The Astrophysical Journal Letters, 893:L27 (8pp), 2020 April 10 https://doi.org/10.3847/2041-8213/ab84e5

A Habitable-zone Earth-sized Planet Rescued from False Positive Status

Andrew Vanderburg[1,9], Pamela Rowden[2], Steve Bryson[3], Jeffrey Coughlin[4], Natalie Batalha[5], Karen A. Collins[6], David W. Latham[6], Susan E. Mullally[7], Knicole D. Colón[8], Chris Henze[3], Chelsea X. Huang[6], and Samuel N. Quinn[6]
[1] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA; avanderburg@utexas.edu
[2] School of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK
[3] NASA Ames Research Center, Moffett Field, CA 94901, USA
[4] SETI Institute, 189 Bernardo Avenue, Suite 200, Mountain View, CA 94043, USA
[5] University of California Santa Cruz, Santa Cruz, CA, USA
[6] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[7] Space Telescope Science Institute, Baltimore, MD 21218, USA
[8] NASA Goddard Space Flight Center, Exoplanets and Stellar Astrophysics Laboratory (Code 667), Greenbelt, MD 20771, USA
Received 2020 February 29; revised 2020 March 25; accepted 2020 March 26; published 2020 April 15

Abstract

We report the discovery of an Earth-sized planet in the habitable zone of a low-mass star called Kepler-1649. The planet, Kepler-1649 c, is $1.06^{+0.15}_{-0.10}$ times the size of Earth and transits its $0.1977 \pm 0.0051\ M_{\odot}$ "mid" M-dwarf host star every 19.5 days. It receives 74% ± 3% the incident flux of Earth, giving it an equilibrium temperature of 234 ± 20 K and placing it firmly inside the circumstellar habitable zone. Kepler-1649 also hosts a previously known inner planet that orbits every 8.7 days and is roughly equivalent to Venus in size and incident flux. Kepler-1649 c was originally classified as a false positive (FP) by the Kepler pipeline, but was rescued as part of a systematic visual inspection of all automatically dispositioned Kepler FPs. This discovery highlights the value of human inspection of planet candidates even as automated techniques improve, and hints that terrestrial planets around mid to late M-dwarfs may be more common than those around more massive stars.

Unified Astronomy Thesaurus concepts: Exoplanet astronomy (486); Habitable planets (695)

Supporting material: data behind figure

1. Introduction

M-dwarf stars ($0.1M_{\odot} \lesssim M \lesssim 0.6M_{\odot}$) are the most common outcome of the star formation process in our galaxy (Kroupa 2001)—they make up two-thirds of all stars and brown dwarfs in the solar neighborhood (Henry et al. 2018).[10] Exoplanet surveys have found that these stars frequently host terrestrial-sized planets, including those in temperate orbits that could possibly support liquid water on their surfaces (Dressing & Charbonneau 2013). Although M-dwarfs may be less hospitable than higher-mass stars like the Sun (e.g., Luger & Barnes 2015; Howard et al. 2018), their sheer numbers make it plausible that planets around M-dwarfs may be the most common habitable environments in the universe.

Our knowledge of planets around M-dwarfs has greatly increased in the last decade thanks to observations from the Kepler space telescope. Kepler was designed to measure how frequently planets are found around Sun-like stars, but a few thousand M-dwarf stars were also observed in its survey. From these data, we have learned that on average, each M-dwarf star hosts more than two sub-Neptune-sized planets with periods shorter than 200 days (Dressing & Charbonneau 2015). Small planets are found more frequently around M-dwarfs than around higher-mass Sun-like stars (Mulders et al. 2015) and there is tentative evidence that this trend holds even within the M-dwarf spectral class: lower-mass ($M_{\star} \approx 0.25M_{\odot}$) "mid" M-dwarfs may host even more small planets than higher-mass ($M_{\star} \approx 0.5M_{\odot}$) "early" M-dwarfs (Muirhead et al. 2015; Hardegree-Ullman et al. 2019).

These and other statistical results have been enabled by the well-characterized planet candidate (PC) catalogs produced by the Kepler mission. Pixel time series were downloaded from the spacecraft and processed by the Kepler pipeline, which performed image calibration, extracted light curves, removed instrumental systematics, and searched for periodic flux decrements that could be due to a transiting planet (Jenkins et al. 2010; Jenkins 2017); these initial detections are known as Threshold Crossing Events (TCEs). The TCE Review Team (TCERT) then reviewed and dispositioned (classified) TCEs as either PCs or false positives (FPs), with TCEs potentially due to any astrophysical transiting/eclipsing object given a Kepler Object of Interest (KOI) number. By the end of the mission, the TCERT process was fully automated via the Robovetter (Coughlin et al. 2016; Thompson et al. 2018), which uses dozens of specialized metrics and a sophisticated decision tree to classify TCEs—it only used Kepler observations, as other measurements (e.g., ground-based spectroscopic follow-up) were not uniformly performed on all targets. This uniform vetting, along with associated synthetic data products (Burke & Catanzarite 2017; Christiansen 2017; Coughlin 2017), allows for the measurement of the final catalogs' completeness and reliability, thus enabling the accurate determination of planetary occurrence rates. As a result, individual disposition correctness was sacrificed for statistical uniformity, and so it was known that some individual KOIs were incorrectly vetted, with interesting planets misclassified as FPs, and vice versa.

FP KOIs are generally not followed up, possibly ignoring true planets that were incorrectly dispositioned as FPs. To address this issue, members of our team formed the Kepler False Positive Working Group (FPWG; Bryson et al. 2017) to visually inspect, using all available data, all KOIs classified as FPs by the Robovetter. Its goals were as follows.

