How common **are habitable planets?**

Jack J. Lissauer

The Earth is teeming with life, which occupies a diverse array of environments; other bodies in our Solar System offer fewer, if any, niches that are habitable by life as we know it. Nonetheless, astronomical studies suggest that many habitable planets may be present within our Galaxy.

One of the most basic questions that has been pondered by Natural Philosophers for many millennia concerns humanity's place in the Universe: are we alone? This question has been approached from many different viewpoints, and similar reasoning has led to widely diverse answers. Copernicus, Kepler, Galileo and Newton each demonstrated convincingly that other planets that were qualitatively similar to Earth orbit the Sun. In the past few years, more than 20 planets have been discovered in orbit about stars other than our Sun; these are 'extrasolar' planets.

The intellectual and technological advances of the past century leave us poised at the turn of the millennium to investigate the possibility of extraterrestrial life along numerous paths of experimental, observational and theoretical studies in both the physical and life sciences.

> Since one of the most wondrous and noble questions in Nature is whether there is one world or many, a question that the human mind desires to understand, it seems desirable for us to inquire about it.
>
> **Albertus Magnus, 13th century**

Prerequisites for habitability

Here I assume that extraterrestrial life would be carbon-based and use liquid water (characteristics common to all life found on Earth), and I define a 'habitable planet' as one capable of supporting such life.

Life on Earth has been able to evolve and thrive thanks to billions of years of benign climate. Mars seems to have had a climate sufficiently mild for liquid water to have flowed on its surface when the Solar System was roughly one-tenth its current age, but at present, its low atmospheric pressure means that liquid water is not stable on the martian surface. Venus is too hot, with a massive atmosphere dominated by carbon dioxide; we cannot say whether or not young Venus had a mild Earth-like climate. Indeed, as models of stellar evolution predict that the young Sun was about 25 per cent less luminous than at present, we do not understand why Earth, much less Mars, was warm enough to be covered by liquid oceans 4 billion years ago, when life is thought to have originated.

Carbon dioxide is important for carbon-based life. On Earth, this compound cycles — on a wide range of timescales — between the atmosphere, the oceans, living organisms, fossil fuels and carbonate rocks. The carbonate rocks form the largest reservoir, and are produced by reactions involving water (and in some cases living organisms). Carbon dioxide is recycled from carbonates back into the atmosphere as tectonic plates descend into the Earth's mantle and are heated. Carbonates are not readily recycled on a geologically inactive planet such as Mars, and they are not formed on planets like Venus, which lack surface water. Larger planets of a given composition remain geologically active for longer, as they have smaller ratios of surface area to mass, and thus retain heat from accretion and radioactive decay for longer. The number of variables involved in determining a planet's habitability precludes a complete discussion, but some of the main issues are summarized in Fig. 1 (compare ref. 1).

Stellar properties and habitability

Stars are huge balls of plasma that radiate energy from their surfaces and liberate energy through thermonuclear fusion reactions in their interiors. During a star's long-lived 'main-sequence phase', hydrogen in its core is gradually 'burned up' to maintain sufficient pressure to balance gravity. The star's luminosity (energy output) grows slowly during this phase, as fusion increases the mean particle mass in the core and greater temperature is required to achieve pressure balance. Once the hydrogen in the core is used up, the star's structure and luminosity change much more rapidly. Both Sun-like stars and larger stars expand and become 'red giants'; those stars with an initial mass greater than about eight solar masses can end their lives in spectacular supernova explosions.

What stars are most likely to have habitable planets? To make the astronomical problem more straightforward, I assume that the single factor required for supporting life on a planet is the presence of water on its surface over a long timescale. The main-sequence phase of low-mass stars (such as our Sun) provides 'continuously habitable zones'; when in orbits within these zones, planets may maintain liquid water on their surfaces for billions of years. High-mass stars are much hotter than low-mass stars and use up their fuel far more rapidly. Thus, even if Earth-like planets form around high-mass stars at distances where liquid water is stable, it is unlikely that benign conditions exist for long enough on these planets to enable life to form and evolve. However, the greater flux of ultraviolet radiation may speed up biological evolution enough to compensate for the shorter lifetime of a moderately massive star. At the other end of the size spectrum, the smallest, faintest stars can live for trillions of years, but they emit almost all of their luminosity at infrared wavelengths and their luminosity varies by tens of per cent owing to flares and large 'starspots' (analogous to sunspots). In addition, habitable-zone planets orbit so close to these faint stars that their rotation is tidally synchronized (as the Moon's rotation is relative to Earth); thus no day–night cycle occurs, and if the planet's atmosphere is thin it would freeze on the perpetually dark, cold hemisphere[2].

