

Three regimes of extrasolar planet radius inferred from host star metallicities

Lars A. Buchhave^{1,2}, Martin Bizzarro², David W. Latham¹, Dimitar Sasselov¹, William D. Cochran³, Michael Endl³, Howard Isaacson⁴, Diana Juncher^{2,5} & Geoffrey W. Marcy⁴

Approximately half of the extrasolar planets (exoplanets) with radii less than four Earth radii are in orbits with short periods¹. Despite their sheer abundance, the compositions of such planets are largely unknown. The available evidence suggests that they range in composition from small, high-density rocky planets to low-density planets consisting of rocky cores surrounded by thick hydrogen and helium gas envelopes. Here we report the metallicities (that is, the abundances of elements heavier than hydrogen and helium) of more than 400 stars hosting 600 exoplanet candidates, and find that the exoplanets can be categorized into three populations defined by statistically distinct ($\sim 4.5\sigma$) metallicity regions. We interpret these regions as reflecting the formation regimes of terrestrial-like planets (radii less than 1.7 Earth radii), gas dwarf planets with rocky cores and hydrogen–helium envelopes (radii between 1.7 and 3.9 Earth radii) and ice or gas giant planets (radii greater than 3.9 Earth radii). These transitions correspond well with those inferred from dynamical mass estimates^{2,3}, implying that host star metallicity, which is a proxy for the initial solids inventory of the protoplanetary disk, is a key ingredient regulating the structure of planetary systems.

Shortly after the discovery of the first exoplanets, host star metallicity was suggested to have a role in the formation of planetary systems⁴. Indeed, the well-established tendency for hot Jupiters to be more frequently found orbiting metal-rich stars has been confirmed by a number of studies^{5,6}. Although it has recently been shown that small planets form for a wide range of host star metallicities^{7–11}, it is clear that the metallicities of stars with small planets are on average lower than those of gas giants. This suggests that subtle differences may exist in the metallicities of the host stars of small exoplanets, and this, in turn, may be linked to distinct physical properties of the underlying planet populations. However, effectively probing this regime requires a large sample of homogeneously derived metallicities for stars with small planets. Therefore, using our stellar parameters classification (SPC) tool⁷ (Methods Summary), we analyse more than 2,000 high-resolution spectra of Kepler Objects of Interest¹² (KOIs) gathered by the Kepler Follow-up Program, yielding precise stellar parameters (provided as a table in machine-readable form), including metallicities, of 405 stars orbited by 600 exoplanet candidates.

Our sample of spectroscopic metallicities of stars hosting small planets is a factor of two larger than any previous sample⁷, allowing us to probe in greater detail for significant differences in the metallicities of stars hosting planets of different sizes. At various radii, we divide the sample into two bins of stars hosting small and large planets, and perform a two-sample Kolmogorov–Smirnov test to determine whether the metallicities of the two distributions of host stars are not drawn randomly from the same parent population. We find two significant features in the Kolmogorov–Smirnov test diagram, one at $1.7R_{\oplus}$ (Earth radii) with a significance of 4.5σ and one at $3.9R_{\oplus}$ with a significance of 4.6σ , suggesting transitions between three exoplanet size regimes (Fig. 1). The average metallicity of the host stars increases with planet

size, yielding average metallicities of -0.02 ± 0.02 , 0.05 ± 0.01 and 0.18 ± 0.02 dex in the respective regimes. To assess the uncertainty in radius at which these transitions occur, we perform a Monte Carlo simulation by drawing 10^5 sets of data, where the host star metallicities and planetary radii are randomly perturbed within the uncertainties (the uncertainty in the planetary radius is assumed to be dominated by the uncertainty in the radius of the host star). We find the two features to be at $1.55^{+0.88}_{-0.04}R_{\oplus}$ ($4.2^{+0.5}_{-0.4}\sigma$) and $3.52^{+0.74}_{-0.28}R_{\oplus}$ ($4.7^{+0.6}_{-0.4}\sigma$), consistent with the original data.

