Corrections

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Correction for "Mechanism of E-cadherin dimerization probed by NMR relaxation dispersion," by Ying Li, Nicole L. Altorelli, Fabiana Bahna, Barry Honig, Lawrence Shapiro, and Arthur G. Palmer III, which appeared in issue 41, October 8, 2013, of *Proc Natl Acad Sci USA* (110:16462–16467; first published September 25, 2013; 10.1073/pnas.1314303110).

The authors note that the second sentence in the Acknowledgments, "This work was supported by National Institutes of Health (NIH) Grants GM059273 (to A.G.P.) and GM062270 (to L.S.)" should instead appear as "This work was supported by National Institutes of Health (NIH) Grants GM059273 (to A.G.P.) and GM062270 (to L.S.) and by National Science Foundation Grant MCB-0918535 (to B.H.)."

www.pnas.org/cgi/doi/10.1073/pnas.1319465110

DEVELOPMENTAL BIOLOGY

Correction for "A Hox gene controls lateral line cell migration by regulating chemokine receptor expression downstream of Wnt signaling," by Marie A. Breau, David G. Wilkinson, and Qiling Xu, which appeared in issue 42, October 15, 2013, of *Proc Natl Acad Sci USA* (110:16892–16897; first published September 30, 2013; 10.1073/pnas.1306282110).

The authors note that, due to a printer's error, refs. 33–38 were numbered incorrectly. The citations to the reference numbers are correct in the text. Below is the correct order for refs. 33–38.

33. Gamba L, Cubedo N, Ghysen A, Lutfalla G, Dambly-Chaudière C (2010) Estrogen receptor ESR1 controls cell migration by repressing chemokine receptor CXCR4 in the zebrafish posterior lateral line system. *Proc Natl Acad Sci USA* 107(14):6358–6363.
34. Westerfield M (1993) *The Zebrafish Book* (Univ of Oregon Press, Eugene).
35. Kwan KM, et al. (2007) The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* 236(11):3088–3099.
36. Xu Q, Wilkinson DG (1998) *In situ hybridization of mRNA with hapten labelled probes. In Situ Hybridization: A Practical Approach*, ed Wilkinson DG (Oxford Univ Press, Oxford), 2nd Ed, pp 87–106.
37. Robu ME, et al. (2007) p53 activation by knockdown technologies. *PLoS Genet* 3(5): e78.
38. Gerety SS, Wilkinson DG (2011) Morpholino artifacts provide pitfalls and reveal a novel role for pro-apoptotic genes in hindbrain boundary development. *Dev Biol* 350(2):279–289.

www.pnas.org/cgi/doi/10.1073/pnas.1319266110

IMMUNOLOGY

Correction for "IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication," by Martin R. Jakobsen, Rasmus O. Bak, Annika Andersen, Randi K. Berg, Søren B. Jensen, Jin Tengchuan, Anders Laustsen, Kathrine Hansen, Lars Østergaard, Katherine A. Fitzgerald, T. Sam Xiao, Jacob G. Mikkelsen, Trine H. Mogensen, and Søren R. Paludan, which appeared in issue 48, November 26, 2013, of *Proc Natl Acad Sci USA* (110:E4571–E4580; first published October 23, 2013; 10.1073/pnas.1311669110).

The authors note that the author name Jin Tengchuan should instead appear as Tengchuan Jin. The corrected author line appears below. The online and print versions have been corrected.

Martin R. Jakobsen, Rasmus O. Bak, Annika Andersen, Randi K. Berg, Søren B. Jensen, Tengchuan Jin, Anders Laustsen, Kathrine Hansen, Lars Østergaard, Katherine A. Fitzgerald, T. Sam Xiao, Jacob G. Mikkelsen, Trine H. Mogensen, and Søren R. Paludan

www.pnas.org/cgi/doi/10.1073/pnas.1320190110

MATHEMATICS

Correction for "Violating the Shannon capacity of metric graphs with entanglement," by Jop Briët, Harry Buhrman, and Dion Gijswijt, which appeared in issue 48, November 26, 2013, of *Proc Natl Acad Sci USA* (110:19227–19232; first published December 24, 2012; 10.1073/pnas.1203857110).

The authors note that, due to a printer's error, on page 19227, left column, first full paragraph, line 6 "G_n" should instead appear as "$a \in R$".

Also, on page 19227, left column, first full paragraph, line 7 "$x \in V(H_n)$" should instead appear as "$x \in S$". Both the online article and the print article have been corrected.

www.pnas.org/cgi/doi/10.1073/pnas.1301191110

CORRECTIONS

ASTRONOMY

Correction for "Prevalence of Earth-size planets orbiting Sun-like stars," by Erik A. Petigura, Andrew W. Howard, and Geoffrey W. Marcy, which appeared in issue 48, November 26, 2013, of *Proc Natl Acad Sci USA* (110:19273–19278; first published November 4, 2013; 10.1073/pnas.1319909110).

The authors note that the following statement should be added as a Note Added in Proof: "Estimates of the occurrence of Earth analog planets appear in several previous works including Catanzarite and Shao (25), Traub (26), and Dong and Zhu (27). These estimates, which range from 1% to 34%, were built upon early catalogs of *Kepler* planet candidates (based on less than 1.3 years of photometry). These estimates did not address survey completeness with injection and recovery or uncertain stellar radii with spectroscopy." The online version has been updated to include the following three references:

25. Catanzarite J, Shao M (2011) The occurrence rate of earth analog planets orbiting sun-like stars. *Astrophys J* 738(2):151–160.
26. Traub W (2012) Terrestrial, habitable-zone exoplanet frequency from Kepler. *Astrophys J* 745(1):20–29.
27. Dong S, Zhu Z (2013) Fast rise of "Neptune-size" planets (4-8 R_Earth) from P ∼ 10 to ∼250 days—Statistics of Kepler planet candidates up to ∼0.75 AU. *Astrophys J* 778(1): 53–63.

www.pnas.org/cgi/doi/10.1073/pnas.1321363110

Prevalence of Earth-size planets orbiting Sun-like stars

Erik A. Petigura[a,b,1], **Andrew W. Howard**[b], **and Geoffrey W. Marcy**[a]