[9] NASA Sagan Fellow.
[10] http://www.recons.org/census.posted.htm

1. Produce a list of known "certified" FPs,[11] which can be used as a ground-truth when testing new classifiers.
2. Diagnose any issues or failure modes in the Robovetter's algorithm which could be corrected to improve its classifications.
3. Identify and rescue any viable PCs that were erroneously classified as FPs.

Over the past five years, the FPWG has inspected nearly 5000 KOIs and certified nearly 4000 KOIs as FPs or false alarms. In the course of this review, we also examined objects that we were unable to certify as FPs and identified the ones most likely to be viable PCs. Most recently, the FPWG finished a review of all Data Release 25 (DR25) FP KOIs, including those at very low signal-to-noise. Among these possible PCs, one signal stood out as both particularly high quality and scientifically interesting: an Earth-sized PC in a temperate orbit around a nearby low-mass star.

Here, we report our investigation of this newly rescued PC signal. We take advantage of the fact that the candidate's host star, Kepler-1649, was already shown to be an exoplanet host by Angelo et al. (2017), who characterized and validated an inner planet called Kepler-1649 b. Angelo et al. (2017) described Kepler-1649 b as a "Venus analog" because it is similar in size and incident flux to our solar system neighbor. In this Letter, we validate our newly rescued candidate as a planet in the system and show that in addition to a Venus analog, Kepler-1649 hosts an Earth analog as well. Section 2 describes the observations and analysis that we used to characterize this new signal, and Section 3 describes our statistical validation of Kepler-1649 c. Finally, we conclude in Section 4 by discussing Kepler-1649 c's characteristics, the system's architecture, and the implications of this detection regarding the occurrence rate of rocky, habitable planets around M-dwarfs.

2. Observations and Analysis

2.1. Kepler Light Curve

The Kepler space telescope observed Kepler-1649 (KIC 6444896/KOI 3138) for a total of 756 days between 2010 and 2013 during its primary mission. Kepler-1649 was observed during Quarters 6–9 (as part of guest investigator proposal GO20031, PI: Di Stefano) and Quarters 12–17 (after KOI 3138.01/Kepler-1649 b was designated a PC).

A signal with period $P \approx 8.689$ days was first detected, designated as KOI 3138.01, and dispositioned as a PC in the Burke et al. (2014) catalog, which used data from Quarter 1 to Quarter 8, or only nine months of data for this target (Quarters 6, 7, and 8). It was also dispositioned as a PC in each subsequent catalog that re-examined it (the last two of which were based on the Robovetter; Coughlin 2017; Thompson et al. 2018). This signal eventually became known as Kepler-1649 b after statistical validation by Angelo et al. (2017).

A second signal with $P \approx 19.535$ days was only detected in the final Kepler pipeline run (TCE 6444896-02; Twicken et al. 2016). The Thompson et al. (2018) Robovetter dispositioned the TCE as a not-transit-like FP (i.e., a false alarm) with a comment of "MOD_NONUNIQ_ALT," which indicates it failed the Model-Shift Uniqueness test (see Section A.3.4 of Thompson et al. 2018). This outcome indicates that the Robovetter judged the signal's significance to be too low compared to the systematic noise level to be a PC. The Robovetter assigned the 19.5 days TCE a "disposition score" (see Section 3.2 of Thompson et al. 2018) of 0.374, which indicates only weak confidence in the FP disposition (scores near 0.0 indicate high-confidence FPs, while scores near 1.0 indicate high-confidence PCs). All TCEs with disposition scores >0.1 were assigned KOI numbers by Thompson et al. (2018), and thus KOI 3838.02 was created. For convenience, we hereafter refer to this new PC (KOI 3138.02) as Kepler-1649 c.

Our team inspected the transit signal of Kepler-1649 c as part of our systematic review of Kepler's FP objects of interest. Though our assessments of the FPs were usually in good agreement with the Robovetter, we were not confident in the Robovetter's FP classification. Unlike most of the false alarms that the Robovetter fails with the Model-Shift Uniqueness test, Kepler-1649 c has a relatively short orbital period and dozens of observed transits, which were consistent in shape and depth over time. We found no compelling evidence that Kepler-1649 c was an FP, and instead identified it as a possible PC.

We hypothesized that Kepler-1649 c failed the Robovetter's Model-Shift Uniqueness test because of the atypical noise properties of the light curve produced by the Kepler pipeline. The Kepler pipeline light curve shows strong quarterly variations in its photometric scatter; for example, the photometric scatter in Quarter 7 (≈3500 ppm) is almost four times greater than the photometric scatter in Quarter 6 (≈900 ppm). We traced this effect to the choice of photometric apertures by the Kepler pipeline. Kepler-1649 has a high proper motion of 168.2 mas per year, but its motion was not taken into account when the Kepler pipeline selected optimal pixels for aperture photometry (Bryson et al. 2010). Instead, the pixel selection algorithm assumed Kepler-1649's J2000 position, about 2″ (or half of a Kepler pixel) from its true position during Kepler's observations. Because Kepler-1649 is faint, its optimal photometric apertures were small enough that this half-pixel error caused Kepler-1649 to fall at the edge or outside of the aperture in some quarters, while remaining within the photometric aperture in other quarters. The variation of the pixel position of Kepler-1649 relative to the pixels selected for photometry significantly weakened the transit signal in the original pipeline photometry.

We therefore chose to produce our own light curves from the Kepler target pixel files (DR25). Following Vanderburg et al. (2016), we extracted light curves from a set of 20 different photometric apertures. Half of these apertures were circular, defined by identifying all pixels within 10 different radii of Kepler-1649's position. The other half were shaped like the Kepler Pixel Response Function (PRF), defined by fitting the PRF to a representative Kepler image for each quarter and selecting all pixels where the model PRF exceeded 10 different fractions of the peak model flux. We selected the aperture that produced the light curve with the highest photometric precision[12] in each quarter after accounting for diluting flux from nearby fainter stars. The resulting light curve had much

[11] https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=fpwg

[12] We measured the photometric precision on 6 hr timescales by applying a 13 point standard deviation filter to the light curve and taking the median value. This is a quick approximation to the Kepler pipeline's Combined Differential Photometric Precision metric (Christiansen et al. 2012). For more information, see the documentation for the PyKE routine `kepstddev` (https://keplerscience.arc.nasa.gov/ContributedSoftwareKepstddev.shtml, Still & Barclay 2012) and Section 3.1 of Vanderburg & Johnson (2014).