Small planets with short periods could undergo significant evaporation of their atmospheres¹³. These planets will therefore not obey the radius–metallicity relation we are studying, because any accumulated gas would have evaporated. Therefore, we remove small ($R_p < 3R_{\oplus}$, where R_p is the exoplanet radius), highly irradiated planets (stellar flux, $F_v > 5 \times 10^5 \text{ J s}^{-1} \text{ m}^{-2}$) from the sample, leaving 463 planets orbiting 324 stars, which increases the significance of the feature at $1.7R_{\oplus}$ from 3.5σ to the reported 4.5σ . For comparison, the rocky planet Kepler-10b, with an 0.8-d period¹⁴, receives a flux of $F_v \approx 48 \times 10^5 \text{ J s}^{-1} \text{ m}^{-2}$, whereas Kepler-11c¹⁵, whose density suggests it is gaseous, receives $F_v \approx 1.3 \times 10^5 \text{ J s}^{-1} \text{ m}^{-2}$.

Recent studies suggest that the masses and radii of small planets ($1.5R_{\oplus}$ – $4R_{\oplus}$) follow a linear relationship, implying that planet density decreases with increasing planet radius². However, this relationship must change significantly for larger planets ($>4R_{\oplus}$) to explain the large mass of gas giant planets such as Jupiter². Our data indicate a statistically significant increase in metallicity at a comparable planetary radius, $R_p = 3.9R_{\oplus}$. This observation is in agreement with the well-established correlation between a star's metallicity and its likelihood to host hot Jupiters^{5,6}, confirming that the formation regime for larger planets ($>3.9R_{\oplus}$) requires exceptionally high metallicity environments⁷. We therefore interpret the regime of larger planets ($R_p > 3.9R_{\oplus}$) to consist of ice and gas giant planets formed beyond the 'snow line' at around $\sim 3 \text{ AU}$ (1 AU is the average Sun–Earth distance), where the availability of solids is a factor of four higher because volatile elements are able to condense and form solids at cooler temperatures¹⁶. The high concentration of heavy elements in the protoplanetary disk, resulting from the higher metallicity and the condensation of volatiles, and the planet's large distance to the host star allow these planets to grow rapidly and amass a gaseous atmosphere before the gas in the protoplanetary disk dissipates. These planets then migrate to their present positions much closer to their host stars¹⁷, yielding the hot Jupiters seen orbiting stars with high metallicities.

To explore the implications of the feature at $1.7R_{\oplus}$ in the metallicity–radius plane (Fig. 1), we compare our results with recent work attempting to determine the radius at which the transition from gaseous to rocky planets occurs. Dynamical masses derived from precise radial velocities of transiting exoplanets indicate that planets with $R_p > 2R_{\oplus}$ have densities that imply increasing amounts by volume of light material, whereas planets with $R_p < 1.5R_{\oplus}$ have densities systematically

¹Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA. ²Centre for Star and Planet Formation, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen, Denmark. ³McDonald Observatory, The University of Texas, Austin, Texas 78712, USA. ⁴University of California, Berkeley, California 94720, USA. ⁵Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark.