[a]Astronomy Department, University of California, Berkeley, CA 94720; and [b]Institute for Astronomy, University of Hawaii at Manoa, Honolulu, HI 96822

Contributed by Geoffrey W. Marcy, October 22, 2013 (sent for review October 18, 2013)

Determining whether Earth-like planets are common or rare looms as a touchstone in the question of life in the universe. We searched for Earth-size planets that cross in front of their host stars by examining the brightness measurements of 42,000 stars from National Aeronautics and Space Administration's *Kepler* mission. We found 603 planets, including 10 that are Earth size ($1-2\ R_{\oplus}$) and receive comparable levels of stellar energy to that of Earth ($0.25-4\ F_{\oplus}$). We account for *Kepler's* imperfect detectability of such planets by injecting synthetic planet–caused dimmings into the *Kepler* brightness measurements and recording the fraction detected. We find that 11 ± 4% of Sun-like stars harbor an Earth-size planet receiving between one and four times the stellar intensity as Earth. We also find that the occurrence of Earth-size planets is constant with increasing orbital period (*P*), within equal intervals of log*P* up to ∼200 d. Extrapolating, one finds $5.7^{+1.7}_{-2.2}$% of Sun-like stars harbor an Earth-size planet with orbital periods of 200–400 d.

extrasolar planets | astrobiology

The National Aeronautics and Space Administration's (NASA's) *Kepler* mission was launched in 2009 to search for planets that transit (cross in front of) their host stars (1–4). The resulting dimming of the host stars is detectable by measuring their brightness, and *Kepler* monitored the brightness of 150,000 stars every 30 min for 4 y. To date, this exoplanet survey has detected more than 3,000 planet candidates (4).

The most easily detectable planets in the *Kepler* survey are those that are relatively large and orbit close to their host stars, especially those stars having lower intrinsic brightness fluctuations (noise). These large, close-in worlds dominate the list of known exoplanets. However, the *Kepler* brightness measurements can be analyzed and debiased to reveal the diversity of planets, including smaller ones, in our Milky Way Galaxy (5–7). These previous studies showed that small planets approaching Earth size are the most common, but only for planets orbiting close to their host stars. Here, we extend the planet survey to *Kepler's* most important domain: Earth-size planets orbiting far enough from Sun-like stars to receive a similar intensity of light energy as Earth.

Planet Survey

We performed an independent search of *Kepler* photometry for transiting planets with the goal of measuring the underlying occurrence distribution of planets as a function of orbital period, P, and planet radius, R_P. We restricted our survey to a set of Sun-like stars (GK type) that are the most amenable to the detection of Earth-size planets. We define GK-type stars as those with surface temperatures $T_{\mathrm{eff}} = 4{,}100$–$6{,}100$ K and gravities log$g$ = 4.0–4.9 (log$g$ is the base 10 logarithm of a star's surface gravity measured in cm s^{-2}) (8). Our search for planets was further restricted to the brightest Sun-like stars observed by *Kepler* (Kp = 10–15 mag). These 42,557 stars (Best42k) have the lowest photometric noise, making them amenable to the detection of Earth-size planets. When a planet crosses in front of its star, it causes a fractional dimming that is proportional to the fraction of the stellar disk blocked, $\delta F = (R_P/R_*)^2$, where R_* is the radius of the star. As viewed by a distant observer, the Earth dims the Sun by ∼100 parts per million (ppm) lasting 12 h every 365 d.

We searched for transiting planets in *Kepler* brightness measurements using our custom-built TERRA software package described in previous works (6, 9) and in *SI Appendix*. In brief, TERRA conditions *Kepler* photometry in the time domain, removing outliers, long timescale variability (>10 d), and systematic errors common to a large number of stars. TERRA then searches for transit signals by evaluating the signal-to-noise ratio (SNR) of prospective transits over a finely spaced 3D grid of orbital period, P, time of transit, t_0, and transit duration, ΔT. This grid-based search extends over the orbital period range of 0.5–400 d.

TERRA produced a list of "threshold crossing events" (TCEs) that meet the key criterion of a photometric dimming SNR ratio SNR > 12. Unfortunately, an unwieldy 16,227 TCEs met this criterion, many of which are inconsistent with the periodic dimming profile from a true transiting planet. Further vetting was performed by automatically assessing which light curves were consistent with theoretical models of transiting planets (10). We also visually inspected each TCE light curve, retaining only those exhibiting a consistent, periodic, box-shaped dimming, and rejecting those caused by single epoch outliers, correlated noise, and other data anomalies. The vetting process was applied homogeneously to all TCEs and is described in further detail in *SI Appendix*.

To assess our vetting accuracy, we evaluated the 235 *Kepler* objects of interest (KOIs) among Best42k stars having P > 50 d, which had been found by the *Kepler* Project and identified as planet candidates in the official Exoplanet Archive (exoplanetarchive.ipac.caltech.edu; accessed 19 September 2013). Among them, we found four whose light curves are not consistent with being planets. These four KOIs (364.01, 2,224.02, 2,311.01, and 2,474.01) have long periods and small radii (*SI Appendix*). This exercise suggests that our vetting process is robust and that careful scrutiny of the light curves of small planets in long period orbits is useful to identify false positives.

Significance

A major question is whether planets suitable for biochemistry are common or rare in the universe. Small rocky planets with liquid water enjoy key ingredients for biology. We used the National Aeronautics and Space Administration *Kepler* telescope to survey 42,000 Sun-like stars for periodic dimmings that occur when a planet crosses in front of its host star. We found 603 planets, 10 of which are Earth size and orbit in the habitable zone, where conditions permit surface liquid water. We measured the detectability of these planets by injecting synthetic planet-caused dimmings into *Kepler* brightness measurements. We find that 22% of Sun-like stars harbor Earth-size planets orbiting in their habitable zones. The nearest such planet may be within 12 light-years.

Author contributions: E.A.P., A.W.H., and G.W.M. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The *Kepler* photometry is available at the Milkulski Archive for Space Telescopes (archive.stsci.edu). All spectra are available to the public on the Community Follow-up Program website (cfop.ipac.caltech.edu).

[1]To whom correspondence should be addressed. E-mail: epetigura@berkeley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1319909110/-/DCSupplemental.