NASA/CORBIS/OAR/NURP

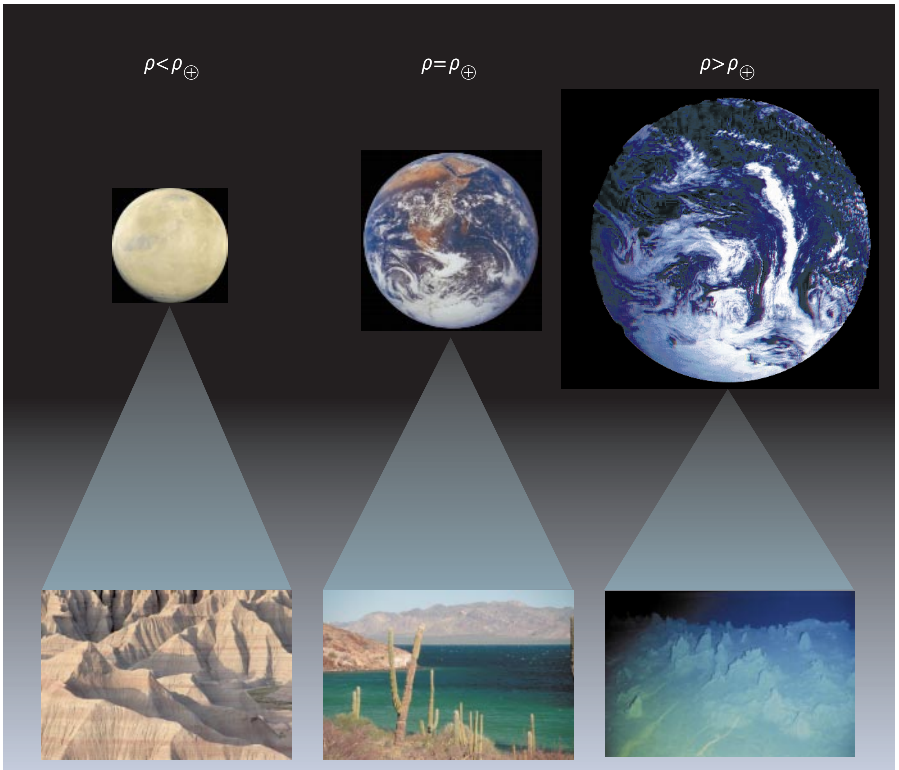


Figure 1 **Environments on small and large Earth-like planets. Earth is depicted in the central panel of the figure. A smaller planet (left) made of the same material as Earth would be less dense, because the pressure in the interior would be lower. Such a planet would have a larger ratio of surface area to mass, so its interior would cool faster. Its lower surface gravity and more rigid crust would allow for higher mountains and deeper valleys than are seen on Earth. Most important to life is that it would have a much smaller surface pressure as a result of four factors: a larger ratio of surface area to mass, lower surface gravity, more volatiles sequestered in crust as there is less crustal recycling, and more atmospheric volatiles escaping to space. Among other things, this would imply a lower surface temperature, because there would be less greenhouse gases in the atmosphere. Some remedial measures that could improve the habitability of such a mass-deprived planet are: (1) move it closer to the star, so less greenhouse effect would be needed to keep the surface temperatures comfortable; (2) add extra atmospheric volatiles; and (3) include a larger fraction of long-lived radioactive nuclei than on Earth, to maintain crustal recycling. A larger planet (right) made of the same material as Earth would be denser and have a hotter interior. Its higher surface gravity and less rigid crust would lead to muted topography. It would have a much greater atmospheric pressure, and, unless its greenhouse effect was strong enough to boil away its water, much deeper oceans, probably covering its entire surface. Some remedial measures that could improve the habitability of such a mass-gifted planet are: (1) move it farther from the star; and (2) include a smaller fraction of atmospheric volatiles. It is not clear that more active crustal recycling would be a problem, within limits.**

Stability of planetary systems

Although isolated single-planet systems are stable essentially indefinitely, mutual gravitational perturbations within multiple-planet systems can lead to collisions and ejections from the system. To a first approximation the star's gravity dominates, but planets exchange orbital energy and angular momentum, so that over millions or billions of orbits, even weak perturbations could move planets out of orbits that were initially in the habitable zone. Resonances among various orbital and precession (that is, rotation of the spatial orientation of the orbit ellipse) frequencies are the main source of chaos in planetary systems. There are no simple criteria for determining the stability of systems with many planets, but in general, larger spacing between orbits, smaller eccentricities and inclinations, and lower-mass planets are more stable[3].

