

# Exploring exoplanet populations with NASA's Kepler Mission

Natalie M. Batalha<sup>1</sup>

National Aeronautics and Space Administration Ames Research Center, Moffett Field, 94035 CA

Edited by Adam S. Burrows, Princeton University, Princeton, NJ, and accepted by the Editorial Board June 3, 2014 (received for review January 15, 2014)

**The Kepler Mission is exploring the diversity of planets and planetary systems. Its legacy will be a catalog of discoveries sufficient for computing planet occurrence rates as a function of size, orbital period, star type, and insolation flux. The mission has made significant progress toward achieving that goal. Over 3,500 transiting exoplanets have been identified from the analysis of the first 3 y of data, 100 planets of which are in the habitable zone. The catalog has a high reliability rate (85–90% averaged over the period/radius plane), which is improving as follow-up observations continue. Dynamical (e.g., velocimetry and transit timing) and statistical methods have confirmed and characterized hundreds of planets over a large range of sizes and compositions for both single- and multiple-star systems. Population studies suggest that planets abound in our galaxy and that small planets are particularly frequent. Here, I report on the progress *Kepler* has made measuring the prevalence of exoplanets orbiting within one astronomical unit of their host stars in support of the National Aeronautics and Space Administration's long-term goal of finding habitable environments beyond the solar system.**

planet detection | transit photometry

Searching for evidence of life beyond Earth is one of the primary goals of science agencies in the United States and abroad. The goal looms closer as a result of exoplanet discoveries made by the National Aeronautics and Space Administration's (NASA) 10th Discovery Mission, *Kepler*. Launched in March 2009, the *Kepler* spacecraft is exploring the diversity of planets and planetary systems within one astronomical unit (AU). The primary mission objective is to determine the prevalence of potentially habitable, Earth-size planets in the galaxy. Discovering exoterans in the habitable zone (HZ), characterizing those that have habitable environments, and then focusing on the signatures of biological chemistry is a path of exploration that stretches decades into the future. It begins by determining if planets like Earth are abundant.

## NASA's 10th Discovery Mission

From 2009 to 2013, *Kepler* monitored a 115-square-degree field in the constellations Cygnus and Lyra, collecting ultrahigh precision photometry of over 190,000 stars simultaneously at a 30-min cadence. Nearly uninterrupted photometry is possible due to a heliocentric orbit and off-ecliptic pointing. The observations yield an evenly sampled, minimally gapped flux time series that can be searched for periodic diminutions of light due to the transit of an exoplanet across the stellar disk in an aligned geometry. The photometer was engineered to achieve 20-ppm relative precision in 6.5 h for a 12th magnitude G-type main-sequence star (1). For reference, the Earth orbiting

the Sun would produce an 84-ppm signal lasting ~13 h.

*Kepler's* pixel and flux measurements (2) are publicly available at the Mikulski Archive for Space Telescopes (MAST) (<http://archive.stsci.edu/kepler>). Transit searches have been performed on successively larger data volumes yielding incremental planet candidate catalogs that are hosted at NASA's Exoplanet Archive (NEA) (<http://exoplanetarchive.ipac.caltech.edu>). To date, approximately three-quarters of the data have been thoroughly searched. As of this writing (April 2014), the archive is host to over 3,500 viable planet candidates (with radii smaller than twice Jupiter). All have been subjected to a series of statistical tests (based on the *Kepler* data itself) that ensure a low rate of instrumental and astrophysical false positives (3).

*Kepler* has a follow-up observation program to increase the reliability of the catalog even further by (i) improving the accuracy of the host star properties which in turn improves the accuracy of the planet properties (or changes the interpretation altogether) and (ii) identifying bound stellar companions and line-of-sight neighbors that might indicate an astrophysical false positive. Ground- and space-based telescopes with apertures ranging from 1.5 to 10 m are being used to acquire high-resolution spectroscopy and high-contrast/high-spatial resolution images. Strategic high-precision Doppler measurements are providing planet masses in an effort to delineate the transition between terrestrial and giant planets.

Translating *Kepler's* discovery catalog into population statistics requires corrections for observation and detection biases. This is a work in progress. However, occurrence rate calculations based on subsets of the data already indicate that nature produces small planets relatively efficiently in the warmer environs of a planetary system. Giant planets in such orbits are orders of magnitude less frequent than their sub-Neptunian counterparts. Ironically, the hot Jupiters that comprised the very first Doppler and transiting exoplanet discoveries are actually quite rare. Current results for habitable-zone planets tell us that we may not have to look very far before happening upon a planet similar to Earth.

A comprehensive review of *Kepler* exoplanet science is beyond the scope of this contribution. Here, I focus on the science leading to the determination of planet occurrence rates, from the discovery catalogs to the first calculations of the prevalence of Earth-like planets.

## *Kepler* Transforms the Discovery Space

Exoplanet discoveries trickled in at a steady rate in the latter half of the 1990s. Approximately 30 were reported with sizes ranging from 0.4 to 8 Jupiter masses and orbital periods ranging from 3 to 3,800 d. Heralding in the new millennium, the first transiting

Author contributions: N.M.B. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission. A.S.B. is a guest editor invited by the Editorial Board.

<sup>1</sup>Email: Natalie.M.Batalha@nasa.gov.

exoplanet was discovered (4, 5). The timing was a boon for *Kepler* as it was proposing to use this detection technique from space. In 2000, *Kepler* was one of the three Discovery Mission proposals invited to submit a Concept Study Report. It was selected for flight on December 20, 2001.

As *Kepler* was being designed and built, exoplanet discoveries were growing at an accelerated pace. By the eve of *Kepler*'s launch, over 300 discoveries had been reported including nearly 70 transiting systems. All non-*Kepler* discoveries up through April 2014 are shown in Fig. 1, *Left*, in a plot of mass (or minimum mass for nontransiting planets) versus orbital period with symbols color-coded by the discovery method. (Methodologies with small numbers of discoveries have been left out for clarity). Collectively, there are 697 (non-*Kepler*) exoplanets (with a measured orbital period and radius or mass) associated with 583 unique stars. Approximately 16% of these host stars are known to harbor multiple planets.

Fig. 1, *Right*, shows the same population together with the *Kepler* planet candidate discoveries in the cumulative table at NEA as of April 2014. Detections are plotted as planet radius versus orbital period, and the non-*Kepler* discoveries are included for comparison. Where planet radii are not available (as is the case for most of the Doppler

detections), they are estimated using a polynomial fit to solar system planets ( $R = M^{0.4854}$ ) (6). Shown here are 3,553 *Kepler* discoveries associated with 2,658 stars. Approximately 22% of the *Kepler* host stars are known to harbor multiple planet candidates. The overall reliability of the catalog (80–90%) is discussed below.

The demographics of the observed population has changed remarkably. *Kepler* has increased the roster of exoplanets by nearly 400%. More remarkable still is the change in the distribution: 86% of the non-*Kepler* discoveries have masses larger than Neptune whereas 85% of the *Kepler* discoveries have radii smaller than Neptune. *Kepler* is filling in an area of parameters space that was not previously accessible. The increase in sensitivity afforded us by *Kepler* has opened the floodgates to the small planets so difficult to detect from ground-based surveys. The most common type of planet known to us is a population that does not exist in our own solar system: the super-Earths and mini-Neptunes between 1 and 4 Earth radii.