Vetting of our TCEs produced a list of 836 eKOIs, which are analogous to KOIs produced by the *Kepler* Project. Each light curve is consistent with an astrophysical transit but could be due to an eclipsing binary (EB), either in the background or gravitationally bound, instead of a transiting planet. If an EB resides within the software aperture of a *Kepler* target star (within ∼10 arcsec), the dimming of the EB can masquerade as a planet transit when diluted by the bright target star. We rejected as likely EBs any eKOIs with these characteristics: radii larger than 20 $R_{\oplus}$, observed secondary eclipse, or astrometric motion of the target star in and out of transit (*SI Appendix*). This rejection of EBs left 603 eKOIs in our catalog.

Kepler photometry can be used to measure R_P/R_* with high precision, but the extraction of planet radii is compromised by poorly known radii of the host stars (11). To determine R_* and T_{eff}, we acquired high-resolution spectra of 274 eKOIs using the HIRES spectrometer on the 10-m Keck I telescope. Notably, we obtained spectra of all 62 eKOIs that have $P > 100$ d. For these stars, the ∼35% errors in R_* were reduced to ∼10% by matching spectra to standards.

To measure planet occurrence, one must not only detect planets but also assess what fraction of planets were missed. Missed planets are of two types: those whose orbital planes are so tilted as to avoid dimming the star and those whose transits were not detected in the photometry by TERRA. Both effects can be quantified to establish a statistical correction factor. The first correction can be computed as the geometrical probability that an orbital plane is viewed edge-on enough (from Earth) that the planet transits the star. This probability is $P_{\mathrm{T}} = R_*/a$, where a is the semimajor axis of the orbit.

The second correction is computed by the injection and recovery of synthetic (mock) planet-caused dimmings into real *Kepler* photometry. We injected 40,000 transit-like synthetic dimmings having randomly selected planetary and orbital properties into the actual photometry of our Best42k star sample, with stars selected at random. We measured survey completeness, $C(P, R_P)$, in small bins of (P, R_P), determining the fraction of injected synthetic planets that were discovered by TERRA (*SI Appendix*). Fig. 1 shows the 603 detected planets and the survey completeness, C, color-coded as a function of P and R_P.

The survey completeness for small planets is a complicated function of P and R_P. It decreases with increasing P and decreasing R_P as expected due to fewer transits and less dimming, respectively. It is dangerous to replace this injection and recovery assessment with noise models to determine C. Such models are not sensitive to the absolute normalization of C and only provide relative completeness. Models also may not capture the complexities of a multistage transit-finding pipeline that is challenged by correlated, nonstationary, and non-Gaussian noise. Measuring the occurrence of small planets with long periods requires injection and recovery of synthetic transits to determine the absolute detectability of the small signals buried in noise.

Planet Occurrence

We define planet occurrence, f, to be the fraction of stars having a planet within a specified range of orbital period, size, and perhaps other criteria. We report planet occurrence as a function of planet size and orbital period, $f(P, R_P)$ and as a function of planet size and the stellar light intensity (flux) incident on the planet, $f(F_P, R_P)$.

Planet Occurrence and Orbital Period. We computed $f(P, R_P)$ in a 6×4 grid of P and R_P shown in Fig. 2. We start by first counting the number of detected planets, n_{cell}, in each P-R_P cell. Then we computed $f(P, R_P)$ by making statistical corrections for planets missed because of nontransiting orbital inclinations and because of the completeness factor, C. The first correction augments each detected transiting planet by $1/P_{\mathrm{T}} = a/R_*$, where P_T is the geometric transit probability, to account for planets missed in inclined orbits. Accounting for the completeness, C, the occurrence in a cell is $f(P, R_P) = 1/n_* \sum_i a_i/(R_{*,i} C_i)$, where $n_* = 42{,}557$ stars, and the sum is over all detected planets within that cell. Uncertainties in the statistical corrections for a/R_* and for completeness may cause errors in the final occurrence rates of ∼10%. Such errors will be smaller than the Poisson uncertainties in the occurrence of Earth-size planets in long period orbits.

Fig. 2 shows the occurrence of planets, $f(P, R_P)$, within the P-R_P plane. Each cell is color-coded to indicate the final planet occurrence: the fraction of stars having a planet with radius and orbital period corresponding to that cell (after correction for both completeness factors). For example, 7.7 ± 1.3% of Sun-like stars have a planet with periods between 25 and 50 d and sizes between 1 and 2 $R_{\oplus}$.

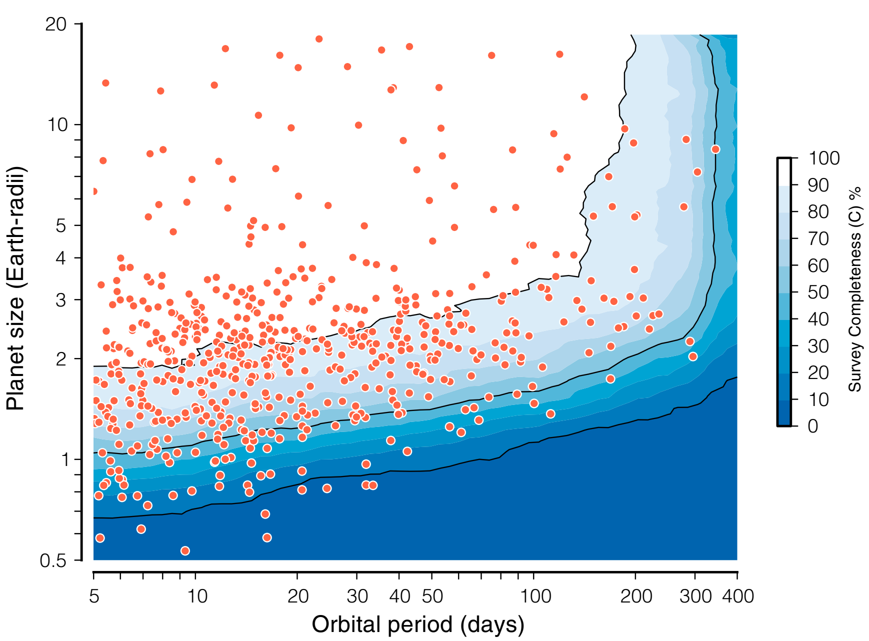

Fig. 1. 2D domain of orbital period and planet size, on a logarithmic scale. Red circles show the 603 detected planets in our survey of 42,557 bright Sun-like stars (Kp = 10–15 mag, GK spectral type). The color scale shows survey completeness measured by injection and recovery of synthetic planets into real photometry. Dark regions represent (P, R_P) with low completeness, C, where significant corrections for missed planets must be made to compute occurrence. The most common planets detected have orbital $P < 20$ d and $R_P \approx 1-3$ $R_{\oplus}$ (at middle left of graph). However, their detectability is favored by orbital tilt and detection completeness, C, that favors detection of such close-in, large planets.