One of these criteria — large spacing between orbits — has important implications. There is a minimum separation required for a system of Earth-mass planets to be stable for long periods of time. This separation is comparable to the width of a star's continuously habitable zone. Thus, arguments based on orbital stability support the possibility that most stars could have one or even two planets with liquid water on their surfaces. But unless greenhouse effects conspire to substantially compensate for increasing distance from the star, larger numbers of habitable planets per system are unlikely.

Planet formation

Stars are observed to be forming within cold regions of our Galaxy called molecular clouds. Even a very slowly rotating molecular-cloud core of stellar mass has far too much angular momentum to collapse down to an object of stellar dimensions. This leads to a (proto)star surrounded by a rotating disk of material; a significant fraction of the material in a collapsing core falls onto this disk. Such a disk has the same initial elemental composition as the growing star. At sufficient distances from the central star, it is cool enough for about 1–2 per cent of this material to be in solid form, either remnant interstellar grains or condensates formed within the disk. The growth from micrometre-sized dust to kilometre-sized solid bodies called planetesimals remains poorly understood[4].

Kilometre-sized and larger planetesimals in protoplanetary disks travel on elliptical orbits that are altered by mutual gravitational interactions and physical collisions. These interactions lead to accretion (and in some cases, erosion and fragmentation) of planetesimals. Gravitational encounters within a 'swarm' of planetesimals can produce velocities that exceed the 'escape velocity' from the largest common planetesimals in the swarm[5], and sufficiently massive and dense planets far enough from the star can hurl material into interstellar space. Comets in the Oort cloud — the vast comet reservoir ~5,000–50,000 AU from the Sun — are believed to be icy planetesimals that were sent outwards at nearly the Solar System escape velocity, and were perturbed into long-lived orbits around the Sun by close stars, interstellar clouds or the tidal forces of the Galactic disk.

We do not at present have enough data to determine the range of planetary systems that occur in nature. The gravitational pull of known extrasolar planets induce velocity variations in their stars much larger than would a planetary system like our own, and surveys so far accomplished have not detected low-mass and long-period planets[6]. Our own Solar System may represent a biased sample of a different kind, because it contains a planet with conditions suitable for life to evolve to the point of being able to ask questions about other planetary systems[7]. Unfortunately, we lack the capability to develop planet-formation models from first principles, or even from observations of protostellar disks, whose detailed properties are poorly known. But theory suggests that a great of diversity of planetary systems are possible[8].

We do, however, possess planet-formation models that can predict the growth times of giant planets (such as Jupiter). Current models[9] predict growth times that are similar to estimates of the lifetime of gaseous protoplanetary disks. Thus, giant

Box 1 Interstellar eavesdropping

The Search for ExtraTerrestrial Intelligence (SETI) is an endeavour to detect signals from alien life forms[21]. A clear detection of such a signal would probably change humanity's world view as much as any other scientific discovery in history. As our society is in its technological infancy, another civilization capable of communicating over interstellar distances is likely to be enormously advanced compared with our own — compare our technology to that of a mere millennium ago and then extrapolate millions or billions of years into the future! Thus, a dialogue with extraterrestrials could alter our society in unimaginable ways (and they could probably answer most if not all of the questions raised in this article).

The primary instrument used by SETI is the radiotelescope. Most radio waves propagate with little loss through the interstellar medium, and many wavelengths also easily pass through the Earth's atmosphere. They are easy to generate and to detect. Radio thus appears to be an excellent means of interstellar communication, whether data are being exchanged between a community of civilizations around different stars or broadcast to the Galaxy in order to reach unknown societies in their technological infancy. Signals used for local purposes, such as radar and television on Earth, also escape and could be detected at great distances.

The first deliberate SETI radiotelescope observations were performed by Frank Drake in 1960. Since that time, improvements in receivers, data processing capabilities and radiotelescopes have doubled the capacity of SETI searches roughly every 8 months. Nonetheless, only a minuscule fraction of directions and frequencies have been searched, so one should not be discouraged by the lack of success so far.

planets might not form in most protoplanetary disks. Although giant planets themselves are unlikely abodes for life, they may harbour habitable moons. Moreover, they affect both the stability of the orbits of Earth-like planets and the flux of materials striking these planets[7]. Such impacts can have a devastating effect on life — a fact that no dinosaur is likely to argue with.