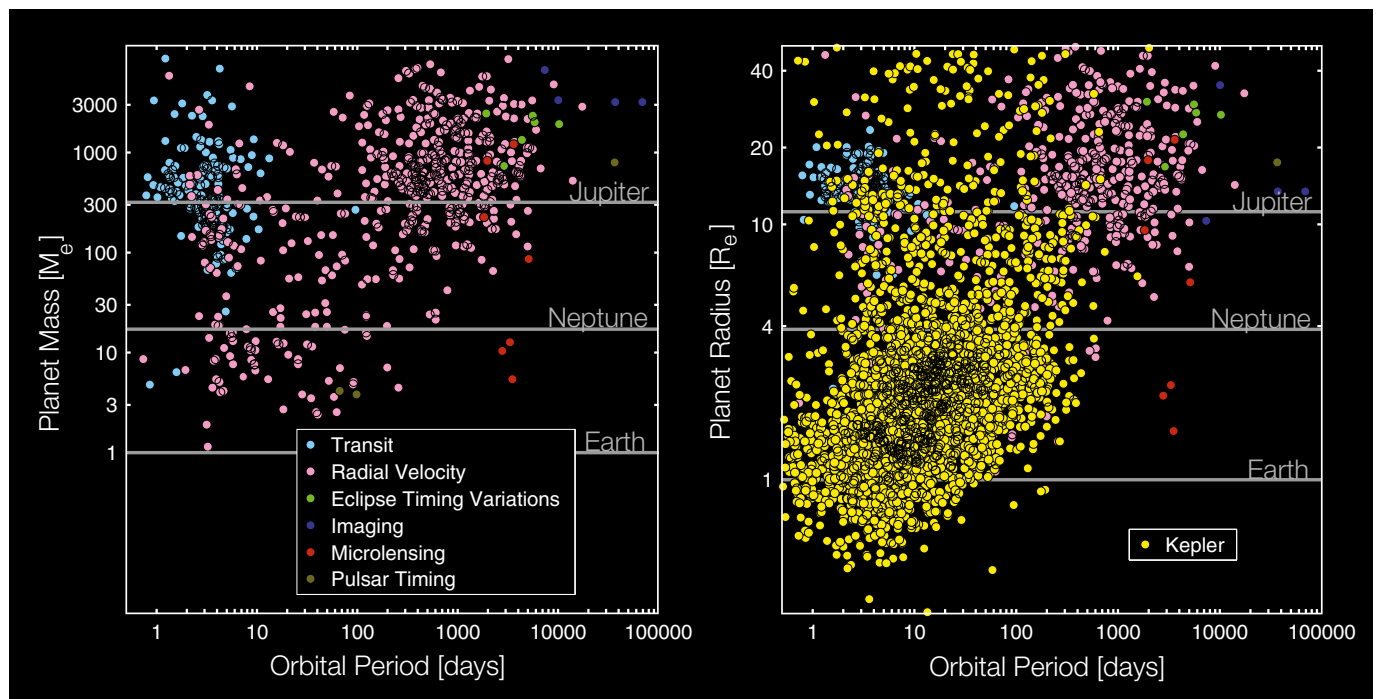
### Status of *Kepler*'s Discovery Catalogs

Catalogs of *Kepler*'s viable planet candidates have been released periodically since launch and have included 312, 1,235, 2,338, 2,738, and 3,553 detections (cumulative counts) associated with 306, 997, 1,797, 2,017, and

2,658 stars based on 1.5, 13, 16, 22, and 34.5 of the ~48 mo of data acquired during the primary mission (7–11). *Kepler* data in the prime mission were downlinked monthly but processed on a quarterly basis. Transit searches and the associated planet candidate catalogs are, therefore, referred to by the quarters bracketing the data. The most recent planet candidates were identified in a search of 12 quarters of data (Q1–Q12) where the first is only slightly longer than one month in duration (hence the 34.5-mo time span).

Previously detected candidates are reexamined as larger data volumes become available. However, this does not occur with every catalog release. Some of the candidates in the cumulative archive at the NEA were discovered with less than 34.5 mo of data and have not yet been reexamined. This nonuniformity will be resolved as *Kepler* completes its final search and vetting of the entire 17 quarters (48 mo) of data acquired during its primary mission lifetime. *Kepler*'s planet candidate catalog is also known as the Kepler Object of Interest (KOI) Catalog. However, KOIs also include events that are classified as false alarms or astrophysical false positives. Only those flagged as planet candidates in the NEA cumulative catalog are shown in Fig. 1.

The catalogs contain the five parameters produced by fitting a limb-darkened Mandel and Agol (12) model to the observed flux



**Fig. 1.** Non-*Kepler* exoplanet discoveries (*Left*) are plotted as mass versus orbital period, colored according to the detection technique. A simplified mass–radius relation is used to transform planetary mass to radius (*Right*), and the >3,500 *Kepler* discoveries (yellow) are added for comparison. Eighty-six percent of the non-*Kepler* discoveries are larger than Neptune, whereas the inverse is true of the *Kepler* discoveries: 85% are smaller than Neptune.

time series assuming zero eccentricity: the transit ephemeris (period and epoch), reduced radius ( $R_p/R_*$ ), reduced semimajor axis ( $d/R_*$ ), and impact parameter. To first order, the reduced semimajor is equivalent to the ratio of the planet–star separation during transit to the stellar radius. Despite its name, it is equivalent to  $a/R_*$  (where  $a$  is the semimajor axis) only in the case of a zero eccentricity orbit.

Planet properties are also tabulated in the discovery catalogs. Planet radius, semimajor axis, and insolation flux are computed from light curve parameters and knowledge of the host star properties (effective temperature, surface gravity, mass, and radius). The Kepler Input Catalog (KIC) (13) contains the properties of stars in the *Kepler* field of view derived from ground-based broad- and narrow-band photometry acquired before launch to support target selection. However, the KIC contains known deficiencies and systematic errors, making it unsuitable for computing accurate planet properties (14–18).

A Kepler working group provides incremental deliveries of updated properties of all stars observed by *Kepler* with the long-term goal of increasing accuracy and quantifying systematics. Accuracy is required for characterizing individual planetary systems. Also,

an understanding of planetary populations via occurrence rate studies requires a homogeneous database of the properties of all observed stars. Toward this aim, the working group coordinates campaigns and collates atmospheric properties (temperature, surface gravity, and metallicity) derived from different observational techniques (photometry, spectroscopy, and asteroseismology), which are then fit to a grid of stellar isochrones to determine fundamental properties like mass and radius.

The planet radii plotted in Fig. 1 (and Fig. 2) are not taken directly from the NEA cumulative table. Rather, the planet radii (and ancillary properties like insolation flux) are recomputed using the modeled light curve parameters and the Q1–Q16 catalog of star properties (also available at the NEA), so called because it is used as input to the Q1–Q16 pipeline run. The provenance of all values in the Q1–Q16 star properties catalog are described by Huber et al. (19) as is the strategy for future updates to the catalog. Published properties of confirmed planets are used where available.

Looking forward, there is 1 y of data left to analyze. The Q1–Q16 pipeline run searched for statistically significant, transit-like signals, also called threshold-crossing events (TCEs).