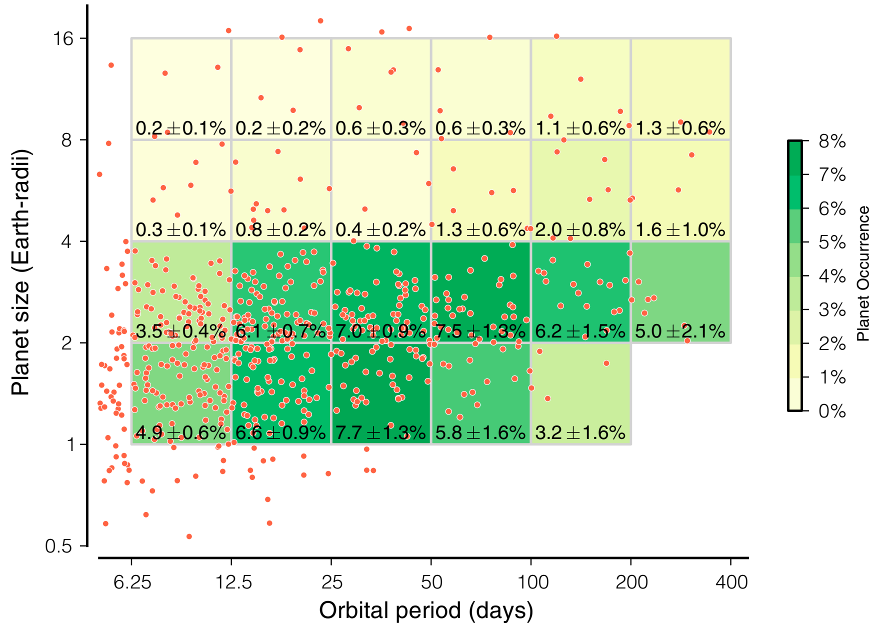

Fig. 2. Planet occurrence, $f(P,R_P)$, as a function of orbital period and planet radius for P = 6.25–400 d and $R_P = 0.5-16$ $R_\oplus$. As in Fig. 1, detected planets are shown as red circles. Each cell spans a factor of 2 in orbital period and planet size. Planet occurrence in a cell is given by $f(P,R_P) = 1/n_\star \sum_i a_i/(R_{\star,i} C_i)$, where the sum is over all detected planets within each cell. Here, a_i/R_i is the number of nontransiting planets (for each detected planet) due to large tilt of the orbital plane, $C_i = C(P_i,R_{P,i})$ is the detection completeness factor, and $n_\star$ = 42,557 stars in the Best42k sample. Cells are colored according to planet occurrence within the cell. We quote planet occurrence within each cell. We do not color cells where the completeness is less than 25%. Among the small planets, 1–2 and 2–4 $R_\oplus$, planet occurrence is constant (within a factor of 2 level) over the entire range of orbital period. This uniformity supports mild extrapolation into the P = 200–400 d, $R_P = 1-2$ $R_\oplus$ domain.

We compute the distribution of planet sizes, including all orbital periods P = 5–100 d, by summing $f(P,R_P)$ over all periods. The resulting planet size distribution is shown in Fig. 3*A*. Planets with orbital periods of 5–100 d have a characteristic shape to their size distribution (Fig. 3*A*). Jupiter-sized planets (11 $R_\oplus$) are rare, but the occurrence of planets rises steadily with decreasing size down to about 2 $R_\oplus$. The distribution is nearly flat (equal numbers of planets per log R_P interval) for 1–2 $R_\oplus$ planets. We find that 26 ± 3% of Sun-like stars harbor an Earth-size planet ($1-2$ $R_\oplus$) with P = 5–100 d, compared with 1.6 ± 0.4% occurrence of Jupiter-size planets ($8-16$ $R_\oplus$).

We also computed the distribution of orbital periods, including all planet sizes, by summing each period interval of $f(P,R_P)$ over all planet radii. As shown in Fig. 3*B*, the occurrence of planets larger than Earth rises from 8.9 ± 0.7% in the P = 6.25–12.5 d domain to 13.7 ± 1.2% in the P = 12.5–25 d interval and is consistent with constant for larger periods. This rise and plateau feature was observed for $\gtrsim 2$ $R_\oplus$ planets in earlier work (5, 12).

Two effects lead to minor corrections to our occurrence estimates. First, some planets in multitransiting systems are missed by TERRA. Second, a small number of eKOIs are false detections. These two effects are small, and they provide corrections to our occurrence statistics with opposite signs. To illustrate their impact, we consider the small and long period (P > 50 d) planets that are the focus of this study.

TERRA detects the highest SNR transiting planet per system, so additional transiting planets that cause lower SNR transits are not included in our occurrence measurement. Using the *Kepler* Project catalog (Exoplanet Archive), we counted the number of planets within the same cells in P and R_P as Fig. 2, noting those that did not yield the highest SNR in the system. Inclusion of