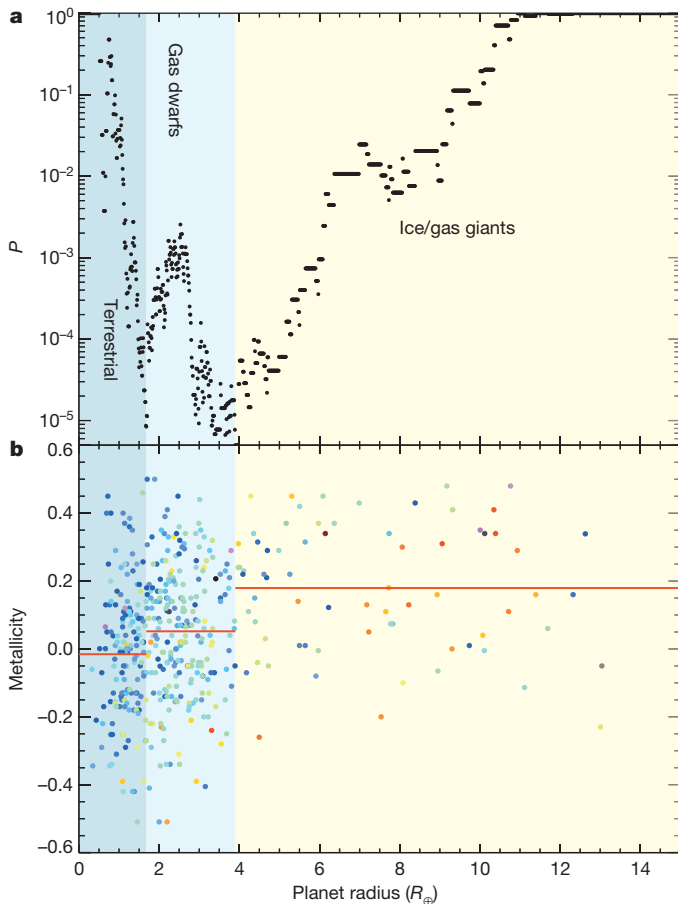


Figure 1 | Host star metallicities and three types of exoplanets with different composition. **a**, P value of the two-sample Kolmogorov–Smirnov test. **b**, Radii of the individual planets and their host star metallicities. Point colour represents the logarithm of the period of the planets (blue, shortest period; red, longest period). The solid red lines are the average metallicities in the three regions (-0.02 ± 0.02 , 0.05 ± 0.01 and 0.18 ± 0.02 dex, where each uncertainty is 1 s.e.m. of the host star metallicities in the corresponding bin).

greater than that of Earth². Moreover, an analysis of data for a larger sample of planets (including masses derived from transit timing variations), has shown that planets with $R_p < 1.5R_{\oplus}$ probably are of rocky composition³. Finally, it has been suggested that $R_p \approx 1.75R_{\oplus}$ is a physically motivated transition point between rocky and gaseous planets, based on reported masses and radii combined with thermal evolutionary atmosphere models¹⁸. The statistically significant peak in the metallicity–radius plane at $1.7R_{\oplus}$ agrees with these findings, suggesting that the compositions of small exoplanets ($R_p < 3.9R_{\oplus}$) in close proximity to their host stars are also regulated by the number density of solids in the protoplanetary disk. Thus, we interpret the two regimes of smaller planets identified by the host star metallicities as reflecting the transition between rocky terrestrial exoplanets that have not amassed a gaseous atmosphere ($R_p < 1.7R_{\oplus}$) and planets with rocky cores that have accumulated an envelope of hydrogen, helium and other volatiles, which we denote gas dwarfs ($1.7R_{\oplus} < R_p < 3.9R_{\oplus}$).

The formation mechanism of the terrestrial and gas dwarf exoplanet regimes in short orbital periods is not fully understood. In one model, these small exoplanets are believed to form *in situ* with little post-assembly migration^{19,20}. Although the *in situ* accretion model seems to be successful in reproducing the observed distribution of the ‘hot Neptune’ and super-Earth systems, including their orbital spacing²¹, it requires unusually large amounts of solids in the innermost protoplanetary disk. A competing model invokes accretion during the inward migration of a population of planetary embryos formed at a range of

orbital distances beyond the snow line^{22,23}. On this view, Mars- and Earth-size embryos migrate inwards owing to tidal interaction with the disk²⁴, and accumulate at the inner edge of the protoplanetary disk, where they complete their assembly²⁵. Irrespective of the formation mechanism, however, the observed peak in the metallicity–radius plane at $1.7R_{\oplus}$ suggests that the final mass and composition of a small exoplanet is controlled by the amount of solid material available in the protoplanetary disk. A higher-metallicity environment promotes a more rapid and effective accretion process, thereby allowing the cores to amass a gaseous envelope before dissipation of the gas. In contrast, lower-metallicity environments may result in the assembly of rocky cores of several Earth masses on timescales greater than that inferred for gas dispersal in protoplanetary disks¹⁹ (< 10 Myr), yielding cores without gaseous hydrogen–helium atmospheres.