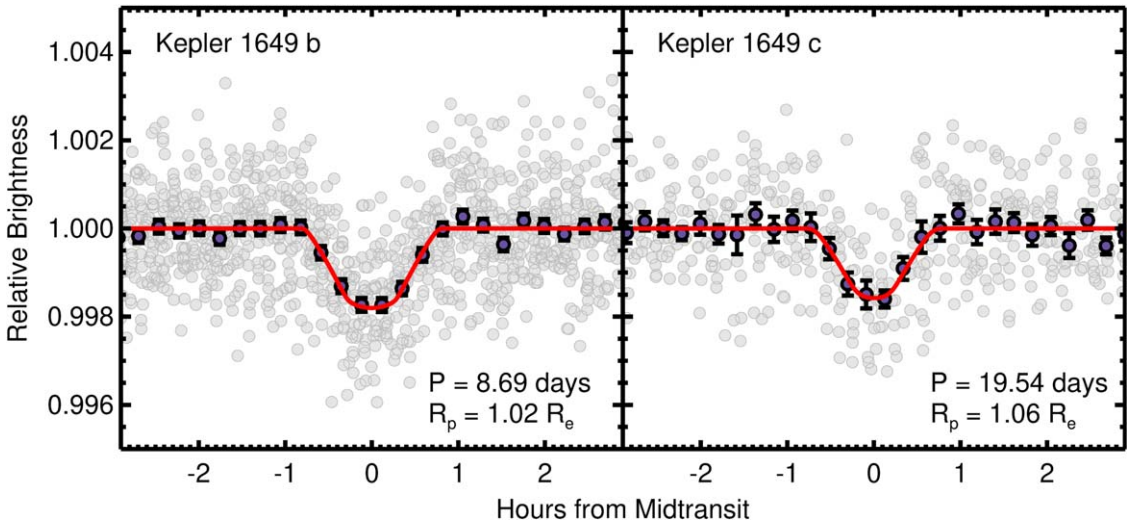

Figure 1. Kepler light curve phase-folded on the transits of Kepler-1649 b (left) and c (right). Gray points are individual Kepler long-cadence exposures, purple points are averages in phase, and the red solid line is the best-fit transit model.

(The data used to create this figure are available.)

more consistent photometric scatter across quarters, improving significantly upon the Kepler pipeline light curve in the quarters where its aperture selection was suboptimal. We use this new light curve in our analysis throughout the rest of the Letter and show the phase-folded transit signals of Kepler-1649 b and c in Figure 1. We also experimented with fitting the Kepler images with a model based on the telescope's measured PRF to produce light curves (as done by Angelo et al. 2017), but never achieved higher photometric precision than our well-optimized aperture photometry.

Finally, we re-ran the Model-Shift Uniqueness test using our new light curve, and found that this time, the PC solidly passed. We therefore consider Kepler-1649 c to be a viable PC.

2.2. Stellar Parameters

Although the star Kepler-1649 was already well characterized by Angelo et al. (2017), we were able to improve upon their stellar parameters thanks to the newly released trigonometric parallax from Gaia Data Release 2 (DR2; Gaia Collaboration et al. 2016, 2018). We derive Kepler-1649's radius and mass using empirical relations between these quantities and the star's absolute K-band magnitude from Mann et al. (2015) and Mann et al. (2019), respectively. These relations yield a mass of $M_\star = 0.1977 \pm 0.0051\ M_\odot$ and radius of $R_\star = 0.2317 \pm 0.0049\ R_\odot$, which are consistent with, but several times more precise than, the estimates from Angelo et al. (2017). We adopt the spectroscopic metallicity and temperature reported by Angelo et al. (2017) of [M/H] $= -0.15 \pm 0.11$ and $T_{\rm eff} = 3240 \pm 61$ K. Using an average of three different bolometric corrections for the 2MASS *JHK* magnitudes from Mann et al. (2015), we find that Kepler-1649 is only about half a percent as luminous as the Sun ($L_\star = 0.00516 \pm 0.00020\ L_\odot$). This value is in good agreement with the luminosity calculated from our adopted effective temperature and stellar radius using the Stefan-Boltzmann law ($L_\star = 0.00533 \pm 0.00046\ L_\odot$), indicating that our stellar parameters are self-consistent. Our adopted stellar parameters are listed in Table 1.