Impacts and dinosaurs

Impacts, like earthquakes, come in various sizes, with the large ones much rarer but vastly more hazardous than the small ones (Table 1). Because of the destruction that impacts may produce, impact frequency is an important factor in planetary habitability. The impact rate on the terrestrial planets of our Solar System was orders of magnitude larger four billion years ago than it is at present. In a planetary system like our own, but with smaller planets replacing Jupiter and Saturn, large impact fluxes could continue, making planets with Earth-like compositions and radiation fluxes hostile abodes for living organisms[7].

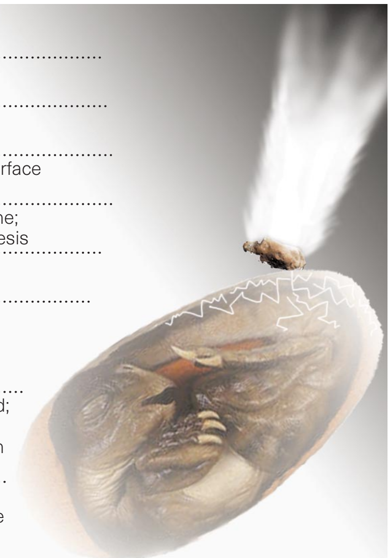
The largest mass extinction of the past 200 million years or so occurred 65 million years ago, when roughly half of the genera of multicellular organisms on Earth, including all of the dinosaurs, suddenly died off. The geological record shows a layer of impact-produced minerals and iridium, an element rare in the Earth's crust but more abundant in primitive meteorites, deposited at the time that the dinosaurs vanished (at the Cretaceous/Tertiary or K/T boundary). Additionally, the largest known crater on Earth dated at less than one billion years old was formed at this time. Taken together, these data imply that this K/T mass extinction was caused by the impact of an asteroid or comet, about 10 km in radius, into the Yucatan peninsula[10].

A deluge of data to come

The speculation about advanced civilizations on Mars that was rampant a century ago has been quashed by better observations. However, evidence for a wet Mars billions of years ago suggests that life might have formed on that planet, and microbes may still be present below the surface[11]. Interplanetary spacecraft will soon attempt to determine whether or not life ever existed on Mars, and if so what were (or are) its characteristics[12]. NASA plans to investigate a possible subsurface ocean in Jupiter's moon Europa[13], and the Cassini mission en route to Saturn will study the properties of Titan's atmosphere[14]. Although Titan's surface is too cold for liquid water, this large moon has a methane-rich atmosphere in which

Table 1 Impacts and life

Size	*Example(s)*	*Most recent*	*Planetary effects*	*Effects on life*
Super colossal $R > 2{,}000$ km	Moon-forming event	4.45×10^9 yr ago	Melts planet	Drives off volatiles; wipes out life on planet
Colossal $R > 700$ km	Pluto 1 Ceres (borderline)	$\gtrsim 4.3 \times 10^9$ yr ago	Melts crust	Wipes out life on planet
Huge $R > 200$ km	4 Vesta (large asteroid)	$\sim 4.0 \times 10^9$ yr ago	Vaporizes oceans	Life may survive below surface
Extra large $R > 70$ km	Chiron (largest active comet)	3.8×10^9 yr ago	Vaporizes upper 100 m of oceans	Pressure-cooks photic zone; may wipe out photosynthesis
Large $R > 30$ km	Comet Hale-Bopp	$\sim 2 \times 10^9$ yr ago	Heats atmosphere and surface to ~1,000 K	Continents cauterized
Medium $R > 10$ km	K/T impactor 433 Eros (largest NEA)	65×10^6 yr ago	Fires, dust, darkness; atmosphere/ocean chemical changes; large temperature swings	Half of species extinct
Small $R > 1$ km	~500 NEAs	~300,000 yr ago	Global dusty atmosphere for months	Photosynthesis interrupted; individuals die but few species extinct; civilization threatened
Very small $R > 100$ m	Tonguska event	91 yr ago	Major local effects; minor hemispheric effects; dusty atmosphere	Newspaper headlines; romantic sunsets increase birth rate



