Properties and Occurrence Rates for *Kepler* Exoplanet Candidates as a Function of Host Star Metallicity from the DR25 Catalog

Mayank Narang[1], P. Manoj[1], E. Furlan[2], C. Mordasini[3], Thomas Henning[1,4,7], Blesson Mathew[5],
Ravinder K. Banyal[6], and T. Sivarani[6]

[1] Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India
mayank.narang@tifr.res.in
[2] IPAC, Mail Code 314-6, Caltech, 1200 E. California Blvd., Pasadena, CA 91125, USA
[3] Physikalisches Institut, Universität Bern, Gesellschaftstrasse 6, 3012 Bern, Switzerland
[4] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[5] Department of Physics, Christ University, Hosur Road, Bangalore 560029, India
[6] Indian Institute of Astrophysics, Bangalore 560034, India

Abstract

Correlations between the occurrence rate of exoplanets and their host star properties provide important clues about the planet formation process. We studied the dependence of the observed properties of exoplanets (radius, mass, and orbital period) as a function of their host star metallicity. We analyzed the planetary radii and orbital periods of over 2800 *Kepler* candidates from the latest *Kepler* data release, DR25 (Q1–Q17), with revised planetary radii based on *Gaia* DR2 as a function of host star metallicity (from the Q1–Q17 (DR25) stellar and planet catalog). With a much larger sample and improved radius measurements, we are able to reconfirm previous results in the literature. We show that the average metallicity of the host star increases as the radius of the planet increases. We demonstrate this by first calculating the average host star metallicity for different radius bins and then supplementing these results by calculating the occurrence rate as a function of planetary radius and host star metallicity. We find a similar trend between host star metallicity and planet mass: the average host star metallicity increases with increasing planet mass. This trend, however, reverses for masses $>4.0\,M_J$: host star metallicity drops with increasing planetary mass. We further examined the correlation between the host star metallicity and the orbital period of the planet. We find that for planets with orbital periods less than 10 days, the average metallicity of the host star is higher than that for planets with periods greater than 10 days.

Key words: methods: statistical – planets and satellites: formation – planets and satellites: general – stars: abundances – stars: fundamental parameters

1. Introduction

The launch of the *Kepler* mission has provided us with an unprecedented view of planetary systems around stars other than the Sun. *Kepler* has added significantly to the number of smaller (planetary radius $R_P \leqslant 4\,R_{\oplus}$) planets known to date (e.g., Borucki et al. 2010, 2011; Batalha et al. 2011, 2013; Lissauer et al. 2011; Howard et al. 2012; Fabrycky et al. 2014). The detection of more than 4000 *Kepler* planet candidates has ushered in an era of statistically significant studies of exoplanet properties (e.g., Howard et al. 2012; Dressing & Charbonneau 2013; Mann et al. 2013; Petigura et al. 2013; Morton & Swift 2014; Burke et al. 2015; Christiansen et al. 2015; Mulders et al. 2015a, 2015b; Wang & Fischer 2015; Fulton et al. 2017; Pascucci et al. 2018; Petigura et al. 2018).

Studies of exoplanet properties as a function of host star properties are particularly interesting, as they provide important clues to the formation of planetary systems. Planetary systems are formed out of protoplanetary disks surrounding young stars. These disks are the by-products of the star formation process, and both the star and the disk (as well as the planetary system) are formed out of the same molecular cloud. The disk properties are known to strongly correlate with the host star properties (e.g., McClure et al. 2010; Muzerolle et al. 2010; Andrews et al. 2011, 2013; Furlan et al. 2011; Manoj

et al. 2011; Kim et al. 2013, 2016; Mohanty et al. 2013; Pascucci et al. 2016).

The observed properties of exoplanets are also correlated with the host star properties. For example, the occurrence rate of exoplanets has a strong dependence on the spectral type of the host stars: the occurrence rate of giant planets ($R_P > 4$–$6\,R_{\oplus}$) is found to be higher for F, G, and K stars compared to that for M-type stars (Cumming et al. 2008; Howard et al. 2012; Bonfils et al. 2013; Gaidos et al. 2013; Mulders et al. 2015b; Winn & Fabrycky 2015; Obermeier et al. 2016; Mulders 2018; Winn 2018).

Radial velocity (RV) studies prior to *Kepler* have demonstrated a strong correlation between the occurrence rate of giant planets ($M_P > 0.5\,M_J$) and the metallicity of their host stars (Gonzalez 1997; Santos et al. 2001, 2004; Fischer & Valenti 2005; Johnson et al. 2007, 2010; Sousa et al. 2008, 2011; Sozzetti et al. 2009; Ghezzi et al. 2010; Mortier et al. 2012; Mulders 2018; Winn 2018). These studies found that as the metallicity of the host stars increases, the fraction/frequency of giant planets around them also increases. Further, Fischer & Valenti (2005) and Johnson et al. (2010) showed that the fraction of stars hosting gas-giant planets scales as a power-law function of the host star metallicity.

The host star metallicity correlation studies for smaller planets ($R_P \leqslant 4\,R_{\oplus}$) began only after the launch of the *Kepler* mission because such studies required a large number of smaller-planet detections (e.g., Buchhave et al. 2012, 2014;

---

[7] Adjunct Professor.

Adibekyan et al. 2013; Batalha et al. 2013; Fressin et al. 2013; Mann et al. 2013; Buchhave & Latham 2015; Mulders et al. 2016; Petigura et al. 2018). These studies showed that the average metallicity of the host stars of small planets ($R_P \leqslant 4 R_\oplus$) is lower than the average metallicity of the stars hosting giant planets. Further, the overall scatter in the distribution of the metallicity of the host stars was larger for those harboring small planets compared to those hosting giant planets, indicating that small planets can form around stars with a wide range of metallicities (e.g., Buchhave et al. 2012, 2014; Mann et al. 2013; Buchhave & Latham 2015; Schlaufman 2015; Wang & Fischer 2015; Mulders 2018; Petigura et al. 2018).

Studies have also shown the presence of a correlation between the metallicity of the host star and the orbital period of the planet around it. For planets orbiting with periods of 10 days or less, the host stars appear to be metal-rich and have a higher average metallicity than host stars with planets orbiting farther out (e.g., Adibekyan et al. 2013; Beaugé & Nesvorný 2013; Mulders et al. 2016; Maldonado et al. 2018; Wilson et al. 2018; Petigura et al. 2018).