2.3. Transit Modeling

We determined planetary parameters with a Markov Chain Monte Carlo (MCMC) analysis of the Kepler light curve. We modeled the transit profiles of both planets simultaneously using Mandel & Agol (2002) curves, oversampling by a factor of 30 and integrating to account for the 29.4 minute Kepler long-cadence exposure times. We fit for the host's stellar density, $\rho_\star$, imposing a Gaussian prior centered at 22.5 g cm^{-3} with width 1.4 g cm^{-3}, and calculated the scaled semimajor axes ($a/R_\star$) for both planets from the $\rho_\star$ value at each MCMC step using Kepler's third law. Because both Kepler-1649 b and c orbit too far from their host star to have undergone tidal circularization, we model eccentric orbits for both planets, fitting in $\sqrt{e}\sin\omega$ and $\sqrt{e}\cos\omega$, where e is the orbital eccentricity, and ω is the argument of periastron. Several groups have shown that planets in multi-transiting systems tend to have lower orbital eccentricity than planets in single-transiting systems, so we imposed a Gaussian prior on each planet's eccentricity centered at 0 with width 0.103, the 2σ upper limit from van Eylen et al. (2019). We assumed a quadratic limb-darkening law, fitting for coefficients using the q_1 and q_2 parameterization recommended by Kipping (2013). Because the transits of Kepler-1649 b and c are both short-duration and heavily distorted by the Kepler long-cadence integration time, we imposed Gaussian priors on Kepler-1649's u_1 and u_2 limb-darkening coefficients. The priors were centered on values ($u_1 = 0.33$, $u_2 = 0.39$) predicted by model atmospheres calculated by Claret & Bloemen (2011) and had widths of 0.07, the empirically measured scatter between these model predictions and measured values (Müller et al. 2013). Finally, we limited the planet radii by enforcing $R_p/R_\star < 1$. In other words, the planets cannot be larger than their host star.

All in all, we fit 16 parameters: the stellar density, two limb-darkening parameters (q_1 and q_2), a constant flux offset, and for each planet, orbital period, transit time, orbital inclination, $\log(R_p/R_\star)$, $\sqrt{e}\sin\omega$, and $\sqrt{e}\cos\omega$. We explored parameter space with an affine invariant MCMC sampler (Goodman & Weare 2010) evolving each of 100 walkers for 100,000 steps, removing the first 10,000 steps as burn-in. The results of our fit

Table 1
System Parameters for Kepler-1649

Parameter	Value	Comment
Other Designations		
KIC 6444896		
KOI 3138		
LSPM J1930 + 4149		
Gaia DR2 2125699062780742016		
Basic Information		
R.A.	19:30:00.9006122986	A
decl.	+41:49:49.513849537	A
Proper Motion in R.A. (mas yr^{-1})	-135.842 ± 0.112	A
Proper Motion in decl. (mas yr^{-1})	-99.232 ± 0.139	A
Distance to Star (pc)	92.5 ± 0.5	A
Gaia G-magnitude	16.2682 ± 0.001	A
K-magnitude	12.589 ± 0.026	B
Stellar Parameters		
Mass, $M_\star$ ($M_\odot$)	0.1977 ± 0.0051	A,B,C
Radius, $R_\star$ ($R_\odot$)	0.2317 ± 0.0049	A,B,C
Surface Gravity, $\log g_\star$ (cgs)	5.004 ± 0.021	A,B,C
Metallicity [M/H]	-0.15 ± 0.11	E
Effective Temperature, $T_{\rm eff}$ (K)	3240 ± 61	E
Luminosity ($L_\odot$)	0.00516 ± 0.00020	C
Kepler-1649 b		
Orbital Period, P (days)	8.689099 ± 0.000025	D
Radius Ratio, $R_P/R_\star$	0.0402 ± 0.0018	D
Scaled Semimajor Axis, $a/R_\star$	44.77 ± 0.93	D
Orbital Inclination, i (deg)	$89.15^{+0.11}_{-0.079}$	D
Transit Impact Parameter, b	$0.65^{+0.072}_{-0.12}$	D
Transit Duration, t_{14} (hr)	$1.184^{+0.085}_{-0.066}$	D
Time of Transit, t_t (BJD)	2455374.6219 ± 0.0016	D
Planet Radius, R_P ($R_\oplus$)	1.017 ± 0.051	A,B,C,D
Incident Flux, S ($S_\oplus$)	2.208 ± 0.094	A,B,C,D,E
Equilibrium Temperature, $T_{\rm eq}$ (K)	307 ± 26	A,B,C,D,E,F
Kepler-1649 c		
Orbital Period, P (days)	19.53527 ± 0.00010	D
Radius Ratio, $R_P/R_\star$	$0.042^{+0.0055}_{-0.0038}$	D
Scaled Semimajor Axis, $a/R_\star$	76.8 ± 1.6	D
Orbital Inclination, i (deg)	89.339 ± 0.056	D
Transit Impact Parameter, b	0.875 ± 0.074	D
Transit Duration, t_{14} (hr)	$1.07^{+0.15}_{-0.21}$	D
Time of Transit, t_t (BJD)	2455410.9777 ± 0.0033	D
Planet Radius, R_P ($R_\oplus$)	$1.06^{+0.15}_{-0.10}$	A,B,C,D
Incident Flux, S ($S_\oplus$)	0.750 ± 0.032	A,B,C,D,E
Equilibrium Temperature, $T_{\rm eq}$ (K)	234 ± 20	A,B,C,D,E,F
Other Fit Parameters		
Linear Limb-darkening Parameter (u_1)	0.330 ± 0.070	G
Quadratic Limb-darkening Parameter (u_2)	0.392 ± 0.069	G
Transformed Limb-darkening Parameter 1 (q_1)	0.52 ± 0.14	G
Transformed Limb-darkening Parameter 2 (q_2)	0.229 ± 0.036	G
Constant Flux Offset Parameter (δF)	0.000030 ± 0.000029	D

Note. A: Parameters come from Gaia DR2. B: Parameters come from 2MASS (Skrutskie et al. 2006). C: Parameters come from empirical relations (Mann et al. 2015, 2019). D: Parameters come from our transit analysis described in Section 2.3. E: Parameters come from Angelo et al. (2017). F: Equilibrium temperatures $T_{\rm eq}$ calculated assuming circular orbits, albedo α uniformly distributed between 0 and 0.7, and perfect heat redistribution. $T_{eq} = T_{\rm eff}(1-\alpha)^{1/4}\sqrt{\frac{R_\star}{2a}}$. G: Constrained by an informative prior on u_1 and u_2 based on model limb-darkening parameters (Claret & Bloemen 2011).

are given in Table 1 and our best-fit model is plotted in Figure 1.