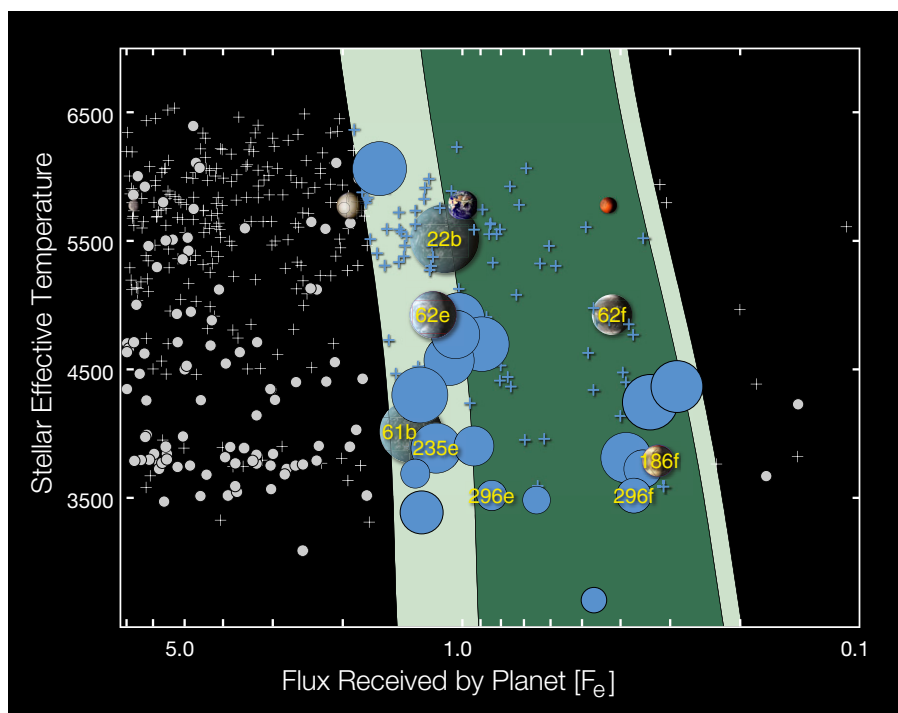
Over 16,000 events were identified. The Q1–Q16 TCE list is archived at the NEA and described in ref. 20. The list contains previously discovered planets, false positives, and eclipsing binaries as well as numerous false alarms. Dispositioning will occur after a vetting process using the validation tests described in ref. 3.

Efforts to produce an updated catalog of planet candidates are underway and should be completed in mid-2014. Hundreds of new discoveries are expected, including the first small planet candidates in the HZ of G-type stars. Moreover, Kepler data are in the public domain thereby enabling many additional discoveries. Both the scientific community (21, 22) and citizen science efforts (23, 24) have yielded new candidates and confirmed planets. Interesting new niches of parameter space have been opened up thanks to such efforts. Notables include the first seven-planet system KOI-351 (25), a planet in a quadruple-star system (26), and objects in ultrashort orbits (27).

### Planets in the HZ

*Kepler's* objective is to determine the frequency of Earth-size planets in the HZ of Sun-like stars. Defined as the region where a rocky planet can maintain surface liquid water, the HZ is a useful starting point for identifying exoplanets that may have an atmospheric chemistry affected by carbon-based life (28). As we broaden our perspective, we stretch and prod the HZ limits. Abe et al. (29) and Zsom et al. (30) consider the extreme case of arid Dune-like planets. LeConte et al. (31) and Yang et al. (32) consider the effects of rotation. And Lissauer (33) considers the dessication of planetary bodies before their M-type host stars settle onto the main sequence. There may not be a simple evolutionary pathway that lands an exoplanet inside of a well-defined HZ. Regardless, it is of interest to understand the prevalence of planets with properties similar to Earth. For *Kepler's* exoplanets, comparisons with Earth are made considering size (radius) and orbital environment (period or semimajor axis), both of which require knowledge of the host star properties. The orbital environment can also be characterized by the irradiation, or insolation flux, defined as  $F = (R_*/R_\odot)^2 (T_*/T_\odot)^4 (a_\oplus/a_p)^2$ . The insolation flux of each planet candidate is shown in Fig. 2, where the y axis is the effective temperature of the host star.

Two definitions of the HZ are included for reference in Fig. 2, both of which are taken from (34). The wider HZ (light green in Fig. 2) is based on the recent Venus and early Mars limits discussed therein and is referred to as



**Fig. 2.** Stellar effective temperature versus insolation (stellar flux at the semimajor axis) for Kepler exoplanets larger than  $2 R_\oplus$  (pluses) and smaller than  $2 R_\oplus$  (circles). Symbols are colored blue if they lie within the HZ and are sized relative to the Earth (represented by a superimposed image) if they represent a planet smaller than  $2 R_\oplus$ . The confirmed HZ exoplanets (Kepler-22b, Kepler-62 e and f, Kepler-61b, and Kepler-186f) are displayed as the artist's conceptions.

the “optimistic” HZ. The optimistic HZ does not extend all of the way in to the venusian orbit. The Sun was  $\sim 92\%$  less luminous 1 Ga at the epoch when Venus may have had liquid water on its surface. The insolation intercepted by Venus during that epoch corresponds to the insolation at 0.75 AU in the present-day Solar System ( $1.78 F_{\oplus}$ ). Similarly, the outer edge of the optimistic HZ extends beyond the martian orbit since the Sun was  $\sim 75\%$  as luminous 3.8 Ga when Mars was thought to have liquid water. The insolation intercepted by Mars at that epoch corresponds to the insolation at 1.77 AU in the present day solar system ( $0.32 F_{\oplus}$ ).

The narrow HZ (dark green in Fig. 2) is defined via climate models assuming an Earth-mass planet with different  $\text{CO}_2$  and  $\text{H}_2\text{O}$  compositions that take the planet to the two extremes. These are the runaway greenhouse and maximum greenhouse limits (34) and are referred to as the “conservative” HZ. According to these models, the highest flux a planet can receive while maintaining surface temperatures amenable to liquid water occurs for a water-saturated atmosphere. The inner edge at  $1.02 F_{\oplus}$  corresponds to rapid water loss and hydrogen dissipation in a water-saturated atmosphere. The outer edge at  $0.35 F_{\oplus}$  corresponds to the maximum possible greenhouse warming from a  $\text{CO}_2$ -dominated atmosphere. Beyond the outer edge of this conservative HZ, models indicate that  $\text{CO}_2$  begins to condense and lose its warming greenhouse properties.

The inner solar system planets line up horizontally in Fig. 2, with Mercury at the extreme left, Venus and Mars bracketing the optimistic HZ, and the Earth near the inner edge of the conservative HZ. The HZ fluxes at the inner and outer edges have a slight dependence on the properties of the host star (note that the green shaded regions in Fig. 2 are not vertical bars). The amount of radiation absorbed/reflected by the planet is wavelength dependent. Therefore, the Bond albedo depends on the spectral energy distribution of the host star, and the limits are adjusted accordingly.

From the first 3 y of data (Q1–Q12), there are over 100 candidates that have an insolation flux that falls within the optimistic HZ. Of those, 21 are smaller than  $2 R_{\oplus}$ . These are shown as circles in Fig. 2. The symbols are sized in proportion to the Earth image to reflect their relative radii. Five of the Kepler HZ discoveries are planets that have been statistically validated at the 99% confidence level or higher: Kepler-22b (35), Kepler-61b (36), Kepler-62 e and f (37), and Kepler-186f (38), with radii of  $2.38 \pm 0.13$ ,  $2.15 \pm 0.13$ ,  $1.61 \pm 0.05$ ,  $1.41 \pm 0.07$ , and  $1.11 \pm 0.14 R_{\oplus}$ , respectively. These are represented

by the artist’s conceptions, also scaled in size with respect to the Earth. Kepler-235e and Kepler-296 e and f are verified planets (39) with uncertain properties. Disparate star properties have been reported in the literature for Kepler-235. The planet properties shown in Fig. 2 are derived assuming a  $0.48 R_{\odot}$  host star (19). Kepler-296 is a diluted (multiple star) system (39). The properties of 296 e and f shown here are derived assuming the planets orbit the primary star.

Kepler’s small HZ candidates orbit predominantly K- and M-type main-sequence stars—perhaps not surprising given the fact that only 34.5 mo of data were used to produce the sample of planet candidates shown in Figs. 1 and 2. HZ planets associated with G-type main-sequence stars produce shallower transits, have longer orbital periods, and therefore require more data for detection compared with those transiting cooler stars of comparable magnitude.

Kepler was designed to achieve a 6.5-h precision of 20 ppm or better for a 12th magnitude Sun-like star. The 20 ppm total allows for the detection of a  $1.0 R_{\oplus}$  planet with four transits, which in turn allows for the detection of Earth analogs in a 4-y mission. The baseline noise budget for a G2-type main-sequence star included a 10-ppm contribution for intrinsic stellar variability consistent with observations of the Sun (40). However, the realized noise for 12th magnitude Sun-like stars has a mode at 30 ppm due to a combination of unanticipated stellar variability and instrument noise (41). Both were a factor of 2 larger than expected and, when added to the shot noise, resulted in a total noise budget that was 50% larger than anticipated.