A prediction from gas accretion on short orbital periods is that the critical mass at which a core can accrete an atmosphere is $M_{cr} \approx 2.6M_{\oplus}(\eta/0.3)^{1/2}(P_{orb}/1 \text{ d})^{5/12}$, where $\eta = M_{atm}/M_{cr}$ is the fractional mass comprised by the atmosphere²⁶. In this model, the planetary mass and, thus, radius indicating the transition from rocky to gaseous planets should increase with orbital period. However, the exact opposite dependence, namely a decrease in core mass with increased orbital period, has also been suggested²⁷. To investigate whether the radius of transition from rocky to gaseous planets found in our data shows a dependence on orbital period, we segregate the sample by period into four bins with approximately equal numbers of planets. We repeat the previously described Kolmogorov–Smirnov test for the planets in each of the period bins: we remove the larger planets ($R_p > 3.9R_{\oplus}$) and perform a Monte Carlo simulation by drawing 10^6 sets of data, where the host star metallicities and planetary radii are randomly perturbed within the uncertainties. The red line in Fig. 2 is a power-law fit to the transition radius, R_{cr} , inferred from our data ($R_{cr} = 1.06R_{\oplus}(P_{orb}/1 \text{ day})^{0.17}$). Although additional data are required to confirm this relationship, the fit is apparently consistent with a critical core mass that increases with orbital period and an atmospheric fraction of 5% (ref. 26; blue dashed line in Fig. 2). If correct, this predicts the existence of more massive rocky exoplanets at longer orbital periods.

Although our analysis of a statistically significant number of planets and their host star metallicities allows us to distinguish between three distinct exoplanet regimes, we emphasize that a multitude of factors can affect the outcome of planet formation. Therefore, the transition radii inferred from our analysis probably represent gradual transitions

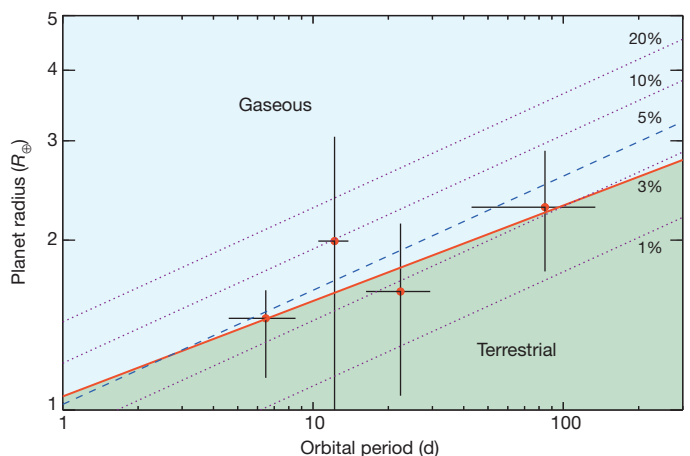


Figure 2 | The radius of transition from rocky to gaseous exoplanets. The transition radii (red points) are the means of the posterior distributions resulting from the Monte Carlo analysis (main text), and the error bars indicate the 1σ uncertainties. Using the mass–radius approximation²⁹ $M/M_{\oplus} = (R/R_{\oplus})^{2.06}$, we plot the radius corresponding to M_{cr} with an atmosphere fraction of 5% as the blue dashed line and those for respective atmosphere fractions of 1%, 3%, 10% and 20% as the dotted purple lines. The solid red line is a power-law fit to the Monte Carlo data: $R_{cr} = 1.06R_{\oplus}(P_{orb}/1 \text{ day})^{0.17}$.

between the different planet regimes and so may not apply to all planetary systems. However, the agreement between the transition radii inferred here and those deduced from dynamical mass measurements of transiting planets^{2,3} implies that host star metallicity—and, by extension, the solids inventory of a protoplanetary disk—is one of the driving factors determining the outcome of planet formation.