NATURAL HISTORY MUSEUM, LONDON

The effects of an impact on life depend in a qualitative way on the impact energy. The smallest space debris to hit Earth's atmosphere is slowed to benign speeds by gas drag or vaporized before it hits the ground. The largest impactors can melt a planet's crust and eliminate life entirely. Strong iron impactors ranging in size from that of a car to that of a house may hit the ground at high velocity, killing living beings in their path. The rocky bolide that exploded over Tonguska, Siberia, in 1908 was about the size of a football field; it produced a blast wave that knocked over trees tens of kilometres away. An impactor a few kilometres in size would throw enough dust into the upper atmosphere to substantially darken the sky for much of a growing season[19]; the threat of such an impactor to human food supplies has led NASA to initiate a programme to detect all near-Earth asteroids (NEAs) larger than about 1 km. Mass extinctions (such as that at the K/T boundary; see text) result from even larger impacts, which load the atmosphere with dust and chemicals (from vapour and pulverized matter originating in both the impactor and the crater ejecta); radiation from high-velocity ejecta re-entering the atmosphere may cause global fires. Even larger impacts fill the atmosphere with enough hot ejecta and greenhouse gases to vaporize part or all of the planet's oceans[20]. Indeed, phylogenetic evidence implies that the last common ancestor of all life on Earth was a thermophilic prokaryote, which would have been most capable of surviving such a scalding impact. Even larger impacts would destroy all life on the planet, although it is possible that some organisms could survive in meteoroids ejected by the impact, and subsequently re-establish themselves on the planet (or another planet orbiting the same star) by a fortuitously gentle return after conditions on the planet had improved.

photochemical reactions create organic molecules. Analogous processes may have occurred within Earth's early atmosphere.

Interstellar probes for *in situ* exploration of extrasolar planets require substantial technological advances. Even with gravity assists by the planets and the Sun, payloads sent with current rocket propulsion systems would require thousands of years to reach the nearest stars. Thus at present, interstellar travel remains in the realm of science fiction, but considering the advances of the past millennium, voyages over such vast distances may become practical in the coming centuries.

Another approach to the detection of habitable planets is to search for signals that have been sent, intentionally or otherwise, by inhabitants of those worlds. This is being done by the Search for ExtraTerrestrial Intelligence (SETI) (see Box 1).

Although we are not yet able to reach the stars, we are nonetheless entering a golden age of extrasolar planetary study by means of telescopic observations. All of the known extrasolar planets have been identified indirectly, through the gravitational force that they exert on their star, and all have been found during the 1990s. The first two extrasolar planets to be discovered orbit a rapidly spinning neutron star, which emits a substantial fraction of its luminosity as X-rays. The extrasolar planets known to orbit main-sequence stars are each more massive than Saturn. They were discovered using the Doppler technique, which measures changes in the radial velocity of the star in response to the planet's gravitational tug.

Various techniques should increase our knowledge of extrasolar planets in the coming decades. Doppler studies should continue to find giant extrasolar planets[6]: the current precision of this technique is sufficient to detect Jupiter-mass planets in Jupiter-like orbits, and Uranus-mass planets orbiting very close to their stars. Better precision may eventually lead to identification of smaller and more distant planets, but turbulence and other variability of stellar atmospheres will make detection of Earth analogues using the Doppler technique impractical if not impossible.

Planets may also be detected through the wobble that they induce in the motion of their stars projected into the plane of the sky. This astrometric technique is most sensitive to massive planets orbiting stars that are relatively close to the Earth. Because the star's motion is detectable in two dimensions, a better estimate of the planet's mass can be obtained than by using radial velocities. No astrometric claim of detecting an extrasolar planet has yet been confirmed, but technological advances (especially the development of optical and infrared interferometry) suggest that Jupiter-mass (and possibly smaller) planets will be detected from the ground using this technique within the next decade. Even higher precision is likely to be achievable from spacecraft observations. The ultimate limit to astrometric precision is likely to be differences between the positions of a star's centre of mass and its centre of light, which result from 'starspots' and other variations in brightness.

Earth-sized extrasolar planets?

If the Earth lies in or near the orbital plane of an extrasolar planet, that planet passes in front of the disk of its star once each orbit as viewed from Earth. Precise photometry can reveal such transits, which can be distinguished from rotationally modulated 'starspots' and intrinsic stellar variability by their periodicity, and would provide the size and orbital period of the detected planet. Scintillation in, and variability of, the Earth's atmosphere limit photometric precision to roughly one-thousandth of a magnitude, allowing detection from the ground of transits by Jupiter-sized planets but not by Earth-sized planets[15]. Far greater precision is achievable above the atmosphere, with planets as small as Earth likely to be detectable[16]. This technique has the greatest potential for detecting Earth analogues within the next ten years.