In this paper, we investigate the correlations between host star metallicity and planetary radius and orbital period using the latest exoplanet data from *Kepler* Data Release 25 (DR25; Thompson et al. 2018). We supplement the DR25 catalog with improved planetary and stellar radius estimates based on *Gaia* DR2 from Berger et al. (2018).

An advantage of the large and uniform DR25 *Kepler* sample is that the observational biases and selection effects are well understood and can be corrected for by calculating the occurrence rate as a function of host star metallicity and planetary properties.

In Section 2, we describe our sample and the selection criteria used in compiling the sample. In Sections 3–5, we investigate how the host star metallicity is related to planetary radius and mass. In Section 6, we discuss the relationship between the orbital period of the planet and host star metallicity. In Section 7, we compare our results obtained using the metallicities from the DR25 stellar catalog (Mathur et al. 2017) with those of Petigura et al. (2018), who used more precise metallicities measured from high-resolution ($R \sim 60,000$) spectra of host stars. We show that our results are consistent with those of Petigura et al. (2018) and summarize them in Section 8.

2. Sample Selection

To ensure a uniform sample with well understood selection biases, we used the *Kepler* DR25 (Thompson et al. 2018; retrieved on 2018 March 18) from the NASA Exoplanet Archive[8] (Akeson et al. 2013). Detailed characterization of the host star properties is important to derive accurate planet properties and understand the planet population. Huber et al. (2014) and later Mathur et al. (2017, 2018) compiled a catalog of all the stars observed by the *Kepler* mission in Quarters 1–16 (Q1–Q16) and 1–17 (Q1–17; DR25), respectively. These catalogs compiled the host star properties (temperature, surface gravity, and metallicity) from various surveys. For the DR25 stellar catalog, Mathur et al. (2017) mainly used results from LAMOST (Luo et al. 2015), APOGEE (Majewski et al. 2015), the *Kepler* community follow-up program (CFOP; spectroscopy), and Kepler Input Catalog (KIC) photometry (Brown

---

8 https://exoplanetarchive.ipac.caltech.edu/
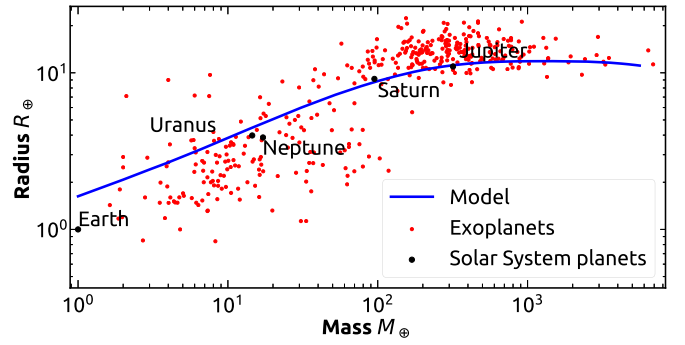
Figure 1. Masses and radii for 403 planets (red circles). Also plotted are the solar system planets (black circles). We also show the theoretical mass–radius curve (without inflation) from Mordasini et al. (2012b) in blue. The median uncertainty in the mass measurements of the exoplanets is about 11%, and the median uncertainty in the radius measurements is about 5%.

et al. 2011). The temperature, surface gravity, and metallicity were then homogeneously fitted to grids of Dartmouth stellar isochrones to derive stellar properties following Serenelli et al. (2013). Mathur et al. (2017) then reported the best-fit value conditioned on the isochrone fitting as the stellar parameters for the DR25 catalog. The fit to isochrones is done to derive the stellar radius and age; as a result of this fitting procedure, the effective temperature, surface gravity, and metallicity are changed somewhat from the input values, since they have to be consistent with the values of the isochrones. For our analysis, we replaced the stellar and planetary radii in DR25 with improved estimates based on *Gaia* DR2 from Berger et al. (2018). The stellar radii estimates from *Gaia* DR2 are a factor of 4–5 better than previous estimates, which translates into a similar improvement in the planetary radii (Berger et al. 2018).

In this paper, we focus on planet candidates (identified by the DR25 transit detection run) around main-sequence stars with spectral types F, G, K, and M with a T_{eff} range from 3200 K (M4) to 7200 K (F0). To ensure that we only pick main-sequence dwarfs, we restrict the sample to log g (in cgs units) values between 4 and 5. We also restrict the sample to an orbital period of 1–365 days.

On further examining our sample, we found that many *Kepler* planet candidates have radii $>20 R_\oplus$. The theoretical mass–radius relations derived from models without special inflation mechanisms suggest that the upper limit of planet radii should be close to about $12 R_\oplus$ (e.g., Bodenheimer et al. 2003; Fortney et al. 2007; Seager et al. 2007; Chabrier et al. 2009; Mordasini et al. 2012b, 2015; Swift et al. 2012). Planets with radii up to $\sim$17–19 $R_\oplus$ can be explained as inflated Jupiters (e.g., Fortney et al. 2007; Miller et al. 2009; Demory & Seager 2011; Hartman et al. 2011, 2016; Barros et al. 2016; Lillo-Box et al. 2016; Spake et al. 2016; Thorngren & Fortney 2018). Any object with a radius above 22–25 $R_\oplus$ is unlikely to be a planet.