An extended mission was awarded, but the loss of two of Kepler’s reaction wheels degraded the pointing stability in the nominal field of view. Very high pointing stability is required to achieve the photometric precision necessary to detect small HZ planets. Consequently, the nominal mission ended with the loss of the second reaction wheel in May 2013. Detection of sizable numbers of small HZ planets may require software solutions to reduce other noise contributions. Numerous improvements to pipeline modules have been implemented, and a full reprocessing of the data is underway.

Considering the possibility of fewer detections than originally anticipated, it is critical to carefully quantify the reliability of the detections in hand.

### Catalog Reliability

The Kepler discoveries are referred to as planet candidates until they are either dynamically

confirmed or statistically validated (see below). The former deals with follow-up observations and/or analyses that seek to identify dynamical evidence of an exoplanet (e.g., radial velocity or transit timing variations), whereas the latter deals with follow-up observations that seek to rule out scenarios produced by astrophysical signals that can mimic a planetary transit. Potential sources of astrophysical false positives include

- i) grazing eclipse of binary stars;
- ii) eclipse of a giant star by a main-sequence star;
- iii) eclipse of an FGK-type main-sequence star by a very late-type star or brown dwarf;
- iv) eclipse of a fore- or background binary near the target as projected on the sky;
- v) eclipse of a binary physically associated with the target;
- vi) transiting planet orbiting a nearby (projected onto the sky) fore- or background star;
- vii) transiting planet orbiting a physical companion of the target star; and
- viii) long-period, eccentric companion (star or giant planet) that yields only the secondary eclipse (or occultation).

Kepler’s target stars are relatively well characterized making it unlikely that an exoplanet transit will be confused by a main-sequence star eclipsing a giant. Moreover, Kepler’s ultrahigh precision photometry allows for statistical tests that eliminate many of the false-positive scenarios that plague ground-based surveys. For example, Kepler readily detects secondary eclipses of grazing and high-mass ratio eclipsing binaries. Moreover, part-per-million differences between the eclipse depths of two nearly equal-mass stars are often discernible. The statistical tests performed on the data to identify these tell-tale signs are described in ref. 3.

By design, Kepler’s pointing stability is better than 0.003 arcseconds (arcsec) on 15-min timescales (1). This allows us to measure relative star positions to millipixel precision (42). The center of light distribution (photocenter) for a photometric aperture can be computed at each cadence producing a time series of row and column photocenter values with submillipixel precision on transit timescales (43). These time series contain information about the location of the source of the transit or eclipse event. However, dilution from multiple flux sources (known and unknown) in the aperture makes the interpretation difficult in some cases. Alternatively, in-transit and out-of-transit pixel images can be used to construct difference images that provide direct information about the location of the transit (or eclipse) source (42). Difference image analysis eliminates

a large fraction of the false-positive scenarios involving dilution from nearby targets.

Follow-up observations further restrict the false-positive parameter space. *Kepler* has made it a priority to collect high-resolution, high-signal-to-noise spectra and high-contrast, high-spatial resolution imaging of as many of the planet–host stars as possible. Difference image analysis rules out the presence of diluting stars outside of a spatial radius (typically about 2 arcsec, or half a pixel). Adaptive optics or speckle imaging can tighten that radius to a fraction of an arc-second, thereby significantly reducing the parameter space where false positives can lurk. Bound stellar-mass companions with subarcsecond separation and flux greater than 1% of the primary can be ruled out by spectroscopy (37).

Numerical simulations provide an estimate of the likelihood of remaining astrophysical false-positive scenarios given the density of stars as a function of magnitude and galactic coordinates as well as the frequency of eclipsing binaries and transiting planets. Morton and Johnson (44) computed the false-positive probability (FPP) for each of the 1,235 planet candidates reported in ref. 8 and find the FPPs to be less than 10% for nearly all candidates. Empirical estimates are a mixed bag. Santerne et al. (45) performed radial velocity follow-up of 46 close-in giant planet candidates and estimated a 34.8% false-positive rate, whereas Désert et al. (46) acquired Spitzer observations of 51 candidates (of primarily subneptunian sizes) and identified only one false positive.

Fressin et al. (47) simulated the global population of astrophysical false positives that would be detectable in the observations of all target stars and would persist even after the careful vetting described above. Two interesting results emerged. Somewhat counterintuitively, the highest false-positive rates (~18%) were found for the close-in giant planets which were qualitatively consistent with the empirical results of ref. 45. Secondly, the most common source of false positives mimicking small planets was a larger planet transiting an unseen physical companion or a background star. Such scenarios were not considered in the Morton and Johnson analysis. Fressin et al. reported a  $9.4 \pm 0.9\%$  global false-positive rate for the Q1–Q6 catalog (9). This value was revised upward (48) to  $11.3 \pm 1.1\%$  upon inclusion of secondary-only false positives.

Even if only 80–90% of the detections are bona fide planets, *Kepler* has quadrupled the number of exoplanets, providing a statistically

significant and diverse population for studying demographics.

### Planet Confirmation and Characterization

The confirmation and characterization of *Kepler*'s exoplanet candidates contribute to planet population studies by increasing the reliability of the planet census and by offering an empirical ground truth to estimates of FPPs as previously discussed. Just as important, however, is the information emerging about the distribution of planet densities. With this information, we can estimate not only the occurrence rate of Earth-size planets in the HZ, but also the occurrence rate of veritably rocky planets in the HZ.

As of this writing, over 962 *Kepler* exoplanet candidates have been either dynamically confirmed or statistically validated. High-precision radial velocity follow-up has yielded ~50 mass determinations from instruments scattered across the northern hemisphere, including the SOPHIE (Spectrographe pour l'Observation des Phénomènes des Intérieurs stellaires et des Exoplanètes) spectrograph at the Observatoire de Haute-Provence (49–51), FIES (fiber-fed echelle spectrograph) on the Nordic Optical Telescope (52), HRS (high-resolution spectrograph) on the Hobby–Eberly Telescope (53), HARPS-N (High Accuracy Radial velocity Planet Searcher–North) on the Telescopio Nazionale Galileo (54), and the HIRES (high-resolution Echelle spectrometer) spectrograph on Keck (55). Of special interest are the measurements for the sub-Neptune-size planets, particularly those that have densities indicative of a rocky composition: Kepler-10b (56) and Kepler-78b (57, 58). A recent report on 4 y of strategic Keck observations (59) has added another six candidate rocky planets to this roster.

Dynamical confirmation is not limited to velocimetry measurements. Approximately half of *Kepler*'s confirmations come from measurement of transit timing variations (60–68). Anticorrelated timing variations exhibited by two planets in a system can place an upper limit on mass thereby supporting the planet interpretation. In some cases, dynamical models of transit timing variations resulting from mutual planetary perturbations yield mass measurements. Such measurements have been obtained for sub-Neptune-sized objects including five planets orbiting Kepler-11 (69, 70), Kepler-20 b and c (71), Kepler-30b (72), Kepler-18b (73), Kepler-87c (74), Kepler-79 b and c (75), Kepler-36c, and its rocky neighbor Kepler-36b (76).