3. Statistical Validation

Like most stars observed by Kepler, Kepler-1649 is too faint for radial velocity observations to confirm its planets' existence. Instead, we statistically show that Kepler-1649 c is almost certainly a genuine exoplanet. To do this, we use the `vespa` software (Morton 2015), which implements the methods described by Morton (2012) to calculate any given planet candidates's FP probability (FPP). `Vespa` uses knowledge of a candidate's orbital characteristics, transit light curve, host stellar parameters, location in the sky, and observational constraints to calculate the relative likelihood that the candidate signal is due to a transiting planet compared to several different FP scenarios.

We ran `vespa` on both Kepler-1649 c and the previously validated Kepler-1649 b. We input our well-determined stellar parameters, improved transit light curve, and the high-resolution imaging of Kepler-1649 obtained by Angelo et al. (2017). Like Angelo et al. (2017), we found that Kepler-1649 b is almost certainly a planet, with an FP probability of 2×10^{-5}. The results for Kepler-1649 c, on the other hand, were more ambiguous. Due to its slightly shorter transit duration and lower signal-to-noise, it is harder to distinguish the transit light curve of Kepler-1649 c from that of an extremely grazing eclipsing binary. `Vespa` reflects this uncertainty by calculating a ≈2% chance that Kepler-1649 c is an eclipsing binary. However, this calculation does not take into account the presence of a nearby validated planet. If Kepler-1649 b is indeed a planet orbiting Kepler-1649, then Kepler-1649 c cannot be an eclipsing binary also orbiting Kepler-1649 or the system would become dynamically unstable.[13] As Kepler-1649 c can only be an eclipsing binary if Kepler-1649 b does not orbit Kepler-1649, we multiply the prior for the eclipsing binary scenario in `vespa` by 1.6×10^{-5} (the probability that Kepler-1649 b does not orbit Kepler-1649). Taking this into account, we find that Kepler-1649 c's FP probability is about 2×10^{-3}, well below the thresholds typically applied to consider a planet statistically validated (e.g., Rowe et al. 2014; Morton et al. 2016). We get a similar result if we apply a "multiplicity boost" to Kepler-1649 c's FP probability, to reflect the fact that planets tend to be found in multi-transiting systems more often than FPs, which are more randomly distributed over the stars observed by Kepler. Applying a multiplicity boost of 10–20 for the DR25 catalog (Burke et al. 2019) yields a similar FP probability of about 1×10^{-3} to 2×10^{-3}.

`Vespa` evaluates almost all astrophysical FP scenarios, but does not consider FPs due to instrumental artifacts. Most Kepler instrumental FPs are for low signal-to-noise, long-period candidates with only a handful ($\lesssim$5) of transits (Thompson et al. 2018), though Burke et al. (2019) recently showed that there may be some instrumental false alarm contaminants at short periods (25–100 days) as well. The mechanism behind these short-period false alarms is unknown, but appears only to be significant at very low signal-to-noise ratios (S/N); the examples identified by Burke et al. (2019) all have S/N $\leqslant$ 8.1. Kepler-1649 c transited 39 times during the Kepler observations, so it does not fall in the main population

[13] Using the analytic stability criterion from Gladman (1993), we calculate that the total mass of Kepler-1649 b and c must be less than about 2 Jupiter masses.

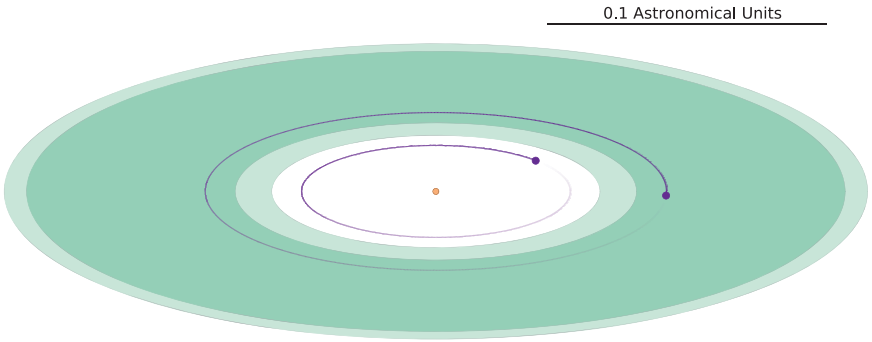

Figure 2. Diagram of the Kepler-1649 system, from the vantage of an observer inclined by 20 degrees from the plane of the system. The host star, Kepler-1649, is shown to scale in the center of the image, colored as a 3200 K blackbody would appear to the naked eye (see http://www.vendian.org/mncharity/dir3/blackbody/UnstableURLs/bbr_color.html). The optimistic and conservative habitable zone defined by Kopparapu et al. (2013) are colored in light and dark green, respectively. The orbits of Kepler-1649 b and c are shown as faded purple curves. The purple dots at the ends of the orbit curves denoting Kepler-1649 b and c are not to scale; the planets' true sizes would be about 4 times smaller than the orbit curve widths.

of long-period Kepler instrumental signals, and has significantly higher S/N (9.3 as calculated by the Kepler pipeline, and 11 in our improved light curve) than the Burke et al. (2019) short-period FPs. We conclude that Kepler-1649 c is not in a regime where instrumental FP signals are a serious concern and consider it validated as a *bona fide* exoplanet.