METHODS SUMMARY

We use SPC⁷ for the spectroscopic analysis yielding the stellar parameters in this work. SPC uses a grid of synthetic library spectra to derive the effective temperature, surface gravity, metallicity and rotational velocity simultaneously by matching the models with the observed spectra originating from a number of different instruments (Methods). The stellar parameters from SPC and Yonsei–Yale stellar evolutionary models²⁸ are used to estimate the radii of the host stars, which we couple with the photometric data from the Kepler mission to infer the planet radii. The majority of planets in the sample have short orbital periods, owing to the observational bias of Kepler towards shorter period planets. The mean and median periods are 38.0 and 12.4 d, respectively.

To investigate whether the radius of transition from rocky to gaseous planets found in our data shows a dependence on orbital period (Fig. 2), we remove the larger planets ($R_p > 3.9R_\oplus$) and segregate the remainder of the sample by period into four bins with approximately equal numbers of planets. The previously described Kolmogorov–Smirnov test (Fig. 1) is repeated for each bin by performing a Monte Carlo simulation in which the host star metallicities and planetary radii are randomly perturbed within the uncertainties. The data points plotted in Fig. 2 are the means of the resulting posterior distributions, and the error bars indicate the 1σ uncertainties.

To ensure that the statistical analyses are sound, we perform further tests evaluating the effects of uneven sampling and contamination (Methods). We find results consistent with those presented in the paper within the uncertainties, and conclude that our statistical analyses are robust.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 November 2013; accepted 12 March 2014.