Collectively, data on subneptunian planets do not support a strict relation between mass and radius. A power-law fit of mass versus radius for 63 exoplanets smaller than  $4 R_{\oplus}$  has a reduced  $\chi^2$  of 3.5 (77). The large dispersion is indicative of a compositional diversity arising from the varied formation, migration, interaction, and irradiation pathways of planetary evolution. Kepler-11d and Kepler-100b exemplify this diversity, having similar masses ( $7.3 \pm 1.2$  and  $7.3 \pm 3.2 M_{\oplus}$ ) but quite different radii ( $3.12 \pm 0.07$  and  $1.32 \pm 0.04 R_{\oplus}$ ). Kepler-11d most likely contains a high H/He and/or ice envelope fraction ( $\rho = 1.28 \pm 0.20$  g/cm<sup>3</sup>), whereas Kepler-100b is consistent with an Earth-like composition ( $\rho = 14.25 \pm 6.33$  g/cm<sup>3</sup>).

Theoretical models of sub-Neptune-sized planets suggest that planetary radius changes very little with increasing mass for a given compositional mix (78). The authors suggest that planetary radius is, to first order, a proxy for planetary composition. However, the observational data serve as a caution. Kepler-11b and Kepler-113b have nearly equal radii ( $1.80 \pm 0.04$  and  $1.82 \pm 0.05 R_{\oplus}$ ) yet different masses ( $1.9 \pm 1.2$  and  $11.7 \pm 4.2 M_{\oplus}$ ) and densities ( $1.72 \pm 1.08$  and  $10.73 \pm 3.9$  g/cm<sup>3</sup>). This occurs as well for planets in the same system. Kepler-138 c and d, for example, have the same radius ( $1.61 \pm 0.16 R_{\oplus}$ ) but different masses ( $1.01^{+0.42}_{-0.34}$  and  $3.83^{+1.51}_{-1.26} M_{\oplus}$ , respectively) (79).

The fraction of planets of a given composition is likely to be a smooth function of planet size, implying no particular radius that marks a clean transition from rocky planets to those with H/He and/or ice envelopes. There are hints, however, that most planets smaller than  $1.5 R_{\oplus}$  are rocky, whereas most planets larger than  $2 R_{\oplus}$  have volatile-rich envelopes (77). Moreover, planets larger than  $3 R_{\oplus}$  are most often less dense than water, implying a higher hydrogen content in the atmosphere (80). This suggests that the (somewhat arbitrary but commonly used) definition of “Earth-size” ( $R_p < 1.25 R_{\oplus}$ ) is in need of revision.

### Requirements for Reliable Planet Occurrence Rates

*Kepler*'s primary mission objective is to study exoplanet populations. Of particular importance is the determination of  $\eta_{\oplus}$ —the frequency of Earth-size HZ planets. Although no discrimination by star type is captured in this definition, *Kepler* was designed with Earth analogs in mind: Earth-size planets in the HZ of G-type main-sequence stars. The determination of reliable planet occurrence rates requires

- i) sensitivity to small HZ planets for sufficiently large numbers of G, K, and M stars;
- ii) a uniform and reliable catalog of exoplanets with well-understood properties (radius, periods, etc.);
- iii) knowledge of *Kepler's* detection efficiency as a function of both planet and star properties;
- iv) knowledge of the catalog reliability as a function of both planet and star properties; and
- v) well-documented and accessible data products for future archive studies.

As previously mentioned, sensitivity to Earth analogs orbiting G-type stars is a challenge that is being tackled with improvements to software. Planet properties depend on knowledge of star properties, and work is underway to construct a catalog of accurate properties and characterize systematics. Catalog uniformity is achieved by removing human subjectivity from the discovery process. Each time a new planet candidate catalog is generated, there are fewer manual processes thereby improving uniformity. A machine-learning algorithm approach based on a random-forest classifier is simultaneously being developed and may eventually replace the manual processes altogether (81).

The detection efficiency is computed by injecting artificial transits at both the pixel level and the flux level. Artificial transits are propagated through the system from pixels to

planets to quantify the completeness end to end. Tests on the back end of the pipeline (pixel calibration, aperture photometry, systematic error correction, harmonic variability removal) demonstrate a 98% fidelity in preserving the signal-to-noise ratio of a single transit (82). The tests will be repeated with longer data volumes. Tests on the front end of the pipeline (whitening filters, signal detection, and vetting) are in progress. The FPP is computed for every planet candidate as described above (83), yielding a quantitative measure of catalog reliability.

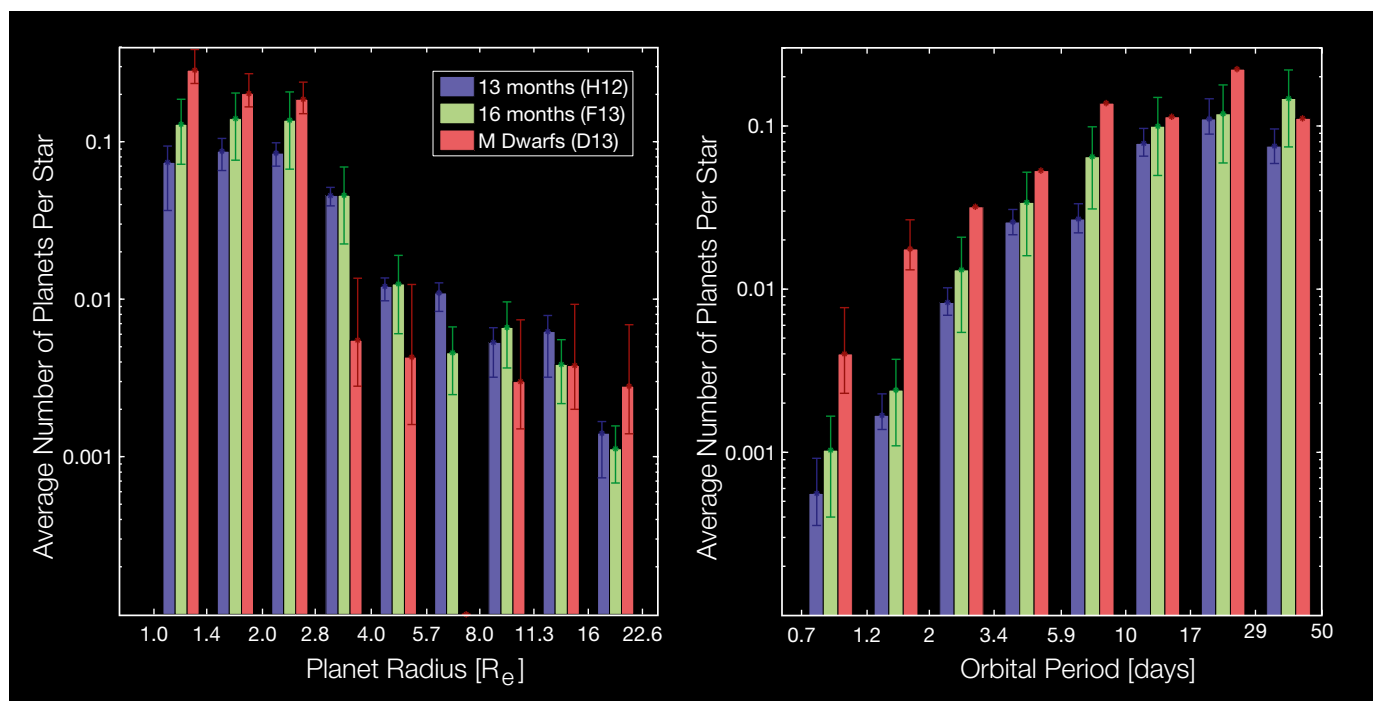
Knowledge of the statistics of multiple star systems is crucial to several key studies. They are used to construct priors for statistical validation, for computing the FPP for planet-hosting stars, and for estimating the catalog reliability. They are also required for computing *Kepler's* detection efficiency. The probability of detecting a planet of a given size and orbital period around a star is degraded in the presence of flux dilution from unresolved nearby stars (either bound or line of sight). Contaminating flux causes transits to appear shallower. We do not know a priori which stars have such dilution. However, the effect on occurrence rates can be quantified via numerical simulation based on multiple star statistics from *Kepler* (84, 85) and other (86) surveys.