4. Discussion

4.1. Kepler-1649 c and the Habitable Zone

Kepler-1649 c is an Earth-sized planet orbiting within its host star's habitable zone (as calculated by Kopparapu et al. 2013, under conservative assumptions). Figure 2 shows a schematic of the Kepler-1649 system along with the location of the circumstellar habitable zone. Though Earth-sized habitable-zone planets are believed to be intrinsically common, they remain difficult to detect, and we only know of a handful today. Based on the NASA Exoplanet Archive Confirmed Planets table,[14] and using Gaia-based radii from Berger et al. (2018) for Kepler planets, only four transiting planets (TRAPPIST-1 e, f, and g, Gillon et al. 2017, and TOI 700 d, Gilbert et al. (2020; Rodriguez et al. 2020; Suissa et al. 2020) and three non-transiting (Proxima Centauri b, Anglada-Escudé et al. 2016, Teegarden's Star c, Zechmeister et al. 2019, and GJ 1061 d Dreizler et al. 2020) have radii smaller than 1.25 $R_\oplus$, or mass less than 2.0 $M_\oplus$, and orbit within their star's conservative habitable zone.[15]

In terms of size and incident bolometric flux, Kepler-1649 c is a near analog of Earth. Its radius ($1.06^{+0.15}_{-0.10} R_\oplus$) is consistent with that of Earth, and the planet receives about 75% of Earth's incident stellar flux. It seems likely, but not guaranteed, that Kepler-1649 c has a rocky composition—hot Earth-sized planets do tend to be rocky (Rogers 2015), but it may be unwise to extrapolate these results to cooler planets like Kepler-1649 c. Indeed, though some of Kepler-1649 c's bulk parameters are similar to Earth, the planet may not be at all "Earth-like." Many of Kepler-1649 c's properties remain uncertain, and planets orbiting M-dwarfs experience a very different environment (an extended era of ultraviolet (UV) irradiation, tidal locking, etc.) from the planets in our own solar system (Shields et al. 2016).

4.2. Kepler-1649 System Architecture

Like many other Kepler systems (Fabrycky et al. 2014), and especially mid M-dwarfs (Muirhead et al. 2015), Kepler-1649 hosts multiple close-in transiting planets. Generally, multi-transiting systems are nearly coplanar, with only a small spread in mutual inclinations (Fabrycky et al. 2014, though not always; see Mills & Fabrycky 2017). The Kepler-1649 system appears to fit this trend, as the inclinations of the two planets are consistent (at the 1σ level) with being perfectly coplanar.

The period ratio between Kepler-1649 b and c is 2.248250 ± 0.000013, only 0.08% inside of a 9/4 period ratio. Often, near-integer period ratios between neighboring planets indicate that the planets were or are in an orbital resonance, but the 9:4 resonance is weak; usually planet pairs are found near stronger resonances like the 2:1 or 3:2 ratios. We therefore suspect that there may be a third planet orbiting between Kepler-1649 b and c, forming a chain of 3:2 resonances. If we assume that Kepler-1649 b and c are in a three-body Laplace resonance with this hypothetical third planet, its period should be close to 13.029593 days (with an uncertainty of a few minutes). We checked the Kepler light curve for a planet with this orbital period, but found no transits deeper than about 600 ppm. Therefore, if there is a third planet orbiting between Kepler-1649 b and c, it is either too small to detect (roughly Mars-sized or less, though transit timing variations could hide a somewhat larger planet), or it is misaligned enough to prevent it from transiting.

4.3. The Frequency of Habitable Zone Earth-sized Planets around Mid M-dwarfs

Kepler-1649 c is the first habitable-zone Earth-sized planet to be found among the mid M-dwarf stars observed by Kepler. Only about 450 such stars were observed during Kepler's primary mission (Hardegree-Ullman et al. 2019), so it is somewhat surprising that Kepler detected a transiting Earth analog in this small sample. We quantified the likelihood that Kepler would find at least one Earth analog around a mid M-dwarf with a Monte Carlo simulation. We generated many ($\approx 10^4$) synthetic planet populations around 461 mid M-dwarfs (with spectral types M3-M5.5) observed by Kepler identified by Hardegree-Ullman et al. (2019), assuming planet occurrence

[14] https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=planets

[15] When published, Kepler-186 f's radius was estimated to be $1.11 \pm 0.14\ R_\oplus$ (Quintana et al. 2014), but Gaia data pushes the planet radius to $1.26 \pm 0.07\ R_\oplus$, just above our cutoff.

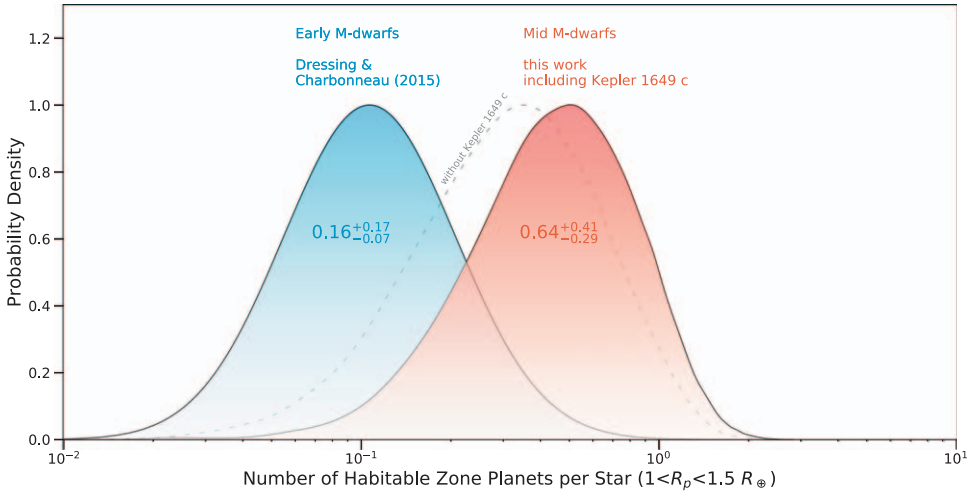

Figure 3. Posterior probability distributions on the occurrence rate of planets between 1 and 1.5 $R_\oplus$ in size in their stars' conservative habitable zones. We show both the occurrence rate for these planets around early M-dwarf stars (blue distribution) from Dressing & Charbonneau (2015), and mid M-dwarf stars (red distribution), including Kepler-1649 c in the calculation. We approximated the posterior distribution for early M-dwarf stars by finding (via least squares minimization) a log-normal distribution whose median and 68% confidence intervals most closely matched (within about 1%) those reported by Dressing & Charbonneau (2015). We also show the posterior probability distribution for mid M-dwarfs without including Kepler-1649 c as a gray dashed line. The addition of Kepler-1649 c slightly increases the measured occurrence rate of small, habitable-zone planets around mid M-dwarfs and adds new evidence suggesting that these planets may be more common than those around higher-mass stars.