In Figure 1, we show the mass–radius relationship for 403 exoplanets (red circles) for which both mass and radius are measured independently. These data are obtained from the confirmed planet table at the NASA Exoplanet Archive. The radii of the planets were measured via transit observations, and the masses of the planets were obtained from RV or transit-timing variation (TTV) measurements. In Figure 1, we also show the theoretical fit to the mass–radius relationship (5 Gyr) proposed by Mordasini et al. (2012b; blue curve). We clearly see that for Jovian-mass planets, the theoretical radius limit is
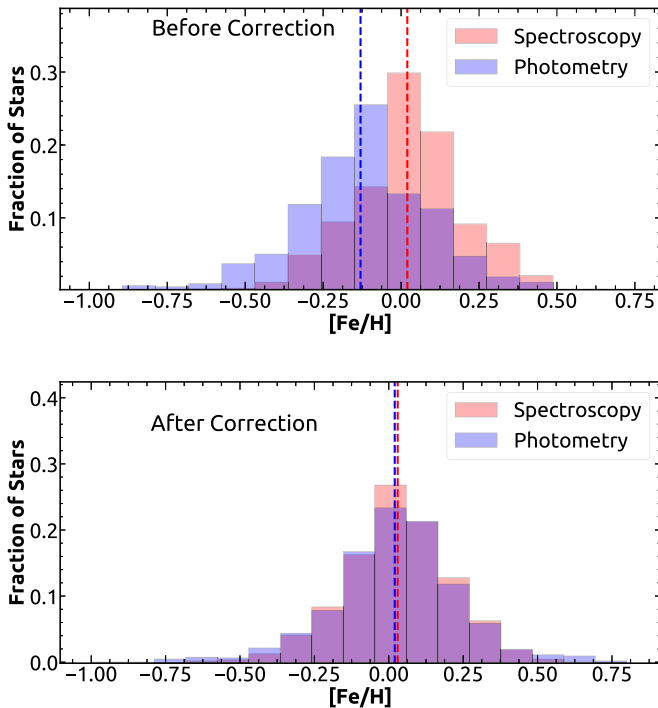
Figure 2. Distribution of spectroscopic and photometric host star metallicities from DR25 before (top panel) and after (bottom panel) correcting for the metallicity offset as described in the text. The dashed blue line represents the mean metallicity of the photometric subsample, and the red dashed line represents the mean metallicity of the spectroscopic subsample.

about $12\,R_\oplus$ and the measured radii are mostly less than about $20\,R_\oplus$. Therefore, in our analysis, we restricted ourselves to planets with radii in the range $1\,R_\oplus \leqslant R_P \leqslant 20\,R_\oplus$.

The stellar parameters derived from spectroscopy are relatively more accurate than those derived from photometry, in particular, the metallicity. In the DR25 sample, about 60% of the host stars of the planet candidates have spectroscopic metallicities. We compared the spectroscopically determined metallicities with those determined photometrically for the planet host stars for our DR25 sample. There appears to be an offset between the two, even though the host stars in both samples have similar spectral type and $\log g$ distribution. This is illustrated in Figure 2, where we show the distribution of spectroscopic and photometric metallicities for the planet host stars from the DR25 catalog. As can be seen from Figure 2, the mean of the spectroscopic and photometric metallicity distributions is offset by $\sim$0.15 dex.

In order to better characterize this offset and correct for it, we compared the DR25 host star metallicities with metallicities from the California Kepler Survey (CKS; Johnson et al. 2017; Petigura et al. 2017, 2018) and LAMOST Data Release 4 (DR4)[9] for sources that are common between DR25 and these two catalogs. We find that the offset between the DR25 spectroscopic metallicities and CKS metallicities (for 640 common sources) is $\sim$0.01 dex. The offset between the DR25 spectroscopic metallicities and LAMOST DR4 metallicities (for 482 common sources) is $\sim$0.03 dex. On the other hand, the DR25 photometric metallicities show offsets of 0.15 dex compared to CKS metallicities (168 common sources) and 0.14 dex compared to LAMOST DR4 metallicities (308 common sources). In order to make the DR25 spectroscopic

and photometric metallicities consistent with each other and with those in CKS and LAMOST, we applied a correction of 0.15 dex to the photometric metallicities and 0.01 dex to the spectroscopic metallicities listed in the DR25 catalog. The metallicity distribution of the DR25 host star sample after applying the correction described above is shown in Figure 2 (bottom panel). Both the spectroscopic and photometric metallicity distributions look very similar after applying the offset correction.

We further carried out a similar comparison of metallicities for the full DR25 stellar sample with metallicities listed in the LAMOST DR4 for sources common between the two. The offsets found between DR25 spectroscopic metallicities and LAMOST DR4 metallicities (for 4435 common sources) and DR25 photometric metallicities and LAMOST DR4 metallicities (for 27,551 common sources) are very similar to those found for planet host star metallicities. Therefore, we applied the same corrections to the spectroscopic and photometric metallicities for the larger DR25 stellar sample as well. We use these corrected metallicities in our analysis. In this analysis, we restrict ourselves to stellar metallicities between -0.8 and 0.5.

After applying all of these filters, our final sample contains a total of 2864 *Kepler* planet candidates around 2142 main-sequence F-, G-, K-, and M-type stars. This is the largest sample to date for which the dependence of observed planet properties (radius and orbital period) on host star metallicity has been studied.

3. Planet Radius and Host Star Metallicity

### 3.1. Planet Classification Based on Radius

Exoplanets come in various radii and masses. Based on their radii and masses, several classifications have been suggested in the literature (e.g., Borucki et al. 2011; Adibekyan et al. 2013; Fressin et al. 2013; Buchhave et al. 2014; Petigura et al. 2018). Fulton et al. (2017) showed that the occurrence rate of planets as a function of radius has a gap at $1.7\,R_\oplus$; however, the improved radius estimates based on *Gaia* DR2 from Berger et al. (2018) show that the gap is at $\sim 2\,R_\oplus$. Keeping this in mind, we classify planets in our sample as super-Earths ($1\,R_\oplus \leqslant R_P \leqslant 2\,R_\oplus$), Neptunes ($2\,R_\oplus < R_P \leqslant 4\,R_\oplus$), sub-Saturns ($4\,R_\oplus < R_P \leqslant 8\,R_\oplus$), and Jupiters ($8\,R_\oplus < R_P \leqslant 20\,R_\oplus$). This binning scheme is quite similar to that of Petigura et al. (2018), with two small differences: the cutoff point between super-Earths and Neptunes is set at $2\,R_\oplus$, and the upper limit of the planetary radius is set to $20\,R_\oplus$.

3.2. Planet Radius and Host Star Metallicity from the DR25 Stellar Catalog

Figure 3 shows the distribution of the planetary radius and host star metallicity for the DR25 sample. The planetary radii are from Berger et al. (2018), and the host star metallicities (corrected) are from the DR25 stellar catalog (Mathur et al. 2017, 2018), as described earlier. In Figure 4, we show host star average metallicity as a function of planet radius. We binned the data in the radius bins described in Section 3.1. Figure 4 shows that although there is large scatter in [Fe/H] in each radius bin, the average host star metallicity rises as the planetary radius increases, indicating that larger planets are preferentially found around host stars with higher metallicity.