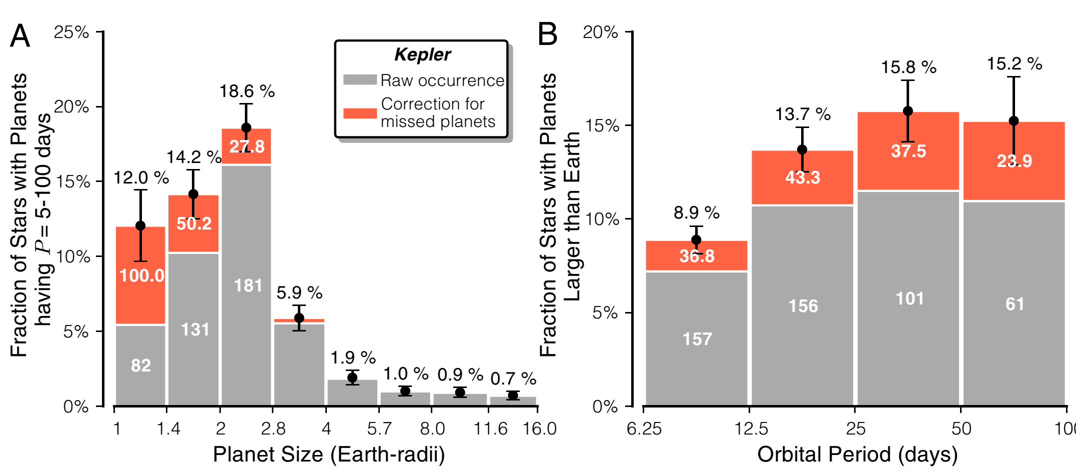

Fig. 3. The measured distributions of planet sizes (*A*) and orbital periods (*B*) for $R_P > 1$ $R_\oplus$ and P = 5–100 d. Heights of the bars represent the fraction of Sun-like stars harboring a planet within a given P or R_P domain. The gray portion of the bars show planet occurrence without correction for survey completeness, i.e., for $C = 1$. The red region shows the correction to account for missed planets, $1/C$. Bars are annotated to reflect the number of planets detected (gray bars) and missed (red bars). The occurrence of planets of different sizes rises by a factor of 10 from Jupiter-size to Earth-sized planets. The occurrence of planets with different orbital periods is constant, within 15%, between 12.5 and 100 d. Due to the small number of detected planets with $R_P = 1-2$ $R_\oplus$ and P > 100 d (four detected planets), we do not include P > 100 d in these marginalized distributions.

these second and third transiting planets boosts the total number of planets per cell (and hence the occurrence) by 21–28% over the P = 50–400 d, R_P = 1–4 $R_\oplus$ domain (*SI Appendix*).

Even with our careful vetting of eKOIs, the light curves of some false-positive scenarios are indistinguishable from planets. Fressin et al. (7) simulated the contamination of a previous KOI (4) sample by false positives that were not removed by the *Kepler* Project vetting process. They determined that the largest source of false positives for Earth-size planets are physically bound stars with a transiting Neptune-size planet, with an overall false-positive rate of 8.8–12.3%. As we have shown (Fig. 2), the occurrence of Neptune-size planets is nearly constant as a function of orbital period, in log$P$ intervals. Thus, this false-positive rate is also nearly constant in period. Therefore, we adopt a 10% false-positive rate for planets having P = 50–400 d and $R_P = 1-2\ R_\oplus$. Planet occurrence, shown in Figs. 2 and 3, has not been adjusted to account for false positives or planet multiplicity. The quoted errors reflect only binomial counting uncertainties. Note that for Earth-size planets in the 50–100 and 100–200 d period bins, planet occurrence is 5.8 ± 1.8% and 3.2 ± 1.6%, respectively. Corrections due to false positives or planet multiplicity are smaller than fractional uncertainties due to small number statistics.

Planet Occurrence and Stellar Light Intensity. The amount of light energy a planet receives from its host star depends on the luminosity of the star (L_*) and the planet-star separation (a). Stellar light flux, F_P, is given by $F_P = L_*/4\pi a^2$. The intensity of sunlight on Earth is $F_\oplus = 1.36$ kW m^{-2}. We compute L_* using $L_* = 4\pi R_*^2 \sigma T_{\rm eff}^4$, where $\sigma = 5.670 \times 10^{-8}$ Wm^{-2}K^{-4} is the Stefan-Boltzmann constant. The dominant uncertainty in F_P is due to R_*. Using spectroscopic stellar parameters, we determine F_P to 25% accuracy and to 80% accuracy using photometric parameters. We obtained spectra for all 62 stars hosting planets with $P >$ 100 d, allowing more accurate light intensity measurements.

Fig. 4 shows the 2D domain of stellar light flux incident on our 603 detected planets, along with planet size. The planets in our sample receive a wide range of flux from their host stars, ranging from 0.5 to 700 $F_\oplus$. We highlight the 10 small ($R_P = 1-2\ R_\oplus$) planets that receive stellar flux comparable to Earth: $F_P = 0.25-4\ F_\oplus$.

Because only two 1–2 $R_\oplus$ planets have $F_P < 1\ F_\oplus$, we measure planet occurrence in the domain, 1–2 $R_\oplus$ and 1–4 $F_\oplus$. Correcting for survey completeness, we find that 11 ± 4% of Sun-like stars have a $R_P = 1-2\ R_\oplus$ planet that receives between one and four times the incident flux as the Earth (*SI Appendix*).

Interpretation

Earth-Size Planets with Year-Long Orbital Periods. Detections of Earth-size planets having orbital periods of P = 200–400 d are expected to be rare in this survey. Low survey completeness ($C \approx 10\%$) and low transit probability (P_T = 0.5%) imply that only a few such planets would be expected, even if they are intrinsically common. Indeed, we did not detect any such planets with TERRA, although the radii of three planets (KIC-4478142, KIC-8644545, and KIC-10593626) have 1σ confidence intervals that extend into the P = 200–400 d, R_P = 1–2 $R_\oplus$ domain. We can place an upper limit on their occurrence: $f < 12\%$ with 95% confidence using binomial statistics. We would have detected one or two such planets if their occurrence was higher than 12%.

However, one may estimate the occurrence of 1–2 $R_\oplus$ planets with periods of 200–400 d by a modest extrapolation of planet occurrence with P. Fig. 5 shows the fraction of stars with 1–2 $R_\oplus$ planets, whose orbital period is less than a maximum period, P, on the horizontal axis. This cumulative period distribution shows that 20.4% of Sun-like stars harbor a 1–2 $R_\oplus$ planet with an orbital period, P < 50 d. Similarly, 26.2% of Sun-like stars harbor a 1–2 $R_\oplus$ planet with a period less than 100 d. The linear increase in cumulative occurrence implies constant planet occurrence per log$P$ interval. Extrapolating the cumulative period distribution predicts $5.7^{+1.7}_{-2.2}\%$ occurrence of Earth-size ($1-2\ R_\oplus$) planets with orbital periods of ~1 y (P = 200–400 d). The details of our extrapolation technique are explained in *SI Appendix*. Extrapolation based on detected planets with P < 200 d predicts that $5.7^{+1.7}_{-2.2}\%$ of Sun-like stars have an Earth-size planet on an Earth-like orbit (P = 200–400 d).