Distant planets are very faint objects which are located near much brighter objects (the star or stars that they orbit), and thus they are extremely difficult to image. Efforts are currently underway to image giant planets directly from the ground using adaptive optics (which adjusts telescope mirrors to compensate for the variability of Earth's atmosphere) and coronagraphs (which block the light from the star that the planet orbits). NASA and the European Space Agency, ESA, are both currently designing space missions for the second decade of the twenty-first century that will use interferometry and nulling to search for Earth-like planets around nearby stars[17]. Assuming such planets are detected, spectroscopic investigations of their atmospheres, to search for gases such as oxygen, ozone and methane, will follow. Obtaining resolved images of the disks of extrasolar Earth-like planets requires substantially better optics and much greater light-gathering capability than available at present, but unlike interstellar travel, such distant resolved images should be achievable within the twenty-first century.

Observations of planetary formation provide information on properties of planets as a class, and typical fluxes of material (impactors) to which Earth-like planets are exposed. The infrared region of the spectrum is likely to provide the greatest information about star and planet formation in the coming decades, because planetary regions radiate most of their energy in the mid-infrared and dusty interstellar clouds transmit far more scattered starlight in the near-infrared than in the visible range. Advances in infrared detectors, interferometry and new large telescopes (on the ground with adaptive optics, and in space[18]) will provide data that are vastly superior to those available today.

Prospects

Predictions have a miserable success rate, and forecasts centuries into the future tend to be particularly conservative. In part, this reflects deficiencies in the human imagination. However, the average rate of change greatly exceeds the median rate. Consider the implication of joining the Galactic club of civilizations a billion years more advanced than our own (assuming such a community exists!). Such a revelation would lead to changes far more fundamental than the invention of movable type, the industrial revolution and the information age have brought to us within the past millennium.

We still do not know whether Earth-like planets on which liquid water flows are rare, are usual for solar-type stars or have intermediate abundances. Nonetheless, I personally believe that, as there are billions of Sun-like stars in our Galaxy, planets with liquid water oceans lasting for long periods of time are common enough to ensure that if we are the only advanced life form in our part of the Galaxy, biological factors are much more likely to be the principal limiting factor than are astronomical causes.

Jack J. Lissauer is at the Space Science Division, MS 245-3, NASA Ames Research Center, Moffett Field, California 94035, USA.
e-mail: lissauer@ringside.arc.nasa.gov

1. Lewis, J. S. *Worlds Without End: The Exploration of Planets Known and Unknown* (Helix, Reading, Massachusetts, 1998).
2. Joshi, M. M., Haberle, R. M. & Reynolds, R. T. *Icarus* **129,** 450–465 (1997).
3. Lissauer, J. J. *Rev. Mod. Phys.* **71,** 835–845 (1999).
4. Weidenschilling, S. J. & Cuzzi, J. N. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 1031–1060 (Univ. Arizona Press, Tucson, 1993).
5. Safronov, V. S. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets* (Nauka Press, Moscow, 1969) (in Russian); Publication TTF-677, NASA, 1972 (English transl.).
6. http://www.physics.sfsu.edu/~gmarcy/planetsearch/planetsearch.html
7. Wetherill, G. W. *Astrophys. Space Sci.* **212,** 23–32 (1994).
8. Lissauer, J. J. *Icarus* **114,** 217–236 (1995).
9. Pollack, J. B. *et al. Icarus* **124,** 62–85 (1996).
10. Chapman, C. R. & Morrison, D. *Nature* **367,** 33–40 (1994).
11. McKay, C. P. *Orig. Life Evol. Biosphere* **27,** 263–289 (1997).
12. http://mpfwww.jpl.nasa.gov/
13. http://www.jpl.nasa.gov/ice_fire//europao.htm
14. http://www.jpl.nasa.gov/cassini/
15. http://web99.arc.nasa.gov/~mars/vulcan/
16. http://www.kepler.arc.nasa.gov/
17. http://astro1.bnsc.rl.ac.uk:80/darwin/
18. http://ngst.gsfc.nasa.gov/
19. Toon, O. B., Zahnle, K., Morrison, D., Turco, R. P. & Covey, C. *Rev. Geophys.* **35,** 41–78 (1997).
20. Sleep, N. H., Zahnle, K. J., Kasting, J. F. & Morowitz, H. J. *Nature* **342,** 139–142 (1989).
21. http://www.seti.org/

Acknowledgements. I thank J. Cuzzi, F. Drake and G. Marcy for insightful discussions.