Although we have applied the correction to the DR25 metallicities to make the spectroscopic and photometric
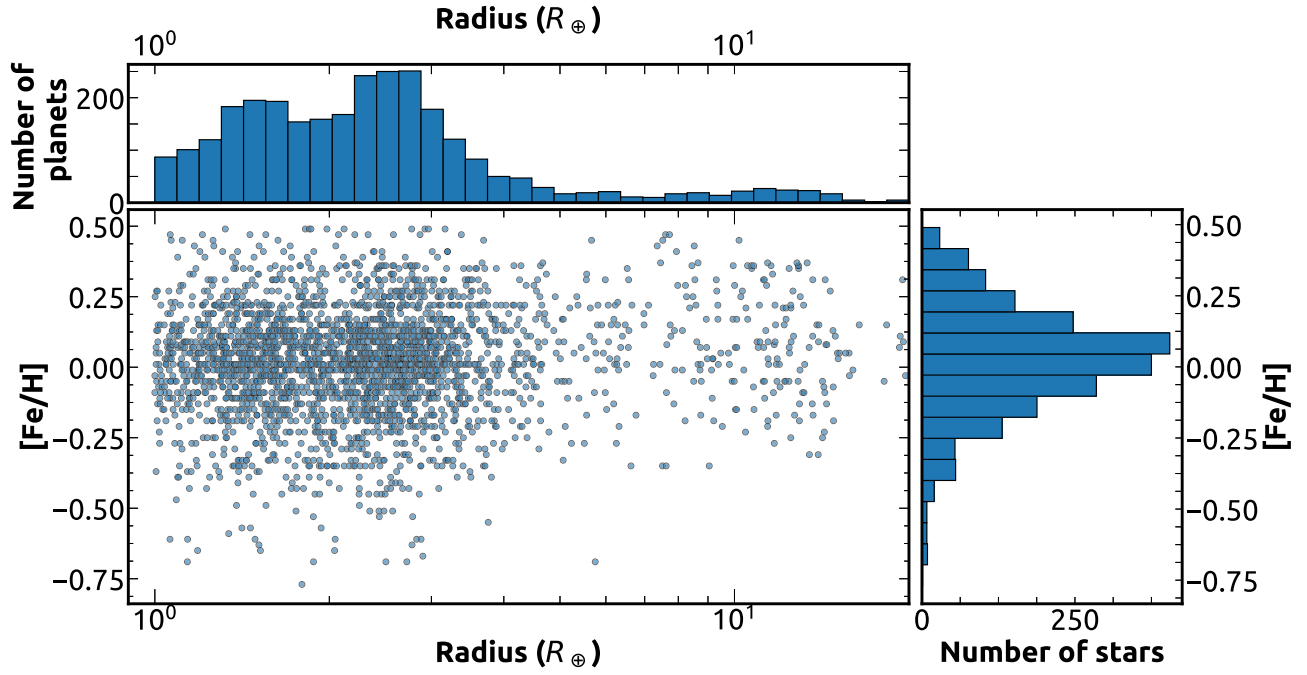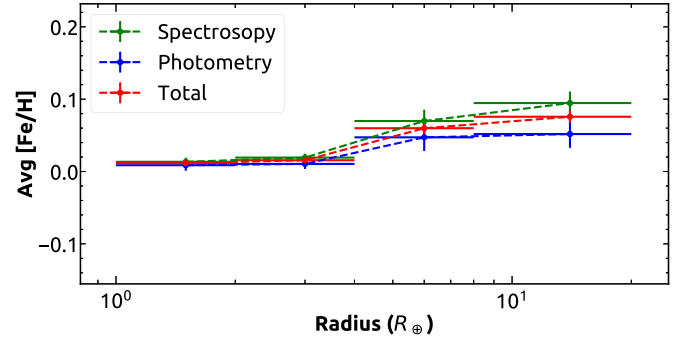

Figure 3. Scatter plot between the planetary radius and the host star metallicity for the *Kepler* DR25 sample. Also shown are the planet radius and host star metallicity distribution on the top and to the right. The host star metallicity histogram list each host star only once so as to avoid overcounting for multiplanetary systems (this convention is followed in all subsequent plots).
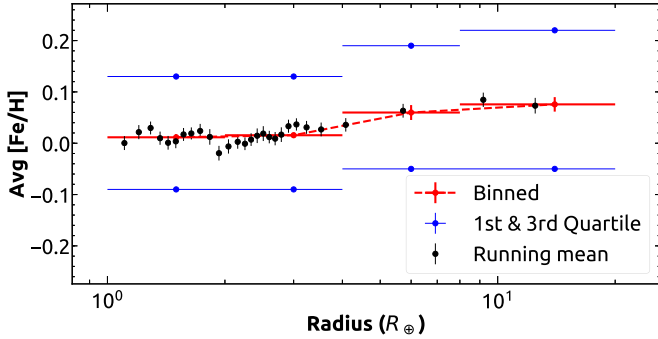
Figure 4. Average host star metallicity as a function of planet radius for the planetary candidate sample (2864 planet candidates in total). The red circles are the mean values of [Fe/H] and radius in each bin. The width of the radius bins is shown as red horizontal lines. The red vertical bars represent the standard deviation of the mean in [Fe/H], and the blue lines represent upper and lower quartiles for each bin. We also show a running mean of the host star metallicity as a function of the planetary radius computed for a box size of 200 with a step size of 100 as black circles, with the vertical (black) lines representing standard deviation of the mean in [Fe/H].

Figure 5. Average host star metallicity planet radius relation obtained after 100,000 random realizations of host star metallicities. The median value of the average metallicity of each bin for 100,000 realizations is shown as a function of planet radius. The standard deviation in the average metallicity of each bin is shown as the vertical error bar. The blue curve represents the average host star metallicities measured via photometry, the green curve represents host star metallicities measured via spectroscopy, and the red curve represents the total sample. The binning scheme is the same as in Figure 4.

metallicities consistent with each other, the photometric metallicities in the DR25 catalog have relatively large uncertainties (mean uncertainty $\sim$0.28 dex) compared to spectroscopically determined metallicities (mean uncertainty $\sim$0.14 dex). We next checked to see if these uncertainties affect the robustness of the statistical results that we find in Figure 4. To do so, we created 100,000 realizations of the metallicity distributions of the DR25 planetary sample with the metallicity (corrected) of each candidate host star chosen randomly such that the metallicity lies within the uncertainty of the measurement. Then we binned the data in radius bins exactly the same way as was done in Figure 4. The average metallicity in each bin was also calculated separately for host stars with spectroscopic and photometric metallicities. We then plot the median value of the average metallicity of

each bin for the 100,000 realizations as a function of planet radius. This is displayed in Figure 5. As can be seen from the figure, both the spectroscopic and photometric samples show a similar trend with planet radius. Also, for the total simulated sample, the average metallicity exhibits almost the same behavior as that in Figure 4. This indicates that even though the uncertainties in individual metallicities could be relatively high, particularly for photometric metallicities, the overall statistical correlation that we find between average host star metallicity and the planetary radius is robust.