statistics from Dressing & Charbonneau (2015). The planets had radii randomly drawn from a uniform distribution between 0.75 and 1.25 $R_\oplus$, orbital periods drawn from a log-uniform distribution inside the habitable zone, and inclinations drawn from a uniform distribution in $\cos(i)$. We calculated the expected transit signal-to-noise for each favorably inclined planet, and used the Kepler detection efficiency curve from Christiansen (2017) to determine which simulated planets were detectable by Kepler. We found only a 3.7% chance that Kepler would detect an Earth-sized planet in the conservative habitable zone around a mid M-dwarf.

Given the long, but not insurmountable odds against its detection, it is possible that Kepler-1649 c's discovery was just lucky chance, but it is also possible that the intrinsic occurrence rate of such objects we used is incorrect. Our previous calculation assumed an occurrence rate calculated by Dressing & Charbonneau (2015) for planets between 1.0 and 1.5 $R_\oplus$ orbiting a sample of mostly early M-dwarf stars. It is plausible that mid M-dwarf stars might form more Earth analogs than their more massive counterparts, thus boosting the chances we would detect such a planet in the Kepler sample. To test this, we calculated the occurrence rates of planets around mid M-dwarf stars observed by Kepler, while including the newly rescued Kepler-1649 c in our calculation. We modeled the population of planets orbiting mid M-dwarf stars (in particular, the sample defined by Hardegree-Ullman et al. 2019) with a joint (un-broken) power-law distribution in planet radius and orbital period (Burke et al. 2015; Bryson et al. 2020) and determined power-law parameters by exploring a Poisson likelihood function with MCMC. Other than the inclusion of Kepler-1649 c, the details of our calculation are identical to that of Bryson (2020).

We note that by including Kepler-1649 c in our calculations, we make an implicit assumption that the pipeline error leading to Kepler-1649 c's incorrect FP classification was rare and therefore not well represented in the Kepler DR25 vetting completeness experiments. This assumption seems reasonable —the root cause of Kepler-1649 c's incorrect FP disposition, the large quarter-to-quarter variations in the light curve's photometric scatter, is unusual in Kepler data, only affecting faint stars with high proper motions unknown to the Kepler pipeline. The quarterly variations in Kepler-1649's light curve are among the worst (95th percentile) of even the faint, high-proper-motion stars in the Kepler mid M-dwarf sample. A visual inspection of the TCEs and the DR25 injection/recovery results for other stars with high quarter-to-quarter variations in photometric scatter showed that no other PCs (real or simulated) were rejected on similar grounds to Kepler-1649 c. We therefore include Kepler-1649 c in our calculations with the caveat that if our assumption is incorrect, our occurrence rate may be biased slightly high.

We integrated the resulting power-law planet occurrence rates over the habitable zones for each star in our sample and calculated an average number of planets per star. Our result is shown in Figure 3, compared to the Dressing & Charbonneau (2015) occurrence rate for such planets around early M-dwarfs. We measure an occurrence rate higher than that of early M-dwarfs, though the statistical significance of this difference is not high. However, when taken together with previous suggestions of increased planet occurrence around mid M-dwarfs (Muirhead et al. 2012; Hardegree-Ullman et al. 2019) and the discovery of three Earth-sized planets in the conservative habitable zone of TRAPPIST-1 (Gillon et al. 2017), the detection of Kepler-1649 c provides further evidence that Earth analogs may be more common around mid M-dwarfs than higher-mass stars.

4.4. Human Inspection of Automatically Vetted Signals

Moving forward, automatic vetting of PCs can only become more important as data volume increases and classification techniques improve. There are likely hundreds of undiscovered planets left in K2 data (Dotson et al. 2019), and NASA's Transiting Exoplanet Survey Satellite (TESS) satellite produces more light curves and TCEs every month

than the entire 10 yr Kepler mission (Guerrero et al. 2019). Expecting humans to keep up with such vast quantities of data is unsustainable, and automatic vetting techniques have already taken the bulk of the triage/vetting workload (Yu et al. 2019; Guerrero et al. 2019). However, as our rescue of Kepler-1649 c reinforces, careful human inspection will remain valuable going forward. Even if inspecting each FP TCE is unfeasible, examining a small but strategically chosen sample[16] of targets could help improve automatic methods and enable new discoveries.

We thank Juliette Becker and Courtney Dressing for helpful conversations. A.V.'s work was performed under contract with the California Institute of Technology/Jet Propulsion Laboratory funded by NASA through the Sagan Fellowship Program executed by the NASA Exoplanet Science Institute.

This research has made use of NASA's Astrophysics Data System, 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, and the SIMBAD database, operated at CDS, Strasbourg, France.

This Letter includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. Some of the data presented in this Letter were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5–26555. Support for MAST for non-Hubble Space Telescope (HST) data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts. This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC; https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

Facility: Kepler.

Software: IDL Astronomy Library (Landsman 1993), `vespa` (Morton 2012, 2015), matplotlib (Hunter 2007), numpy (Oliphant 2006).

