0.5\mathrm{By})$] extra terms were added to the ODE for planetary temperature. For each one, the term added to the ODE was $\left(k_i . \sin\left(2\pi t/\tau\right)\right)$, where k_i is the randomly-determined magnitude $\sim U(0, (\sigma_f/5))$ of each cyclical forcing and t is time}*

13. The increase over time in the luminosity of the Sun happens to be almost linear[9] but temporal forcings in general need not be. {*The shape of the temporal forcing was made to be non-linear by multiplying the overall magnitude of the forcing (∅) by $(t/t_{\max})^{\gamma}$ where t is time, $t_{\max}$ is 3 By and γ is (50% probability) either $\sim U(1.0, 6.0)$ (exponential shape) or $1/{\sim}U(1.0, 6.0)$ (logarithmic shape)*}

14. Early life on Earth was probably thermophilic, perhaps because it first evolved under hot conditions. Perhaps a hot environment is required for the origin of life[10]. {*This possibility was implemented by choosing the initial temperature from a narrower and hotter range: $T_0 \sim U((T_{\max} - 20), T_{\max})$}*

15. The upper limit to thermal habitability is arguably either a function of the boiling point of water or else of the temperature at which enzymes including DNA denature. The former is assumed in this model variant, with the boiling point of water made a function of atmospheric pressure rather than constant as

in the standard model. {*An additional environmental variable, atmospheric pressure at sea-level, was added through the addition of an ODE of the form* $dX_P/dt = f_P(X_P) + \emptyset_P.t$ *, where* f_P *are randomly assigned feedbacks and* $\emptyset_P$ *is a randomly assigned temporal forcing.* $T_{\max}$ *(°C) was made to be a function of pressure (*P*, units of Pa):* $T_{\max} = [4908/(24.67 - \ln P)] - 273.15$ }

16. Although temperature is the primary factor determining habitability[11], other environmental variables (such as pH, ionic strength for life in water) also need to stay within habitable limits for intelligence to be able to evolve. {*Three additional environmental variables (*X_1*,* X_2 *and* X_3*) were added through the addition of three ODEs of the form* $dX_i/dt = f_i(X_i) + \emptyset_i.t$ *, where* f_i *are randomly assigned feedbacks as for temperature in the standard model and* $\emptyset_i$ *are randomly assigned temporal forcings. The planet was deemed to have gone uninhabitable when either temperature (simulated in the normal way) or any other variable left its habitable range (-100 to +100 for the additional variables, in arbitrary units).*}

17. Perhaps life can survive periods of inhospitable conditions within refuges (for instance, the survival of photosynthetic algae through Snowball Earth episodes lasting for many millions of years[12,13] may have been within persistent polynyas[14]. If so, then average global conditions may need to become much more extreme in order for conditions everywhere to be intolerable for life. {$T_{\min}$ *and* $T_{\max}$ *were both changed, to -70°C and +150°C*}

18. There is room for debate about exactly what should be encompassed by the term destiny and what should be considered as part of luck. In this sensitivity analysis an alternative categorisation is used, whereby chance is deemed to consist only of perturbations (random events driving sudden changes in temperature) with all else (including initial temperature) included as part of destiny. *{Rather than being calculated separately for each rerun of each planet,* T_0 *was calculated only once for each planet, with each rerun of a planet being started from the same initial temperature}*

19. As for the previous sensitivity analysis, the definitions of luck and destiny are altered. This time a planet's feedbacks (*f*) are considered to be the only factors that might contribute to its destiny, with all else being considered part of chance and therefore made to vary between reruns of the same planet. Repeat runs of the same planet can therefore be thought of as taking place in different locations (in different galaxies or in different regions of the same galaxy, around different stars or in different planetary systems). *{Rather than being calculated only once for each planet,* $\emptyset$*,* λ_L*,* λ_M *and* λ_S *were all recalculated for each and every rerun}*

20. It is necessary to consider whether the results obtained are biased by the degree of hazard built into the model. It should be considered whether the results would come to resemble H1 (outcomes controlled only by chance) if the perturbations and forcings were made larger and the feedbacks smaller. *{Although this was partially addressed in other SAs with changes to individual parameter values, in this SA it was purposefully addressed through a combination of changes: twice as many perturbations, 50% larger perturbations and forcings, and 25% smaller feedbacks*}

21. Conversely, it should also be considered whether the results would come to resemble H2 (outcomes determined only by the nature of the planet's feedbacks) if the perturbations and forcings were made smaller and the feedbacks larger. *{Again, this question was partially tested in other SAs but was purposefully assessed in this SA through: omitting the larger class of perturbations altogether[63], half the numbers of other perturbations, 25% smaller magnitudes of perturbations and forcings, and feedbacks 50% larger than in the standard version*}

Results (sensitivity to structural assumptions)

After each structural change there was always a continuum of success probabilities. In the equivalent of Figure 2b for each sensitivity analysis the blue line was always separate and distinct from the red and black lines and somewhere between them. As shown in Table S2, the observed value of $N_{\mathrm{occ}}^{\mathrm{p}}$ was always much greater than (nearly always more than twice as large as) the value according to H2, and always much less than (always less than half as large as) the values according to H1. Finally, ‘runstest’ results were again always 1 (p « 0.001), confirming rejection of H1 for all SAs.