3.3. Planet Radius and Host Star Metallicity Relationship from the CKS

To make use of the more precisely determined host star properties, we repeated the analysis using the data from the
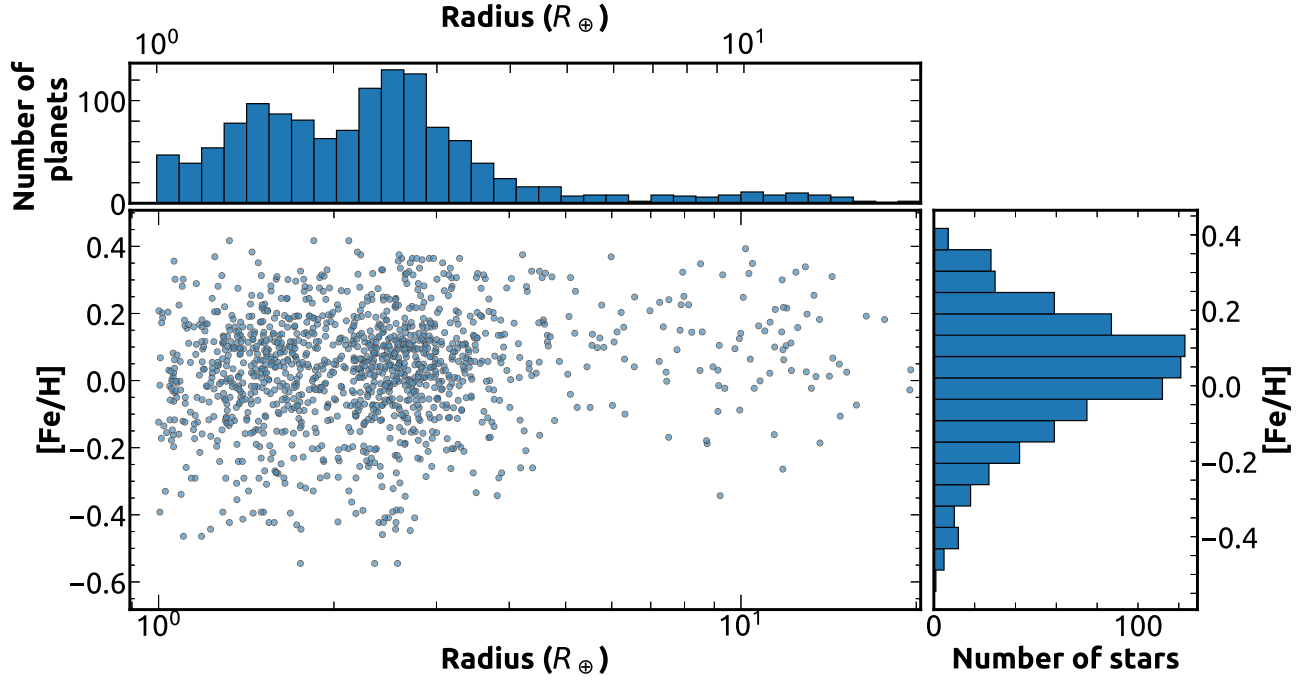
Figure 6. Scatter plot between planetary radius and host star metallicity for the CKS sample. See Figure 3 for a description of the plot.

CKS. The stellar properties for the CKS have been determined using high-resolution spectroscopy (Johnson et al. 2017; Petigura et al. 2017).

We examined the correlation between host star metallicities from the CKS and planet radii estimates from Berger et al. (2018; in order to maintain consistency). Using *Kepler* DR25, we updated the disposition status of all the planets in the CKS sample and only selected planets that were classified as candidates based on DR25. We then applied the same filters to the CKS sample as those used for our DR25 sample. The $T_{\rm eff}$ range is from 3200 to 7200 K, $\log g$ between 4 and 5, and planetary radius between 1 and 20 $R_\oplus$ (see Section 2). After applying these filters, there were a total of 1317 *Kepler* candidates with CKS stellar parameters. Petigura et al. (2018) had a magnitude-limited sample. We do not put any constraints on the magnitude of the host star. We take a much wider $T_{\rm eff}$ range and do not put any constraints on the impact factor of the planet. We also do not exclude planets with host stars that have a nearby stellar companion. Figure 6 shows the distribution of planetary radius and host star metallicity for the CKS sample.

In Figure 7, we show host star metallicity as a function of planet radius for the CKS sample. Here again, we binned the data into various radius bins as described in Section 3.1. We follow the same convention for plotting as in Figure 4. For the running mean, we used a box size of 100 with a step size of 50. The results from Figure 7 are consistent with those of Figure 4 and also with the results presented in Petigura et al. (2018).

4. Planet Mass and Host Star Metallicity

Studies prior to *Kepler* showed that giant planets are preferentially found around metal-rich host stars (e.g., Gonzalez 1997; Santos et al. 2001; Fischer & Valenti 2005; Johnson et al. 2010). Recently, Santos et al. (2017) found that, on average, the host star metallicity of more massive giant planets ($M_{\rm P} > 4\,M_{\rm J}$) is lower than the host star metallicity for
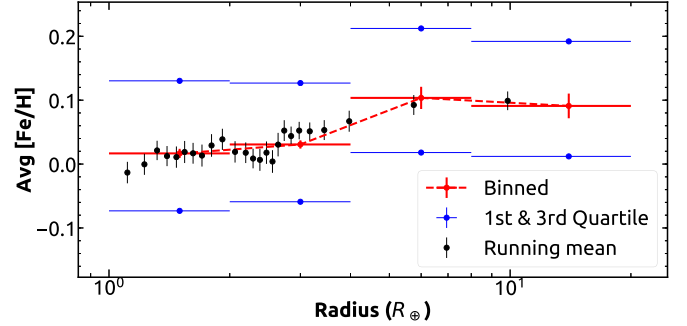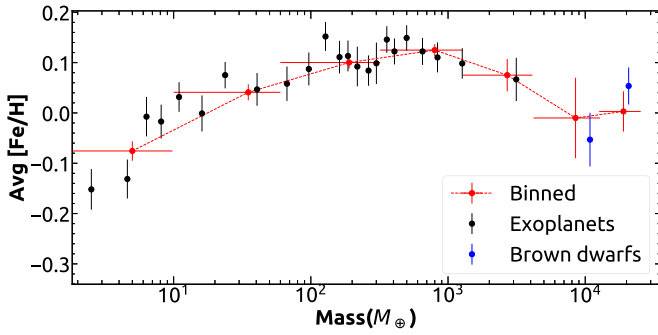