Naturally, such an extrapolation carries less weight than a direct measurement. However, the loss of *Kepler's* second reaction wheel in May 2013 ended observations shortly after the completion of the nominal 3.5-y mission. We cannot count on any additional *Kepler* data to improve the low completeness to Earth

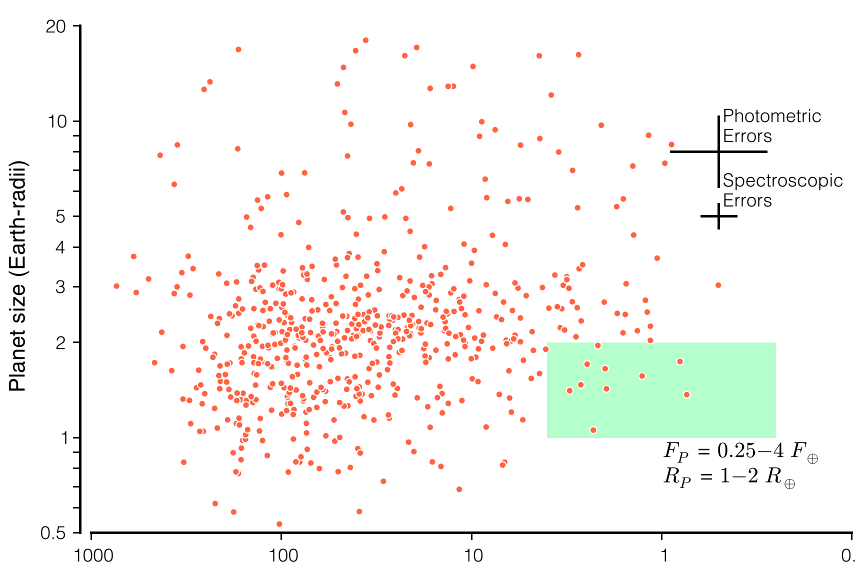

Fig. 4. The detected planets (dots) in a 2D domain similar to Figs. 1 and 2. Here, the 2D domain has orbital period replaced by stellar light intensity, incident flux, hitting the planet. The highlighted region shows the 10 Earth-size planets that receive an incident stellar flux comparable to the Earth: flux = 0.25–4.0 times the flux received by the Earth from the Sun. Our uncertainties on stellar flux and planet radii are indicated at the top right.

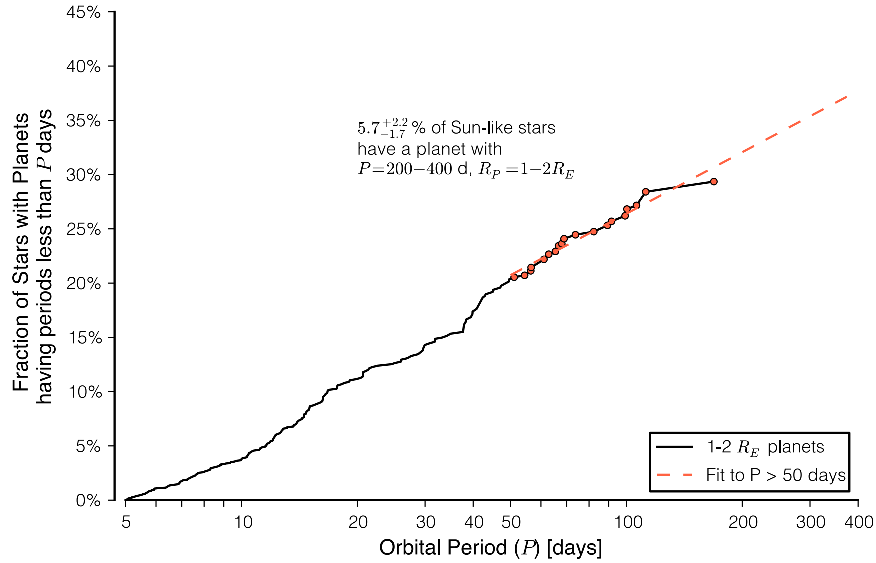

Fig. 5. The fraction of stars having nearly Earth-size planets ($1-2\ R_\oplus$) with any orbital period up to a maximum period, P, on the horizontal axis. Only planets of nearly Earth size ($1-2\ R_\oplus$) are included. This cumulative distribution reaches 20.2% at P = 50 d, meaning 20.4% of Sun-like stars harbor a 1–2 $R_\oplus$ planet with an orbital period, P < 50 d. Similarly, 26.2% of Sun-like stars harbor a 1–2 $R_\oplus$ planet with a period of P < 100 d. The linear increase in this cumulative quantity corresponds to planet occurrence that is constant in equal intervals of log$P$. One may perform a modest extrapolation into the P = 200–400 d range, equivalent to assuming constant occurrence per log$P$ interval, using all planets with P > 50 d. Such an extrapolation predicts that $5.7^{+1.7}_{-2.2}$% of Sun-like stars have a planet with size 1–2 $R_\oplus$, with an orbital period between P = 200 and 400 d.

ASTRONOMY

analog planets beyond what is reported here. Indeed, low survey sensitivity to Earth analogs was the primary reason behind a 4-y extension to the *Kepler* mission. Modest extrapolation is required to understand the prevalence of Earth-size planets with Earth-like orbits around Sun-like stars.

We offer empirical and theoretical justification for extrapolation out to 400 d. As shown in Fig. 2, the prevalence of small planets as a function of log$P$ is remarkably uniform. To test the reliability of our extrapolation into a region of low completeness, we used the same technique to estimate occurrence in more complete regions of phase space and compared the results with our measured occurrence values.