It is important to note that the sample of stars observed by *Kepler* is not representative of the galactic population (87). Exoplanet occurrence rates must be broken out by star type to reconstruct a volume-limited representation of planetary populations in the galaxy. Finally, if future missions need to know how deeply they must probe before happening on a potentially habitable terrestrial planet, we must consider how the *Kepler* planets cluster into multiplanet systems and compute the fraction of stars with planets in addition to the average number of planets per star.

### Estimates of Planet Occurrence Rates

There has yet to be a study that addresses all of the requirements described above using all of the available data. Nevertheless, numerous population estimates have been reported in the literature and patterns are beginning to emerge. The most dramatic is the sharp rise in the (log) radius distribution for planets smaller than about three times the size of Earth (47, 88).

Fig. 3, *Left* shows the planet occurrence rate distribution marginalized over periods less than 50 d reported by independent teams (0.68–50 d being the common domain). A power-law distribution would be a straight line on this logarithmic display. Close-in giants are orders of magnitude less common



**Fig. 3.** The radius distribution (*Left*) and period distribution (*Right*) of planet occurrence rates expressed as the average number of planets per star. The distributions have been marginalized over periods between 0.68 and 50 d (radius distribution) and radii between 0.5 and 22.6  $R_{\oplus}$  (period distribution). H12 refers to ref. 88, F13 refers to ref. 47, and D13 refers to ref. 92. The reported one-sigma uncertainties are shown.

than planets smaller than Neptune. However, a power-law increase toward smaller sizes is not observed. The distribution flattens out for planets smaller than  $2 R_{\oplus}$ . This may be an artifact of catalog incompleteness for the smallest planets, especially at longer orbital periods.

Marginalizing over radius ( $0.5$ – $22.6 R_{\oplus}$ ), we observe a power-law increase in occurrence rate as a function of (log) period up to  $\sim 10$  d. At longer orbital periods, the distribution flattens (Fig. 3, *Right*). The trend can be explored with a larger sample that includes longer period planets. The flat distribution persists out to  $\sim 250$  d (89), at least for planets smaller than Neptune. The giants, however, appear to be gaining ground, slowly increasing in frequency (compare with figure 7 of ref. 89)—a trend that is consistent with Doppler surveys (90) and predicted by core-accretion models (91).

The HZ of M-type dwarfs corresponds to orbital periods of a few weeks to a few months. *Kepler's* current planet catalog is sufficient for addressing statistics of HZ exoplanets orbiting M stars. The results indicate that the average number of small ( $0.5$ – $1.4 R_{\oplus}$ ) HZ (optimistic) planets per M-type main-sequence star is  $\sim 0.5$  (92–94). An estimate of HZ occurrence rates for G and K stars has been made via extrapolation to longer orbital periods (95). An independent planet detection pipeline was applied to a sample of G and K stars observed by *Kepler*, and the survey completeness was quantified via signal injection. An occurrence rate of  $11 \pm 4\%$  was recovered for 1- to  $2 R_{\oplus}$  planets receiving insolation fluxes of 1 to  $4 F_{\oplus}$ .

Assuming the true occurrence rate distribution is approximately constant in (log) period for  $P > 10$  d and in (log) radius for  $R_p < 2.8 R_{\oplus}$ , the planet occurrence over any interval within that domain is proportional to the logarithmic area bounded by the interval. For a homogeneous star sample, a distribution that is constant in (log) period will, to first order, be constant in (log) insolation flux. An orbital period of 10 d corresponds to an insolation flux of  $\sim 100 F_{\oplus}$  for a Sun-like star ( $\sim 20 F_{\oplus}$  for a late K).

Under these assumptions, the reported occurrence rate of  $11 \pm 4\%$  can be scaled for small planets ( $1$ – $1.4 R_{\oplus}$ ) in the optimistic HZ ( $0.27$ – $1.70 F_{\oplus}$  for a K0-type main-sequence star). This yields an occurrence rate of  $7 \pm 3\%$ . If we assume that the (log) radius distribution remains constant down to  $0.5 R_p$ , we can estimate the occurrence rate for an interval comparable to that of the M-dwarf

calculations ( $0.5$ – $1.4 R_{\oplus}$  optimistic HZ). The G and K occurrence rate for this interval is  $22 \pm 8\%$ . At first glance, planets orbiting in the HZ of G- and K-type stars are less common than those orbiting M-type stars. We must proceed cautiously, however, because the results are based on extrapolation to longer periods to account for very high incompleteness.

Collectively, the statistics emerging from the *Kepler* data suggest that every late-type main-sequence star has at least one planet (of any size), that one in six has an Earth-size planet within a Mercury-like orbit, and that small HZ planets around M dwarfs abound. Already, the *Kepler* data suggest that a potentially habitable planet resides within 5 parsec at the 95% confidence level.

### Summary

Our blinders to small planets have been lifted, and the exoplanet landscape looks dramatically different from what it did before the launch of NASA's *Kepler* Mission. A picture is forming in which small planets abound and close-in giants are few, in which the HZs of cool stars are heavily populated with terrestrial planets and the diversity of systems challenges preconceived ideas. The picture will continue to evolve over the next

few years as we analyze the remaining data, refine the sample, and quantify the observational biases. Characterization instruments will continue to gain sensitivity ensuring that *Kepler's* exoplanet discoveries will be studied for years to come. Although *Kepler's* primary data collection has officially ended, the most significant discovery and analysis phase is underway, enabling the long-term goal of exoplanet exploration: the search for habitable environments and life beyond the solar system.

**ACKNOWLEDGMENTS.** The significant science results based on *Kepler* data described in this volume would not have been possible without the decades of persistence, hard work, creativity, and expertise of William Borucki (Principal Investigator), Dave Koch (Deputy Principal Investigator), Jon Jenkins (Analysis Lead), and Doug Caldwell (Instrument Scientist). *Kepler's* co-Investigators, Science Working Group members, Follow-Up Observers, and Participating Scientists (Participating Scientist Program) provide critical expertise and analyses that help *Kepler* meet its baseline objectives. The Science Operations Center provides the software pipeline for constructing light curves and identifying planet candidates. The Science Office, led by Michael Haas, produces the well-vetted and highly reliable catalogs that *Kepler's* planet occurrence rates are derived from. This research has made use of the National Aeronautics and Space Administration (NASA) Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the NASA under the Exoplanet Exploration Program. *Kepler* was competitively selected as the 10th Discovery Mission. Funding for this mission is provided by NASA's Science Mission Directorate.