Figure 7. Average host star metallicity as a function of planet radius for the 1317 planets selected in the same manner as the DR25 sample, using stellar metallicities from the CKS. The symbols and lines have the same meaning as in Figure 4.

less massive giant planets with ($M_{\rm P} \leqslant 4\,M_{\rm J}$) (also see Schlaufman 2018).

To study the relationship between host star metallicity and planetary mass, we retrieved data from the confirmed planets table from the NASA Exoplanet Archive (retrieved on 2018 May 11). The catalog lists upper limits for some of the planetary masses, so we searched the literature for better mass estimates and replaced them. When only the upper limits were available, we did not include the planet in the sample. Planets with RV and inclination angle measurements allow us to derive the true mass (e.g., Batalha et al. 2011; Marcy et al. 2014; Gettel et al. 2016). If the inclination angle was not available, $M\sin i$, the minimum mass, was used as a proxy of planet mass. We also included planets with masses determined from TTV studies (e.g., Weiss et al. 2013; Hadden & Lithwick 2014) to increase the sample size.

The stellar parameters $T_{\rm eff}$, $\log g$, and metallicity were taken from the SWEET-Cat catalog (Santos et al. 2013). The SWEET-Cat catalog derives the host star properties from high-resolution spectra. Care has been taken while maintaining

Figure 8. Average host star metallicity as a function of planet/brown dwarf mass. The average host star metallicity for various (companion) mass bins is shown in red, with the horizontal red lines representing the bin width and the vertical red lines indicating the standard deviation of the mean. We also show a running mean of the host star metallicity as a function of the planetary mass computed for a box size of 60 with a step size of 30 as black circles, with the vertical black lines representing the standard deviation of the mean in [Fe/H]. We also show the average metallicities for host stars with brown dwarf companions (blue circles), with vertical blue lines representing the standard deviation of the mean in the [Fe/H] of the host stars of brown dwarfs.

and updating this catalog to make sure that the determination of stellar properties is as uniform as possible. The stellar parameters listed in the NASA Exoplanet Archive are compiled from various different sources. We, however, wanted to use a uniform set of metallicity values for our analysis; hence, we opted for the SWEET-Cat catalog for the stellar parameters. We cross-matched the positions of the host stars from the confirmed planet table from the NASA Exoplanet Archive to those listed in the SWEET-Cat catalog with a search radius of 30″. We then applied the same filters as those used for the *Kepler* DR25 sample: a host star $T_{\rm eff}$ range of 3200–7200 K and log g between 4 and 5 and the orbital period of the planet between 1 and 365 days. After applying these filters, we had a total of 705 confirmed planets with masses listed in the catalog. Out of these, 175 planets are in the *Kepler* field, and 109 have TTV measurements.

In Figure 8, we show the average host star metallicity as a function of planetary mass. We include binary brown dwarfs along with the metallicities of their host stars in Figure 8. The brown dwarf data is from Ma & Ge (2014). We applied the same filters that we applied for the *Kepler* sample. The average host star metallicity for planets/brown dwarfs in various mass bins (1–10 $M_\oplus$, 10–50 $M_\oplus$, 50 $M_\oplus$–1 $M_{\rm J}$, 1–4 $M_{\rm J}$, 4–13 $M_{\rm J}$, 13–35 $M_{\rm J}$, and 35–80 $M_{\rm J}$) is shown by red circles. The brown dwarfs are shown separately as blue circles, with the blue vertical bars being the standard deviation in the mean of the [Fe/H] of their host star.

As in the case for the radius–metallicity relation, the average host star metallicity increases as the mass of the planet increases until about 4 $M_{\rm J}$ (∼1200 $M_\oplus$). For planet masses >4 $M_{\rm J}$, the trend appears to reverse: the host star metallicity begins to drop as the mass of the planet increases. Figure 8 clearly shows that the drop in the metallicity of the host star as the companion mass increases extends into the brown dwarf regime as well. Similar results have also been reported by Santos et al. (2017) and Schlaufman (2018) for a sample of binaries with giant planets, brown dwarfs, and low-mass stars as secondary companions. These results suggest a similar mechanism for the formation of super-Jupiter (>4 $M_{\rm J}$), brown dwarf, and low-mass star binary systems.

From the simulations, it has been shown that for Jupiters (1–4 $M_{\rm J}$), [Fe/H] is not the only parameter that determines the final planet mass. The [Fe/H] determines whether or not a giant planet can form, but not the mass (Mordasini et al. 2012a). The mass is correlated with the disk gas mass, not [Fe/H]. For very massive giant planets (⩾10 $M_{\rm J}$), this is no more true: in order to become a very massive giant planet, the critical core mass must form very fast, before the gas in the disk is dissipated by accretion onto the star or photo-evaporation (Mordasini et al. 2012a). This is only possible at high [Fe/H]. In fact, Mordasini et al. (2012a), based on their theoretical planet population models, found that the most massive giant planets (⩾10 $M_{\rm J}$) only form at high [Fe/H]. This prediction of the core accretion model is quite contrary to what we find. This may indicate another formation mechanism, e.g., gravitational instability, for the formation of super-Jupiters ($M_{\rm P} > 4\,M_{\rm J}$).

5. Occurrence Rate

The analysis described in the previous sections does not take the completeness of the survey or the detector efficiency into account. The true trend might not be the one that simple binning or a running average shows. In order to derive the correlation between host star metallicity and planet size that is free of observational biases and other selection effects, we use the *Kepler* DR25 catalog to calculate the occurrence rate of exoplanets as a function of radius and metallicity. We calculate the occurrence rates using the methods described extensively in the literature (e.g., Youdin 2011; Howard et al. 2012; Dressing & Charbonneau 2013; Fressin et al. 2013; Morton & Swift 2014; Burke et al. 2015; Mulders et al. 2015a, 2016).