Again, assuming uniform occurrence per log$P$ interval, extrapolating the occurrence of 2–4 $R_\oplus$ planets from 50 to 200 d orbits out to 400 d predicts $6.4^{+0.5}_{-1.2}$% occurrence of planets in the domain of $R_P = 2-4\ R_\oplus$ and P = 200–400 d. The extrapolation is consistent with the measured value of 5.0 ± 2.1% to within 1σ uncertainty. Futhermore, extrapolation based on 1–2 $R_\oplus$ planets with P = 12.5–50 d predicts $6.5^{+0.9}_{-1.7}$% occurrence within the P = 50–100 d bin. Again, the measured value of 5.8 ± 1.6% agrees to better than 1σ.

Although planet size is governed by nonlinear processes such as runaway gas accretion (13), which favors certain planet sizes over others, no such nonlinear processes occur as a function of orbital period in the range of 200–400 d. Extrapolation out to orbital periods of 200–400 d, although dangerous, seems unlikely to be unrealistic by more than a factor of 2.

Earth-Size Planets in the Habitable Zone. Although the details of planetary habitability are debated and depend on planet-specific properties as well as the stochastic nature of planet formation (14), the habitable zone (HZ) is traditionally defined as the set of planetary orbits that permit liquid water on the surface. The precise inner and outer edges of the HZ depend on details of the model (15–18). For solar analog stars, Zsom et al. (17) estimated that the inner edge of the HZ could reside as close as 0.38 astronomical unit (AU) for planets having either a reduced greenhouse effect due to low humidity or a high reflectivity. One AU is the average distance between the Earth and Sun and is equal to 1.50×10^{11} m. Pierrehumbert and Gaidos (18) estimated that the outer edge of the HZ may extend up to 10 AU for planets that are kept warm by efficient greenhouse warming with an H_2 atmosphere.

A planet's ability to retain surface liquid water depends, in large part, on the energy received from its host star. We consider a planet to reside in the HZ if it is bathed in a similar level of starlight as Earth. One may adopt F_P = 0.25–4 $F_\oplus$ as a simple definition of the HZ, which corresponds to orbital separations of 0.5–2.0 AU for solar analog stars. This definition is more conservative than the range of published HZ boundaries that extend from 0.38 to 10 AU (17, 18). This HZ includes Venus (0.7 AU) and Mars (1.5 AU), which do not currently have surface liquid water. However, Venus may have had liquid water in its past, and there is strong geochemical and geomorphological evidence of liquid water earlier in Mars' history (14).

Previously, we showed that 11 ± 4% of stars harbor a planet having an $R_P = 1-2\ R_\oplus$ and $F_P = 1-4\ F_\oplus$. Using the definition of F_P and Kepler's third law, F_P is proportional to $P^{-4/3}$. Therefore, uniform occurrence in log$P$ translates to uniform occurrence in logF_P. We find that the occurrence of 1–2 $R_\oplus$ planets is constant per logF_P interval for $F_P = 100-1\ F_\oplus$. If one was to adopt $F_P = 0.25-4\ F_\oplus$ as the HZ and extrapolate from the $F_P = 1-4\ F_\oplus$

Table 1. Occurrence of small planets in the habitable zone

HZ definition	a_{inner}	a_{outer}	$F_{P,inner}$	$F_{P,outer}$	f_{HZ} (%)
Simple	0.5	2	4	0.25	22
Kasting (1993)	0.95	1.37	1.11	0.53	5.8
Kopparapu et al. (2013)	0.99	1.70	1.02	0.35	8.6
Zsom et al. (2013)	0.38		6.92		26*
Pierrehumbert and Gaidos (2011)		10		0.01	~50†

*Zsom et al. (17) studied the inner edge of the habitable zone. Here we adopt an outer edge of 2 AU from the Simple model.

†Pierrehumbert and Gaidos (18) studied the outer edge of the habitable zone. Here we adopt an inner edge of 0.5 AU from the Simple model. Extrapolation out to 10 AU is severely underconstrained. This estimate is highly uncertain and is included for completeness.

domain, then the occurrence of Earth-size planets in the HZ is $22 \pm 8\%$ for Sun-like stars.

One may adopt alternative definitions of both the properties of Earth-size planets and the domain of the HZ. We showed previously that the occurrence of planets is approximately constant as a function of both R_P (for $R_P < 2.8\ R_\oplus$) and P (in logarithmic intervals). Thus, the occurrence of planets in this domain is proportional to logarithmic area in the R_P–F_P parameter space being considered. For example, the occurrence of planets of size 1.0–1.4 $R_\oplus$ in orbits that receive 0.25–1.0 $F_\oplus$ in stellar flux is 22%/4 = 5.5%. We offer a number of estimates for the prevalence of Earth-size planets in the HZ based on different published definitions of the HZ in Table 1.

Cooler, M dwarf stars also have a high occurrence of Earth-size planets. Based on the *Kepler* planet catalog, Dressing et al. (19) found that $15^{+15}_{-6}\%$ of early M dwarfs have an Earth-size planet $(0.5 - 1.4\ R_\oplus)$ in the HZ using a conservative definition of 0.5–1.1 $F_\oplus$ (15) and three times that value when the HZ is expanded to 0.25–1.5 $F_\oplus$ (20). This result is consistent with a Doppler survey that found that $41^{+54}_{-13}\%$ of nearby M dwarfs have planets with masses 1–10 Earth masses $(M_\oplus)$ in the HZ (21). Thus, Earth-size planets appear to be common in the HZs of a range of stellar types.

Conclusions

Using *Kepler* photometry of Sun-like stars (GK-type), we measured the prevalence of planets having different orbital periods and sizes, down to the size of the Earth and out to orbital periods of 1 y. We gathered Keck spectra of all host stars of planets having periods greater than 100 d to accurately determine their radii. The detection of planets with periods longer than 100 d is challenging, and we characterized our sensitivity to such planets by using injection and recovery of synthetic planets in the photometry. After correcting for orbital tilt and detection completeness, we find that $26 \pm 3\%$ of Sun-like stars have an Earth-size $(1 - 2\ R_\oplus)$ planet with $P = 5$–100 d. We also find that $11 \pm 4\%$ of Sun-like stars harbor an Earth-size planet that receives nearly Earth levels of stellar energy $(F_P = 1 - 4\ F_\oplus)$.