Supplementary Table 2: Results of sensitivity analyses to examine the effect on the main conclusion of structural changes to the model.

Sensitivity analysis number	Number of Successes			Number of Habitable Planets			1<2<3?
	Planets*		Runs†				
	$N^{\mathrm{p}}_{\mathrm{occ}}$	$N^{\mathrm{p}}_{\mathrm{always}}$	N^{r}	1.Destiny‡	2.Simulation	3.Chance§	
standard model	*82*	*0*	*619*	*12*	*82*	*464*	✓
1	108	1	1,234	25	108	714	✓
2	81	0	934	19	81	611	✓
3	75	0	749	15	75	530	✓
4	75	0	691	14	75	501	✓
5	171	0	1,850	37	171	849	✓
6	60	0	699	14	60	505	✓
7	83	0	770	15	83	540	✓
8	104	0	604	12	104	455	✓
9	94	1	1,215	24	94	708	✓
10	79	0	267	5	79	235	✓
11	163	0	510	10	163	401	✓
12	40	0	394	8	40	327	✓
13	80	0	690	14	80	690	✓
14	71	0	689	14	71	500	✓
15	282	4	7,424	148	282	1,000	✓
16	37	0	319	6	37	274	✓
17	411	8	10,264	205	411	1,000	✓
18	84	0	967	19	84	623	✓
19	215	0	745	15	215	528	✓
20**	37	0	255	5	37	251	✓
21	379	26	8,971	179	379	1,000	✓

*$N^{\mathrm{p}}_{\mathrm{occ}}$ is the number of planets out of 1,000 that were successful in at least 1 rerun, $N^{\mathrm{p}}_{\mathrm{always}}$ the number that were successful in all 50 reruns; † N^{r} is the number of runs out of 50,000 that were successful; ‡ (from equation 12); § (from equation 15); ** from 10,000 planets x 50 reruns (10 times more than normal).

Supplementary Table 3: Some heating and cooling fluxes on Earth, and associated potential rates of temperature change. The upper half of the table shows the rates from individual components of the radiation balance if they were to act in isolation. Net fluxes are likely to be smaller because, in reality, there are always two or more processes acting, some of which oppose each other (stellar heating and outgoing infrared radiation for instance). The lower half of the table shows net fluxes associated with different amounts of energy imbalance.

	Magnitude of energy flux		Associated (dT/dt) ‡
Heating or cooling process	($W\ m^{-2}$) *	($10^{24}\ J\ y^{-1} = YJ\ y^{-1}$) †	($°C\ ky^{-1}$)
Solar radiation (sunlight) at the top of the atmosphere	340	5.5	1000
Reflected sunlight (due to albedo)	100	1.6	300
Emitted infrared radiation (heat loss to space)	240	3.9	700
Total greenhouse effect (water vapour, CO_2, clouds, O_3, CH_4, etc)[15]	160	2.6	500
Heat flow from Earth's interior[16]	0.09	0.001	0.3
10% energy imbalance at Earth's surface	53	0.85	150
Greenhouse warming due to CO_2 doubling[17]	3.4	0.055	10
Anthropogenic net climate forcing (since 2005)[15]	0.6	0.037	2

* averaged across the whole of the Earth's surface.
† summed over Earth's surface area of $510 \times 10^{12}\ m^2$; $1\ W = 1\ J\ s^{-1}$; yotta (Y) = 10^{24}.
‡ calculated assuming that Earth's long-term heat capacity is the same as the heat capacity of the whole ocean, $\sim 5.6 \times 10^{24}\ J\ K^{-1}$ (ref. 18).

Supplementary Table 4: Estimated temperature changes associated with some catastrophic events (upper half) and times of sudden climate change (lower half) on Earth.