5.1. Number of Stars with Detectable Planets

Whether we can detect a planet around a given star depends on various factors: the geometric probability of the planet transiting the host star along our line of sight; the signal-to-noise ratio (S/N) that the transit would produce, which, in turn, is a factor of the transit depth, $\sigma_{\rm CDPP}$; and the number of transits. In the following subsections, we will describe how these quantities are calculated.

5.1.1. Transit Probability

The transit probability is defined as the geometric probability that we will be able to observe a given transit of a planet around a star of radius R_* at an orbital distance or semimajor axis a.

Using simple geometry, one can derive the transit probability $\eta_{\rm tr}$ as

$$\eta_{\rm tr} = \frac{R_* + R_{\rm P}}{a}\frac{1}{1 - e^2} \approx \frac{R_*}{a(1 - e^2)}.$$

Here R_* is the radius of the host star, $R_{\rm P}$ is the radius of the planet, e is the eccentricity, and a is the semimajor axis of the planet. Here we assume negligible eccentricity (i.e., circular orbits) and further approximative $\eta_{\rm tr}$ as

$$\eta_{\rm tr} = \frac{R_*}{a(1 - e^2)} \approx \frac{R_*}{a}. \tag{1}$$

Since a, the semimajor axis, is not an observable, we convert a into the orbital period of the planet around the host star using Kepler's third law and stellar mass M_* taken from the stellar

catalog,

$$\eta_{\text{tr}} \approx \frac{R_*}{a} = R_* \left(\frac{4\pi^2}{GM_*P^2} \right)^{1/3}. \tag{2}$$

5.1.2. Multiple Event Statistic

The multiple event statistic (MES) is a measure of how reliable the transit detection is. It is a measure of how noisy the signal would be given that our planet of radius R_P orbits a star with radius R_* with a period of P days, provided the star has a noise level $\sigma_{\text{CDPP}(\tau)}$ (combined differential photometric precision interpolated to transit duration). Here σ_{CDPP} can be described as the effective white noise that a transit signal would see given that the signal has a duration of τ. A σ_{CDPP} of 30 ppm for a 1 hr transit indicates that a 1 hr transit of depth 30 ppm would, on average, have an S/N of 1 (Christiansen et al. 2012). The S/N depends on the transit depth, transit duration, $\sigma_{\text{CDPP}(\tau)}$, and number of transits.

The transit depth δ is the fraction of flux blocked by the planet ΔF to the total flux F emitted by the star and is given by

$$\frac{\Delta F}{F} = \delta = \frac{\text{Area of planet}}{\text{Area of star}} = \frac{R_P^2}{R_*^2}. \tag{3}$$

Transit duration, the time duration in which any part of the planet obscures the disk of the star, is given by

$$\tau_{\text{dur}} = 6 \left(\frac{P}{1 \text{ day}} \right) \left(\frac{R_*}{a} \right) \sqrt{1 - e^2} \text{ hr}$$

$$\approx 6 \left(\frac{P}{1 \text{ day}} \right) \left(\frac{R_*}{a} \right) \text{hr}. \tag{4}$$

Since σ_{CDPP} is only available as an array of time values for [1.5, 2.0, 2.5, 3.0, 3.5, 4.5, 5.0, 6.0, 7.5, 9.0, 10.5, 12.0, 12.5, 15.0] hr, we use Equation (4) (Burke et al. 2015) to calculate the transit duration and interpolate the value of σ_{CDPP} to the transit duration to produce the effective white noise on the detector for the transit duration. To calculate the overall MES for the transit, we also need the total number of transits observed by *Kepler* during the total quarters observed. But not all the stars were observed for the whole *Kepler* mission duration. Due to the movement of the spacecraft, sometimes some of the stars were not observed, as the signal from the stars did not fall on the CCDs. The duty cycle of the star is the ratio of the amount of time the star was observed by the total time period of the *Kepler* mission. We restrict the sample to stars and Kepler Object of Interests (KOIs) with duty cycles greater than 60% and the star's data span on the CCD of at least 2 yr. Due to these restrictions, we calculated the occurrence rate for 2775 *Kepler* candidates around 2080 main-sequence stars.

The total number of transits N_{tr} observed by *Kepler* during the total quarters observed is given by

$$N_{\text{tr}} = \frac{\text{total quarters observed} \ast \text{ duty cycle}}{\text{orbital period}}. \tag{5}$$

We then calculate MES by averaging the transit signal strength over multiple transit events,

$$\text{MES} = \frac{\delta}{\sigma_{\text{CDPP}(\tau)}} \sqrt{N_{\text{tr}}}. \tag{6}$$

5.2. Calculating the Occurrence Rate

In order to investigate how the planetary radius and orbital period of the planets depend on the host star metallicity, we calculate the planet occurrence rate for three host star metallicity bins: sub-solar ($-0.8 \leqslant [\text{Fe/H}] < -0.2$), solar ($-0.2 \leqslant [\text{Fe/H}] \leqslant 0.2$), and super-solar ($[\text{Fe/H}] > 0.2$; following Beaugé & Nesvorný 2013).

The occurrence rate for planets within a radius bin R_0 and period bin P_0 is given as (Hsu et al. 2018)

$$F_{R_0 P_0} = \Sigma_{i \, R_0 P_0} C_i / N_{\text{targ}}, \tag{7}$$

where N_{targ} is the total number of stars in that metallicity bin and C_i is the estimate of the number of planets with radius R_0 and period P_0. To account for incompleteness due to the geometric probability of transit and detection efficiency, C_i is defined as

$$C_i = \frac{1}{\eta_{\text{tr}} \eta_{\text{det}}}, \tag{8}$$

where η_{det} accounts for the fraction of stars around which a planet with a radius R_0 and period P_0 can be detected. Here η_{det} times N_{targ} is the effective number of stars around which we can detect such a planet,

$$\eta_{\text{det}} = \Sigma_{j=1}^{N_{\text{targ}}} \eta_{R_0, P_0, j} / N_{\text{targ}}, \tag{9}$$

where $\eta_{R_0, P_0, j}$ is the probability that we can detect a planet with radius R_0 and period P_0 around the jth target star. Following Burke et al. (2015), the detection probability from the *Kepler* pipeline is a function of pipeline completeness and window function.