- 1 Koch DG, et al. (2010) *Kepler* Mission design, realized photometric performance, and early science. *Astrophys J Lett* 713(2):L79–L86.
- 2 Jenkins JM, et al. (2010) Initial characteristics of *Kepler* long cadence data for detecting transiting planets. *Astrophys J Lett* 713(2):L120–L125.
- 3 Wu H, et al. (2010) Data validation in the *Kepler* Science Operations Center pipeline. *Proc SPIE* 7740:774019.
- 4 Henry GW, Marcy GW, Butler RP, Vogt SS (2000) A transiting “51 peg-like” planet. *Astrophys J* 529(1):L41–L44.
- 5 Charbonneau D, Brown TM, Latham DW, Mayor M (2000) Detection of planetary transits across a Sun-like star. *Astrophys J* 529(1):L45–L48.
- 6 Lissauer JJ, et al. (2011) Architecture and dynamics of *Kepler's* candidate multiple transiting planet systems. *Astrophys J Suppl Ser* 197(1):8.
- 7 Borucki WJ, et al. (2011) Characteristics of *Kepler* planetary candidates based on the first data set. *Astrophys J* 728(2):117.
- 8 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.
- 9 Batalha NM, et al. (2013) Planetary candidates observed by *Kepler*. III. Analysis of the first 16 months of data. *Astrophys J Suppl Ser* 204(2):24.
- 10 Burke CJ, et al. (2014) Planetary candidates observed by *Kepler* IV: Planet sample from Q1–Q8 (22 Months). *Astrophys J Suppl Ser* 210(2):19.
- 11 Rowe J; *Kepler* Team. (2014) The *Kepler* Q1–Q12 Planet Candidate Catalogue, *American Astronomical Society Meeting Abstracts* (American Astronomical Society, Washington, DC), Vol 223, p 228.02.
- 12 Mandel K, Agol E (2002) Analytic light Curves for Planetary Transit Searches. *Astrophys J Lett* 580(2):L171–L175.
- 13 Brown TM, Latham DW, Everett ME, Esquerdo GA (2011) *Kepler* Input Catalog: Photometric calibration and stellar classification. *Astron J* 142(4):112.
- 14 Verner GA, et al. (2011) Global asteroseismic properties of solar-like oscillations observed by *Kepler*: A comparison of complementary analysis methods. *Mon Not R Astron Soc* 415(4):3539–3551.
- 15 Gaidos E, Mann AW (2013) Objects in *Kepler's* mirror may be larger than they appear: Bias and selection effects in transiting planet surveys. *Astrophys J* 762(1):41.
- 16 Pinsonneault MH, et al. (2012) A revised effective temperature scale for the *Kepler* Input Catalog. *Astrophys J Suppl Ser* 199(2):30.
- 17 Mann AW, Gaidos E, Lépine S, Hilton EJ (2012) They might be giants: Luminosity class, planet occurrence, and planet-metallicity relation of the coolest *Kepler* target stars. *Astrophys J* 753(1):90.
- 18 Muirhead PS, et al. (2012) Characterizing the cool *Kepler* Objects of Interest (KOIs). New effective temperatures, metallicities, masses, and radii of low-mass *Kepler* planet-candidate host stars. *Astrophys J Lett* 750(2):L37.
- 19 Huber D, et al. (2014) Revised stellar properties of *Kepler* targets for the quarter 1–16 transit detection run. *Astrophys J Suppl Ser* 211(1):2.
- 20 Tenenbaum P, et al. (2013) Detection of potential transit signals in sixteen quarters of *Kepler* Mission data. *Astrophys J Suppl Ser* 206(1):5.
- 21 Ofir A, Dreizler S (2013) An independent planet search in the *Kepler* dataset. I. One hundred new candidates and revised *Kepler* objects of interest. *Astron Astrophys* 555:A58.
- 22 Petigura EA, Marcy GW, Howard AW (2013) A plateau in the planet population below twice the size of Earth. *Astrophys J* 770(1):69.
- 23 Lintott CJ, et al. (2013) Planet hunters: New *Kepler* planet candidates from analysis of quarter 2. *Astron J* 145(6):151.
- 24 Wang J, et al. (2013) Planet hunters. V. A confirmed Jupiter-size planet in the habitable zone and 42 planet candidates from the *Kepler* Archive Data. *Astrophys J* 776(1):10.
- 25 Schmitt JR, et al. (2013) Planet hunters VI: The first *Kepler* seven planet candidate system and 13 other planet candidates from the *Kepler* Archival Data. arXiv:1310.5912.
- 26 Schwamb ME, et al. (2013) Planet hunters: A transiting circumbinary planet in a quadruple star system. *Astrophys J* 768(2):127.
- 27 Sanchis-Ojeda R, et al. (2014) A study of the shortest-period planets found with *Kepler*. *Astrophys J* 787(1):47.
- 28 Kasting JF, Kopparapu R, Ramirez RM, Harman C (2013) Remote life detection criteria, habitable zone boundaries, and the frequency of Earthlike planets around M and late-K stars. arXiv:1312.1328.