ORCID iDs

Andrew Vanderburg https://orcid.org/0000-0001-7246-5438
Pamela Rowden https://orcid.org/0000-0002-4829-7101
Steve Bryson https://orcid.org/0000-0003-0081-1797
Jeffrey Coughlin https://orcid.org/0000-0003-1634-9672
Natalie Batalha https://orcid.org/0000-0002-7030-9519
Karen A. Collins https://orcid.org/0000-0001-6588-9574
David W. Latham https://orcid.org/0000-0001-9911-7388
Susan E. Mullally https://orcid.org/0000-0001-7106-4683
Knicole D. Colón https://orcid.org/0000-0001-8020-7121
Chelsea X. Huang https://orcid.org/0000-0003-0918-7484
Samuel N. Quinn https://orcid.org/0000-0002-8964-8377

[16] Some possible subsets of FPs to examine manually could include those around the most likely stars to host detectable transiting planets (like nearby dwarf stars or known transiting planet hosts), those around stars with unusual properties that might trip up automatic classifiers (young stars, very low-mass M-dwarfs, or white dwarfs), and TCEs that just barely missed the cutoff to become PCs.

References

Angelo, I., Rowe, J. F., Howell, S. B., et al. 2017, AJ, 153, 162
Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Natur, 536, 437
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Bryson, S. 2020, RNAAS, 4, 32
Bryson, S. T., Abdul-Masih, M., Batalha, N., et al. 2017, Kepler Science Document, KSCI-19093-003
Bryson, S. T., Coughlin, J., Batalha, N. M., et al. 2020, AJ, in press (arXiv:1906.03575)
Bryson, S. T., Jenkins, J. M., Klaus, T. C., et al. 2010, Proc. SPIE, 7740, 77401D
Burke, C. J., Bryson, S. T., Mullally, F., et al. 2014, ApJS, 210, 19
Burke, C. J., & Catanzarite, J. 2017, Kepler Science Document, KSCI-19111-002
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
Burke, C. J., Mullally, F., Thompson, S. E., Coughlin, J. L., & Rowe, J. F. 2019, AJ, 157, 143
Christiansen, J. L. 2017, Kepler Science Document, KSCI-19110-001
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279
Claret, A., & Bloemen, S. 2011, A&A, 529, A75
Coughlin, J. L. 2017, Kepler Science Document, KSCI-19114-001
Coughlin, J. L., Mullally, F., Thompson, S. E., et al. 2016, ApJS, 224, 12
Dotson, J. L., Barentsen, G., Hedges, C., & Coughlin, J. L. 2019, RNAAS, 3, 23
Dreizler, S., Jeffers, S. V., Rodríguez, E., et al. 2020, MNRAS, 493, 536
Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gilbert, E. A., Barclay, T., Schlieder, J. E., et al., 2020 arXiv:2001.00952
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Natur, 542, 456
Gladman, B. 1993, Icar, 106, 247
Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65
Guerrero, N. M., Seager, S., Huang, C. X., et al. 2019, ApJS, submitted
Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, AJ, 158, 75
Henry, T. J., Jao, W.-C., Winters, J. G., et al. 2018, AJ, 155, 265
Howard, W. S., Tilley, M. A., Corbett, H., et al. 2018, ApJL, 860, L30
Hunter, J. D. 2007, CSE, 9, 90
Jenkins, J. M. 2017, Kepler Science Document, KSCI-19081-002
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, ApJL, 713, L87
Kipping, D. M. 2013, MNRAS, 435, 2152
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Kroupa, P. 2001, MNRAS, 322, 231
Landsman, W. B. 1993, in ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco, CA: ASP), 246
Luger, R., & Barnes, R. 2015, AsBio, 15, 119
Mandel, K., & Agol, E. 2002, ApJL, 580, L171
Mann, A. W., Dupuy, T., Kraus, A. L., et al. 2019, ApJ, 871, 63
Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, ApJ, 804, 64
Mills, S. M., & Fabrycky, D. C. 2017, AJ, 153, 45
Morton, T. D. 2012, ApJ, 761, 6
Morton, T. D. 2015, VESPA: False positive probabilities calculator, Astrophysics Source Code Library, ascl:1503.011
Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86
Muirhead, P. S., Johnson, J. A., Apps, K., et al. 2012, ApJ, 747, 144
Muirhead, P. S., Mann, A. W., Vanderburg, A., et al. 2015, ApJ, 801, 18
Mulders, G. D., Pascucci, I., & Apai, D. 2015, ApJ, 814, 130
Müller, H. M., Huber, K. F., Czesla, S., Wolter, U., & Schmitt, J. H. M. M. 2013, A&A, 560, A112
Oliphant, T. E. 2006, A Guide to NumPy (USA: Trelgol Publishing)
Quintana, E. V., Barclay, T., Raymond, S. N., et al. 2014, Sci, 344, 277
Rodriguez, J. E., Vanderburg, A., Zieba, S., et al. 2020, arXiv:2001.00954
Rogers, L. A. 2015, ApJ, 801, 41
Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, ApJ, 784, 45
Shields, A. L., Ballard, S., & Johnson, J. A. 2016, PhR, 663, 1

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Still, M., & Barclay, T. 2012, PyKE: Reduction and Analysis of Kepler Simple Aperture Photometry data, Astrophysics Source Code Library, ascl:1208.004
Suissa, G., Wolf, E. T., Kopparapu, R. K., et al. 2020, arXiv:2001.00955
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, AJ, 152, 158
van Eylen, V., Albrecht, S., Huang, X., et al. 2019, AJ, 157, 61
Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948
Vanderburg, A., Latham, D. W., Buchhave, L. A., et al. 2016, ApJS, 222, 14
Yu, L., Vanderburg, A., Huang, C., et al. 2019, AJ, 158, 25
Zechmeister, M., Dreizler, S., Ribas, I., et al. 2019, A&A, 627, A49