We showed that small planets far outnumber large ones. Only $1.6 \pm 0.4\%$ of Sun-like stars harbor a Jupiter-size $(8 - 16\ R_\oplus)$ planet with $P = 5$–100 d compared with a $23 \pm 3\%$ occurrence of Earth-size planets. This pattern supports the core accretion scenario in which planets form by the accumulation of solids first and gas later in the protoplanetary disk (13, 22–24). The details of this family of models are hotly debated, including the movement of material within the disk, the timescale for planet formation, and the amount of gas accretion in small planets. Our measurement of a constant occurrence of 1–2.8 $R_\oplus$ planets per log$P$ interval establishes an important observational constraint for these models.

The occurrence of Earth-size planets is constant with decreasing stellar light intensity from 100 $F_\oplus$ down to 1 $F_\oplus$. If one was to assume that this pattern continues down to 0.25 $F_\oplus$, then the occurrence of planets having flux levels of 1–0.25 $F_\oplus$ is also $11 \pm 4\%$.

Earth-size planets are common in the *Kepler* field. If the stars in the *Kepler* field are representative of stars in the solar neighborhood, then Earth-size planets are common around nearby Sun-like stars. If one were to adopt a 22% occurrence rate of Earth-size planets in habitable zones of Sun-like stars, then the nearest such planet is expected to orbit a star that is less than 12 light-years from Earth and can be seen by the unaided eye. Future instrumentation to image and take spectra of these Earths need only observe a few dozen nearby stars to detect a sample of Earth-size planets residing in the HZs of their host stars.

ACKNOWLEDGMENTS. We acknowledge extraordinary help from Howard Isaacson, John Johnson, David Ciardi, Steve Howell, Natalie Batalha, Jon Jenkins, William Borucki, Francois Fressin, David Charbonneau, and G. Willie Torres. We extend special thanks to those of Hawaiian ancestry on whose sacred mountain of Mauna Kea we are privileged to be guests. We thank NASA Exoplanet Science Institute (NExScI) and the University of California Observatories at University of California–Santa Cruz for their administration of the Keck Observatory. *Kepler* was competitively selected as the 10th Discovery mission, with funding provided by National Aeronautics and Space Administration's (NASA's) Science Mission Directorate. E.A.P. gratefully acknowledges support from a National Science Foundation graduate fellowship. A.W.H. and G.W.M. gratefully acknowledge funding from NASA Grants NNX12AJ23G and NNX13AJ59G, respectively. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. The Keck Observatory was made possible by the generous financial support of the W. M. Keck Foundation. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy under Contract DE-AC02-05CH11231.

1. Borucki WJ, et al. (2010) Kepler planet-detection mission: Introduction and first results. *Science* 327(5968):977–980.
2. Koch DG, et al. (2010) Kepler mission design, realized photometric performance, and early science. *Astrophys J Lett* 713(2):L79–L86.
3. Borucki WJ, et al. (2011) Characteristics of planetary candidates observed by Kepler. II. Analysis of the first four months of data. *Astrophys J* 736(1):19–40.
4. Batalha NM, et al. (2013) Planetary candidates observed by Kepler. III. Analysis of the first 16 months of data. *Astrophys J* 204(2):24–44.
5. Howard AW, et al. (2012) Planet occurrence within 0.25 AU of solar-type stars from Kepler. *Astrophys J* 201(2):15–39.
6. Petigura EA, Marcy GW, Howard AW (2013) A plateau in the planet population below twice the size of Earth. *Astrophys J* 770(1):69–89.
7. Fressin F, et al. (2013) The false positive rate of Kepler and the occurrence of planets. *Astrophys J* 766(2):81–100.
8. Pinsonneault MH, et al. (2012) A revised effective temperature scale for the Kepler input catalog. *Astrophys J* 199(2):30–51.
9. Petigura EA, Marcy GW (2012) Identification and removal of noise modes in Kepler photometry. *Publ Astron Soc Pac* 124(920):1073–1082.
10. Mandel K, Agol E (2002) Analytic light curves for planetary transit searches. *Astrophys J Lett* 580(2):L171–L175.
11. Brown TM, Latham DW, Everett ME, Esquerdo GA (2011) Kepler input catalog: Photometric calibration and stellar classification. *Astron J* 142(4):112–129.
12. Youdin AN (2011) The exoplanet census: A general method applied to Kepler. *Astrophys J* 742(1):38–50.
13. Ida S, Lin DNC, Nagasawa M (2013) Toward a deterministic model of planetary formation. VII. Eccentricity distribution of gas giants. *Astrophys J* 775(1):42–64.
14. Seager S (2013) Exoplanet habitability. *Science* 340(6132):577–581.
15. Kasting JF, Whitmire DP, Reynolds RT (1993) Habitable zones around main sequence stars. *Icarus* 101(1):108–128.
16. Kopparapu RK, et al. (2013) Habitable zones around main-sequence stars: New estimates. *Astrophys J* 765(2):131–146.
17. Zsom A, Seager S, de Wit J, Stamenkovic V Towards the minimum inner edge distance of the habitable zone *Astrophys J.,* in press.
18. Pierrehumbert R, Gaidos E (2011) Hydrogen greenhouse planets beyond the habitable zone. *Astrophys J Lett* 734(1):L13–L17.
19. Dressing CD, Charbonneau D (2013) The occurrence rate of small planets around small stars. *Astrophys J* 767(1):95–114.
20. Kopparapu RK (2013) A revised estimate of the occurrence rate of terrestrial planets in the habitable zones around Kepler M-dwarfs. *Astrophys J Lett* 767(1):L8–L12.
21. Bonfils X, et al. (2013) The HARPS search for southern extra-solar planets. XXXI. The M-dwarf sample. *Astron Astrophys* 549:A109.
22. Levy EH, Lunine JI, eds (1993) *Growth of Planets from Planetesimals* (University of Arizona Press, Tucson, AZ), pp 1061–1088.
23. Pollack JB, et al. (1996) Formation of the giant planets by concurrent accretion of solids and gas. *Icarus* 124(1):62–85.
24. Mordasini C, et al. (2012) Characterization of exoplanets from their formation. II. The planetary mass-radius relationship. *Astron Astrophys* 547:A112.