- 29 Abe Y, Abe-Ouchi A, Sleep NH, Zahnle KJ (2011) Habitable zone limits for dry planets. *Astrobiology* 11(5):443–460.
- 30 Zsom A, Seager S, de Wit J, Stamenković V (2013) Toward the minimum inner edge distance of the habitable zone. *Astrophys J* 778(2):109.
- 31 Leconte J, Forget F, Charnay B, Wordsworth R, Pottier A (2013) Increased insolation threshold for runaway greenhouse processes on Earth-like planets. *Nature* 504(7479):268–271.
- 32 Yang J, Boué G, Fabrycky DC, Abbot DS (2014) Strong dependence of the inner edge of the habitable zone on planetary rotation rate. *Astrophys J Lett* 787(1):L2.
- 33 Lissauer JJ (2007) Planets formed in habitable zones of M dwarf stars probably are deficient in volatiles. *Astrophys J Lett* 660(2):L149–L152.
- 34 Koppapu RK, et al. (2014) Habitable zones around main-sequence stars: Dependence on planetary mass. *Astrophys J Lett* 787(2):L29.
- 35 Borucki WJ, et al. (2012) Kepler-22b: A 2.4 Earth-radius planet in the habitable zone of a Sun-like star. *Astrophys J* 745(2):120.
- 36 Ballard S, et al. (2013) Exoplanet characterization by proxy: A transiting 2.15 R<sub>J</sub> planet near the habitable zone of the late K dwarf Kepler-61. *Astrophys J* 773(2):98.
- 37 Borucki WJ, et al. (2013) Kepler-62: A five-planet system with planets of 1.4 and 1.6 Earth radii in the habitable zone. *Science* 340(6132):587–590.
- 38 Quintana EV, et al. (2014) An Earth-sized planet in the habitable zone of a cool star. *Science* 344(6181):277–280.
- 39 Rowe JF, et al. (2014) Validation of Kepler's multiple planet candidates. III. Light curve analysis and announcement of hundreds of new multi-planet systems. *Astrophys J* 784(1):45.
- 40 Jenkins JM (2002) The impact of solar-like variability on the detectability of transiting terrestrial planets. *Astrophys J* 575(1):493–505.
- 41 Gilliland RL, et al. (2011) Kepler Mission stellar and instrument noise properties. *Astrophys J Suppl Ser* 197(1):6.
- 42 Bryson ST, et al. (2013) Identification of background false positives from Kepler data. *Publ Astron Soc Pac* 125(930):889–923.
- 43 Batalha NM, et al. (2010) Pre-spectroscopic false-positive elimination of Kepler planet candidates. *Astrophys J Lett* 713(2):L103–L108.
- 44 Morton TD, Johnson JA (2011) On the low false positive probabilities of Kepler planet candidates. *Astrophys J* 738(2):170.
- 45 Santerne A, et al. (2012) SOPHIE velocimetry of Kepler transit candidates (2012) VII. A false-positive rate of 35% for Kepler close-in giant candidates. *Astron Astrophys* 545:A76.
- 46 Désert J-M, Charbonneau D, Fressin F, Torres G (2012) Using Spitzer to estimate the Kepler false positive rate and to validate Kepler Candidates. *American Astronomical Society Meeting Abstracts* (American Astronomical Society, Washington, DC), Vol 219, p 414.02.
- 47 Fressin F, et al. (2013) The false positive rate of Kepler and the occurrence of planets. *Astrophys J* 766(2):81.
- 48 Santerne A, et al. (2013) The contribution of secondary eclipses as astrophysical false positives to exoplanet transit surveys. *Astron Astrophys* 557:A139.
- 49 Santerne A, et al. (2011) SOPHIE velocimetry of Kepler transit candidates. II. KOI-428b: A hot Jupiter transiting a subgiant F-star. *Astron Astrophys* 528:A63.
- 50 Bouchy F, et al. (2011) SOPHIE velocimetry of Kepler transit candidates. III. KOI-423b: An 18 MJup transiting companion around an F7IV star. *Astron Astrophys* 533:A83.
- 51 Bonomo AS, et al. (2012) SOPHIE velocimetry of Kepler transit candidates. V. The three hot Jupiters KOI-135b, KOI-204b, and KOI-203b (alias Kepler-17b). *Astron Astrophys* 538:A96.
- 52 Buchhave LA, et al. (2011) Kepler-14b: A massive hot Jupiter transiting an F star in a close visual binary. *Astrophys J Suppl Ser* 197(1):3.
- 53 Endl M, et al. (2011) Kepler-15b: A hot Jupiter enriched in heavy elements and the first Kepler Mission planet confirmed with the Hobby-Eberly telescope. *Astrophys J Suppl Ser* 197(1):13.
- 54 Hébrard G, et al. (2013) KOI-200 b and KOI-889 b: Two transiting exoplanets detected and characterized with Kepler, SOPHIE, and HARPS-N. *Astron Astrophys* 554:A114.
- 55 Borucki WJ, et al. (2010) Kepler planet-detection mission: Introduction and first results. *Science* 327(5968):977–980.
- 56 Batalha NM, et al. (2011) Kepler's first rocky planet: Kepler-10b. *Astrophys J* 729(1):27.
- 57 Howard AW, et al. (2013) A rocky composition for an Earth-sized exoplanet. *Nature* 503(7476):381–384.
- 58 Pepe F, et al. (2013) An Earth-sized planet with an Earth-like density. *Nature* 503(7476):377–380.
- 59 Marcy GW, et al. (2014) Masses, radii, and orbits of small Kepler planets: The transition from gaseous to rocky planets. *Astrophys J Suppl Ser* 210(2):20.
- 60 Holman MJ, et al. (2010) Kepler-9: A system of multiple planets transiting a Sun-like star, confirmed by timing variations. *Science* 330(6000):51–54.
- 61 Ford EB, et al. (2012) Transit timing observations from Kepler. II. Confirmation of two multiplet systems via a non-parametric correlation analysis. *Astrophys J* 750(2):113.
- 62 Steffen JH, et al. (2012) Transit timing observations from Kepler - III. Confirmation of four multiple planet systems by a Fourier-domain study of anticorrelated transit timing variations. *Mon Not R Astron Soc* 421(3):2342–2354.
- 63 Fabrycky DC, et al. (2012) Transit timing observations from Kepler. IV. Confirmation of four multiple-planet systems by simple physical models. *Astrophys J* 750(2):114.
- 64 Steffen JH, et al. (2013) Transit timing observations from Kepler - VII. Confirmation of 27 planets in 13 multiplet systems via transit timing variations and orbital stability. *Mon Not R Astron Soc* 428(2):1077–1087.
- 65 Wu Y, Lithwick Y (2013) Density and eccentricity of Kepler planets. *Astrophys J* 772(1):74.
- 66 Xie J-W (2013) Transit timing variation of near-resonance planetary pairs: Confirmation of 12 multiple-planet systems. *Astrophys J Suppl Ser* 208(2):22.
- 67 Xie J-W (2014) Transit timing variation of near-resonance planetary pairs. II. confirmation of 30 planets in 15 multiple planet systems. *Astrophys J Suppl Ser* 210(2):25.
- 68 Ming Y, Hui-Gen L, Hui Z, Jia-Yi Y, Ji-Lin Z (2013) Eight planets in four multi-planet systems via transit timing variations in 1350 days. *Astrophys J* 778(2):110.
- 69 Lissauer JJ, et al. (2011) A closely packed system of low-mass, low-density planets transiting Kepler-11. *Nature* 470(7332):53–58.
- 70 Lissauer JJ, et al. (2013) All six planets known to orbit Kepler-11 have low densities. *Astrophys J* 770(2):131.
- 71 Gautier TN III, et al. (2012) Kepler-20: A Sun-like star with three sub-Neptune exoplanets and two Earth-size candidates. *Astrophys J* 749(1):15.
- 72 Sanchis-Ojeda R, et al. (2012) Alignment of the stellar spin with the orbits of a three-planet system. *Nature* 487(7408):449–453.
- 73 Cochran WD, et al. (2011) Kepler-18b, c, and d: A system of three planets confirmed by transit timing variations, light curve validation, Warm-Spitzer photometry, and radial velocity measurements. *Astrophys J Suppl Ser* 197(1):7.
- 74 Ofir A, Dreizler S, Zechmeister M, Husser T-O (2014) An independent planet search in the Kepler dataset. II. An extremely low-density super-Earth mass planet around Kepler-87. *Astron Astrophys* 561:A103.
- 75 Jontof-Hutter D, Lissauer JJ, Rowe JF, Fabrycky DC (2014) Kepler-79's low density planets. *Astrophys J Lett* 783(1):L6.
- 76 Carter JA, et al. (2012) Kepler-36: A pair of planets with neighboring orbits and dissimilar densities. *Science* 337(6094):556–559.
- 77 Weiss LM, Marcy GW (2014) The mass-radius relation for 65 exoplanets smaller than 4 earth radii. *Astrophys J Lett* 783(1):L6.
- 78 Lopez ED, Fortney JJ (2013) Understanding the mass-radius relation for sub-Neptunes: Radius as a proxy for composition. arXiv:1311.0329.
- 79 Kipping DM, et al. (2014) The hunt for exomoons with Kepler (HEK). IV. A search for moons around eight M dwarfs. *Astrophys J* 784(1):28.
- 80 Hadden S, Lithwick Y (2014) Densities and eccentricities of 139 Kepler planets from transit time variations. *Astrophys J* 787(1):80.
- 81 Jenkins JM, et al. (2014) Likely planet candidates identified by machine learning applied to four years of Kepler data. *American Astronomical Society Meeting Abstracts* (American Astronomical Society, Washington, DC), Vol 223, p 206.02.
- 82 Christiansen JL, et al. (2013) Measuring transit signal recovery in the Kepler pipeline. I. Individual events. *Astrophys J Suppl Ser* 207(2):35.
- 83 Morton TD (2012) An Efficient automated validation procedure for exoplanet transit candidates. *Astrophys J* 761(1):6.
- 84 Prša A, et al. (2011) Kepler eclipsing binary stars. I. Catalog and principal characterization of 1879 eclipsing binaries in the first data release. *Astron J* 141(3):83.
- 85 Slawson RW, et al. (2011) Kepler eclipsing binary stars. II. 2165 eclipsing binaries in the second data release. *Astron J* 142(5):160.
- 86 Raghavan D, et al. (2010) A survey of stellar families: Multiplicity of solar-type Stars. *Astrophys J Suppl Ser* 190(1):1–42.
- 87 Batalha NM, et al. (2010) Selection, prioritization, and characteristics of Kepler target stars. *Astrophys J Lett* 713(2):L109–L114.
- 88 Howard AW, et al. (2012) Planet occurrence within 0.25 AU of solar-type stars from Kepler. *Astrophys J Suppl Ser* 201(2):15.
- 89 Dong S, Zhu Z (2013) Fast rise of "Neptune-size" planets (4–8 R<sub>J</sub>) from P ~ 10 to ~250 days – statistics of Kepler planet candidates up to ~0.75 AU. *Astrophys J* 778(1):53.
- 90 Udry S, Santos NC (2007) Statistical properties of exoplanets. *Annu Rev Astron Astrophys* 45(1):397–439.
- 91 Ida S, Lin DNC (2004) Toward a deterministic model of planetary formation. I. A desert in the mass and semimajor axis distributions of extrasolar planets. *Astrophys J* 604(1):388–413.
- 92 Dressing CD, Charbonneau D (2013) The occurrence rate of small planets around Small Stars. *Astrophys J* 767(1):95.
- 93 Koppapu 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.
- 94 Gaidos E (2013) Candidate planets in the habitable zones of Kepler stars. *Astrophys J* 770(2):90.
- 95 Petigura EA, Howard AW, Marcy GW (2013) Prevalence of Earth-size planets orbiting Sun-like stars. *Proc Natl Acad Sci USA* 110(48):19273–19278.