- Fressin, F. *et al.* The false positive rate of Kepler and the occurrence of planets. *Astrophys. J.* **766**, 81 (2013).
- Marcy, G. W. *et al.* Masses, radii, and orbits of small Kepler planets: the transition from gaseous to rocky planets. *Astrophys. J. Suppl. Ser.* **210**, 20 (2014).
- Weiss, L. M. & Marcy, G. W. The mass-radius relation for 65 exoplanets smaller than 4 Earth radii. *Astrophys. J.* **783**, L6 (2014).
- Gonzalez, G. The stellar metallicity-giant planet connection. *Mon. Not. R. Astron. Soc.* **285**, 403–412 (1997).
- Santos, N. C., Israelian, G. & Mayor, M. Chemical analysis of 8 recently discovered extra-solar planet host stars. *Astron. Astrophys.* **363**, 228–238 (2000).
- Fischer, D. A. & Valenti, J. The planet-metallicity correlation. *Astrophys. J.* **622**, 1102–1117 (2005).
- Buchhave, L. A. *et al.* An abundance of small exoplanets around stars with a wide range of metallicities. *Nature* **486**, 375–377 (2012).
- Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M. & Udry, S. Spectroscopic stellar parameters for 582 FGK stars in the HARPS volume-limited sample. Revising the metallicity-planet correlation. *Astron. Astrophys.* **533**, A141 (2011).
- Sousa, S. G. *et al.* Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes. *Astron. Astrophys.* **487**, 373–381 (2008).
- Ghezzi, L. *et al.* Stellar parameters and metallicities of stars hosting Jovian and Neptunian mass planets: a possible dependence of planetary mass on metallicity. *Astrophys. J.* **720**, 1290–1302 (2010).
- Everett, M. E., Howell, S. B., Silva, D. R. & Szkody, P. Spectroscopy of faint Kepler mission exoplanet candidate host stars. *Astrophys. J.* **771**, 107 (2013).
- Batalha, N. M. *et al.* Planetary candidates observed by Kepler. III. Analysis of the first 16 months of data. *Astrophys. J. Suppl. Ser.* **204**, 24 (2013).
- Owen, J. E. & Wu, Y. Kepler planets: a tale of evaporation. *Astrophys. J.* **775**, 105 (2013).
- Batalha, N. M. *et al.* Kepler's first rocky planet: Kepler-10b. *Astrophys. J.* **729**, 27 (2011).
- Lissauer, J. J. *et al.* All six planets known to orbit Kepler-11 have low densities. *Astrophys. J.* **770**, 131 (2013).
- Ida, S. & Lin, D. N. C. Toward a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities. *Astrophys. J.* **616**, 567–572 (2004).
- Lin, D. N. C., Bodenheimer, P. & Richardson, D. C. Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature* **380**, 606–607 (1996).
- Lopez, E. D. & Fortney, J. J. Understanding the mass-radius relation for sub-Neptunes: radius as a proxy for composition. Preprint at <http://arxiv.org/abs/1311.0329> (2013).
- Hansen, B. M. S. & Murray, N. Migration then assembly: formation of Neptune-mass planets inside 1 AU. *Astrophys. J.* **751**, 158 (2012).
- Chiang, E. & Laughlin, G. The minimum-mass extrasolar nebula: in situ formation of close-in super-Earths. *Mon. Not. R. Astron. Soc.* **431**, 3444–3455 (2013).
- Hansen, B. M. S. & Murray, N. Testing in situ assembly with the Kepler planet candidate sample. *Astrophys. J.* **775**, 53 (2013).
- Terquem, C. & Papaloizou, J. C. B. Migration and the formation of systems of hot super-Earths and Neptunes. *Astrophys. J.* **654**, 1110–1120 (2007).
- McNeil, D. S. & Nelson, R. P. On the formation of hot Neptunes and super-Earths. *Mon. Not. R. Astron. Soc.* **401**, 1691–1708 (2010).
- Goldreich, P. & Tremaine, S. The excitation of density waves at the Lindblad and corotation resonances by an external potential. *Astrophys. J.* **233**, 857–871 (1979).
- Raymond, S. N. & Cossou, C. No universal minimum-mass extrasolar nebula: Evidence against in-situ accretion of systems of hot super-Earths. Preprint at <http://arxiv.org/abs/1401.3743> (2014).
- Rafikov, R. R. Atmospheres of protoplanetary cores: critical mass for nucleated instability. *Astrophys. J.* **648**, 666–682 (2006).
- Papaloizou, J. C. B. & Terquem, C. Critical protoplanetary core masses in protoplanetary disks and the formation of short-period giant planets. *Astrophys. J.* **521**, 823–838 (1999).
- Yi, S. *et al.* Toward better age estimates for stellar populations: the Y^2 isochrones for solar mixture. *Astrophys. J. Suppl. Ser.* **136**, 417–437 (2001).
- Lissauer, J. J. *et al.* Architecture and dynamics of Kepler's candidate multiple transiting planet systems. *Astrophys. J. Suppl. Ser.* **197**, 8 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements L.A.B. acknowledges support from the Harvard Origins of Life Initiative. M.B. acknowledges funding from the Danish National Research Foundation (grant number DNR97) and from the European Research Council under ERC Consolidator grant agreement 616027- STARDUST2ASTEROIDS. D.W.L. acknowledges support from the Kepler Mission under NASA Cooperative Agreements NCC2-1390, NNX11AB99A and NNX13AB58A with the Smithsonian Astrophysical Observatory, and thanks the observers who helped obtain the TRES observations reported here, especially R. Stefanik, G. Esquerdo, P. Berlind and M. Calkins.