Perturbation	Effect on global surface temperature	Literature Source
Mount Pinatubo volcanic eruption	-0.5 °C	19
Tambora volcanic eruption	-1.1 °C	19
Toba volcanic supereruption	-2.0 °C	19
Cretaceous/Paleogene boundary (asteroid impact)	up to -10 °C	20
Deccan Traps large igneous province (outpouring of basalts)	+2 °C	21
Paleocene-Eocene boundary (cause uncertain)	+5 °C	22
Eocene-Oligocene Boundary (cause uncertain)	-3 to -6 °C	23,24
Permian-Triassic boundary, largest mass extinction (cause uncertain)	+8 °C	25,26

Supplementary References

1. Bradley, J.V. *Distribution-Free Statistical Tests*. Prentice-Hall (1968).
2. Tyrrell, T. *On Gaia: A Critical Investigation of the Relationship Between Life and Earth*. Princeton University Press, Princeton (2013).
3. Kasting, J.F. & Catling, D. Evolution of a habitable planet. *Annual Review of Astronomy and Astrophysics* **41**, 429-463 (2003).
4. Hansen, C.J., Kawaler, S.D., Trimble, V. *Stellar interiors: physical principles, structure, and evolution*. 2^{nd} edition. Springer, New York (2012).
5. Holland, H.D. The oxygenation of the atmosphere and oceans. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 903-915 (2006).
6. McKay, C.P. Time for intelligence on other planets. In: *Circumstellar Habitable Zones*, (Ed. Doyle, L.), Travis House, Menlo Park, CA, pp. 405-419 (1996).
7. Pierrehumbert, R.T. High levels of atmospheric carbon dioxide necessary for the termination of global glaciation. *Nature* **429**, 646-649 (2004).
8. Doolittle, W.F. Natural selection through survival alone, and the possibility of Gaia. *Biology & Philosophy* **29**, 415-423 (2014).
9. Feulner, G. The faint young Sun problem. *Reviews of Geophysics* **50**, RG2006, *doi:*10.1029/2011RG000375 (2012).
10. Lineweaver, C.H. & Chopra, A. The habitability of our Earth and other Earths: astrophysical, geochemical, geophysical, and biological limits on planet habitability. *Annual Review of Earth and Planetary Sciences* **40**, 597-623 (2012).
11. McKay, C.P. Requirements and limits for life in the context of exoplanets. *Proceedings of the National Academy of Sciences* **111**, 12628-12633 (2014).
12. Rooney, A.D., Strauss, J.V., Brandon, A.D., Macdonald, F.A. A Cryogenian chronology: Two long-lasting synchronous Neoproterozoic glaciations. *Geology* **43**, 459-462 (2015).
13. Prave, A.R., Condon, D.J., Hoffmann, K.H., Tapster, S., Fallick, A.E. Duration and nature of the end-Cryogenian (Marinoan) glaciation. *Geology* G38089-1 (2016).
14. Olcott, A.N., Sessions, A.L., Corsetti, F.A., Kaufman, A.J. & De Oliviera, T.F. Biomarker evidence for photosynthesis during Neoproterozoic glaciation. *Science* **310**, 471-474 (2005).
15. Stephens, G.L. et al. An update on Earth's energy balance in light of the latest global observations. *Nature Geoscience* **5**, 691-696 (2012).
16. Davies, J.H. & Davies, D.R. Earth's surface heat flux. *Solid Earth* **1***,* 5-24 (2010).
17. Martínez-Botí, M.A*. et al.* Plio-Pleistocene climate sensitivity evaluated using high-resolution CO_2 records. *Nature* **518**, 49-54 (2015).
18. Sloan, D., Batista, R.A., Loeb, A. The resilience of life to astrophysical events. *Scientific Reports* **7**, doi: 10.1038/s41598-017-05796-x (2017).
19. Sigl, M. *et al.* Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* 523, 543-549 (2015).

20. Schulte, P. *et al.* The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. *Science* **327**, 1214-1218 (2010).

21. Wilf, P., Johnson, K.R., Huber, B.T. Correlated terrestrial and marine evidence for global climate changes before mass extinction at the Cretaceous–Paleogene boundary. *Proceedings of the National Academy of Sciences* **100**, 599-604 (2003).

22. Jones, T.D. *et al.* Climate model and proxy data constraints on ocean warming across the Paleocene–Eocene Thermal Maximum. *Earth-Science Reviews* **125**, 123-145(2013).

23. Liu, Z. *et al.* Global cooling during the Eocene-Oligocene climate transition. *Science* **323**, 1187-1190 (2009).

24. Wade, B.S. *et al.* Multiproxy record of abrupt sea-surface cooling across the Eocene-Oligocene transition in the Gulf of Mexico. *Geology* **40**, 159-162 (2012).

25. Joachimski, M. *et al.* Climate warming in the latest Permian and the Permian–Triassic mass extinction. *Geology* **40**, 195-198 (2012).

26. Cui, Y. & Kump, L.R. Global warming and the end-Permian extinction event: proxy and modeling perspectives. *Earth-Science Reviews* **149**, 5-22 (2015).