Author Contributions L.A.B. led the project and developed the classification tools for the metallicity analysis. M.B., D.W.L. and D.S. contributed to the discussion of the theoretical implications of the data. L.A.B., D.W.L., W.D.C., M.E., H.I., D.J. and G.W.M. worked on gathering the spectroscopic observations. All authors discussed the results and commented on the manuscript. L.A.B. and M.B. wrote the paper with input from D.W.L. and D.S.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.A.B. (lbuchhave@cfa.harvard.edu).

METHODS

Observations and stellar parameters. This study is based on stellar classifications by SPC⁷ of 2,297 spectra observed using the Fibre-fed Echelle Spectrograph on the 2.6-m Nordic Optical Telescope on La Palma, Spain (488 spectra), the fibre-fed Tillinghast Reflector Echelle Spectrograph on the 1.5-m Tillinghast Reflector at the Fred Lawrence Whipple Observatory on Mt Hopkins, Arizona (985 spectra), the Tull Coudé Spectrograph on the 2.7-m Harlan J. Smith Telescope at the McDonald Observatory Texas (653 spectra) and the HIRES spectrograph on the 10-m Keck I telescope at Mauna Kea, Hawaii (171 spectra). We included only the most secure stellar classifications by limiting our sample to stars with effective temperatures of $4,800\text{ K} < T_{\text{eff}} < 6,500\text{ K}$, projected rotational velocities of $v\sin(i) < 20\text{ km s}^{-1}$, spectra with signal-to-noise ratios per resolution element of more than 25, and normalized cross-correlation function peak heights of more than 0.9 (indicating the quality of the stellar classification).

We improve on the determination of the surface gravity, known to be prone to degeneracies with effective temperature and metallicity³⁰, by imposing a prior on the surface gravity from stellar evolutionary models and an initial estimate of the star's effective temperature and metallicity. This is particularly useful for cooler stars, where the evolutionary models put tight constraints on the surface gravity. We use the stellar parameters from SPC and the Yonsei–Yale stellar evolutionary models²⁸ to estimate the radii of the host stars, and, using the photometrically derived planet radii from Kepler, we correct the planetary radii based on the Kepler Input Catalogue photometry, which are known to be prone to systematic biases and large uncertainties. The improved stellar radii reduce the uncertainties in the planetary

radii from an average error of 34% to one of 11%, assuming that the major contribution to the uncertainty in the planetary radii originates from the stellar radii.

Uneven sampling. To investigate the effect of uneven sample size, we performed a Monte Carlo test with 10^6 realizations where we randomly drew observations from the smaller of the two samples, making each of the two samples equal in size at all times. We find the ice or gas giant transition to be at $3.58^{+0.75}_{-0.38}R_{\oplus}$ with a significance of $4.9^{+0.5}_{-0.4}\sigma$ and the rocky transition to be at $1.60^{+0.83}_{-0.10}R_{\oplus}$ with a significance of $4.4^{+0.5}_{-0.4}\sigma$. Both values are consistent with the Monte Carlo analysis reported in the paper. We conclude that the uneven sample size does not affect the significance of our statistical analysis.

Contamination. We are using the two-sample Kolmogorov–Smirnov test for the statistical analysis, but we find three distinct populations of exoplanets. To establish whether contamination from the third sample affects our results, we remove the planets from the third sample (removing $R_p > 3.9R_{\oplus}$ when searching for the peak at $1.7R_{\oplus}$ and removing $R_p < 1.7R_{\oplus}$ for the peak at $3.9R_{\oplus}$) and subsequently carry out the Kolmogorov–Smirnov test in an attempt to recover each of the peaks in Fig. 1. We find the first transition to be at the same radius ($1.7R_{\oplus}$), albeit with a lower significance (3.1σ). We find the ice/gas giant transition close to the one reported in the paper, again with a slightly lower significance ($4.2R_{\oplus}$ at 3.9σ). Again, evaluation of our data using a different approach supports our results and conclusions.

30. Torres, G. *et al.* Improved spectroscopic parameters for transiting planet hosts. *Astrophys. J.* **757**, 161 (2012).