The modeled pipeline detection probability (from the recovery of injected transit signals) is given by

$$P_{\text{gamma}}(x|a, b, c) = \frac{c}{b^a \Gamma(a)} \int_0^x t^{a-1} e^{-t/b} dt, \tag{10}$$

where Γ is the gamma function; $a = 30.87$, $b = 0.271$, and $c = 0.940$ (for pipeline version 9.3 (DR25)); and x is the expected MES (Christiansen 2017).

The window function P_{win} accounts for the probability that a requisite number of transits required for detection occurs (Burke et al. 2015),

$$P_{\text{win}} = 1 - (1 - f_{\text{duty}})^M - M f_{\text{duty}} (1 - f_{\text{duty}})^{M-1}$$

$$- \frac{M(M-1)}{2} f_{\text{duty}}^2 (1 - f_{\text{duty}})^{M-2}, \tag{11}$$

where $M = T_{\text{obs}} / P_{\text{orb}}$, with T_{obs} being the data span of the star on the detector and P_{orb} being the orbital period of the planets. Here f_{duty} is the duty cycle of the star.

The detection probability around the jth target star is then given as

$$\eta_{R_0, P_0, j} = P_{\text{win}} \ast P_{\text{gamma}}. \tag{12}$$
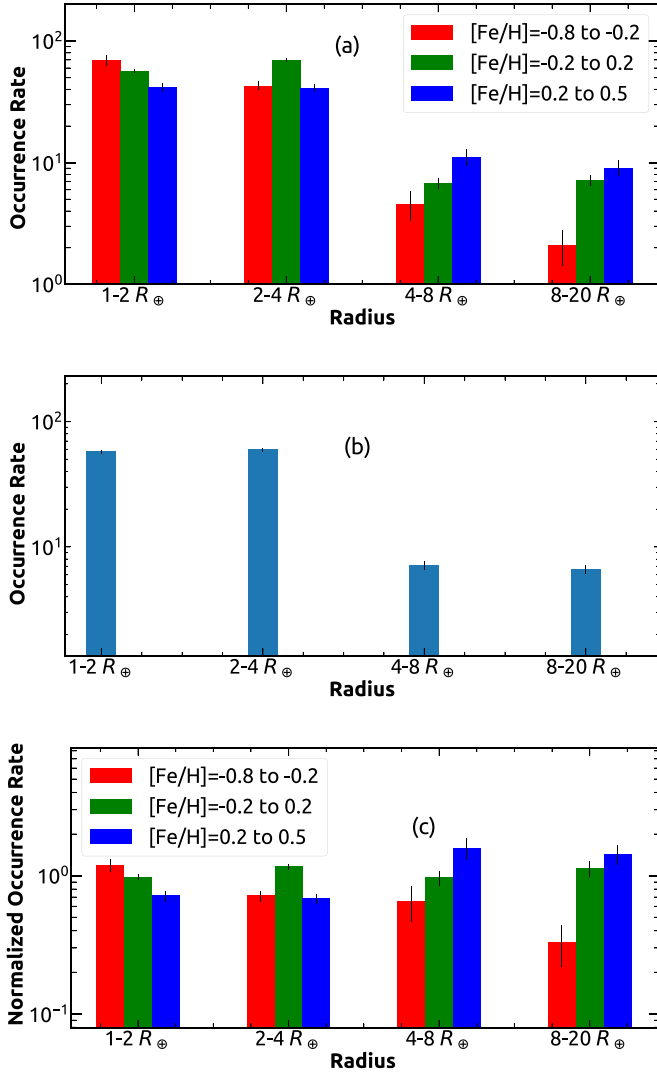
Figure 9. (a) Occurrence rate of exoplanets as a function of planetary radius and host star metallicity. (b) Total occurrence rate of the sample without subdividing it into different metallicity bins. (c) Normalized occurrence rate of exoplanets as a function of planetary radius and host star metallicity. The error bars in these plots are the Poissonian errors based on the number of planets in each bin.

Equation (7) takes into account all the observational biases and gives us a true measure of the occurrence rate. To get the occurrence rate as a function of radius and host star metallicity, we simply sum over the period

$$F|_{R_p=R_0} = \Sigma_i F_{R_0 P_i}.$$

To get the occurrence rate as a function of period and host star metallicity, we simply sum over the radius

$$F|_{P=P_0} = \Sigma_i F_{R_i P_0}.$$

5.3. Occurrence Rate of Exoplanets as a Function of Planet Radius and Host Star Metallicity

Using the above mathematical prescription, we calculated the occurrence rate (per 100 stars) as a function of host star metallicity and planetary radius for the DR25 sample (Figure 9(a)). We also calculated the total occurrence rate as a function of planet radius only for the DR25 sample (Figure 9(b)). Figure 9(a) is a function of both the planetary radius and host star metallicity. In Figure 9(b), we have removed the dependence on host star metallicity and calculated the occurrence rate only as a function of planetary radius. Figures 9(a) and (b) show that the occurrence rate is a much stronger function of the planetary radius than the host star metallicity.

To remove the underlying dependence of the occurrence rate of planets on the planetary radius, we normalize the occurrence rate in Figure 9(a) with the total occurrence rate (Figure 9(b)) to get the normalized occurrence rate in Figure 9(c). The normalized occurrence rate should only be a function of host star metallicity.

In Figure 9(c), the normalized occurrence rate of exoplanets is shown for various host star metallicities. This occurrence rate for super-solar-metallicity host stars (blue bin) increases as a function of planetary radius. A reverse trend is seen for the normalized occurrence rate of planets around sub-solar-metallicity host stars (red bin), where the occurrence rate decreases as a function of planetary radius. The normalized occurrence rate of planets around solar-metallicity host stars (green bin) is mostly constant. Figure 9(c) shows that the host stars with super-solar metallicities have higher occurrence rates for giant planets than metal-poor (solar- and sub-solar-metallicity) host stars. These results are consistent with those of Figures 4, 7, 8, and 9, which indicate that giant planets preferentially form around metal-rich stars. These results are also consistent with previous results in the literature (e.g., Gonzalez 1997; Santos et al. 2001, 2004; Fischer & Valenti 2005; Johnson et al. 2007; Buchhave et al. 2012, 2014; Mann et al. 2013; Buchhave & Latham 2015) and those of Petigura et al. (2018).

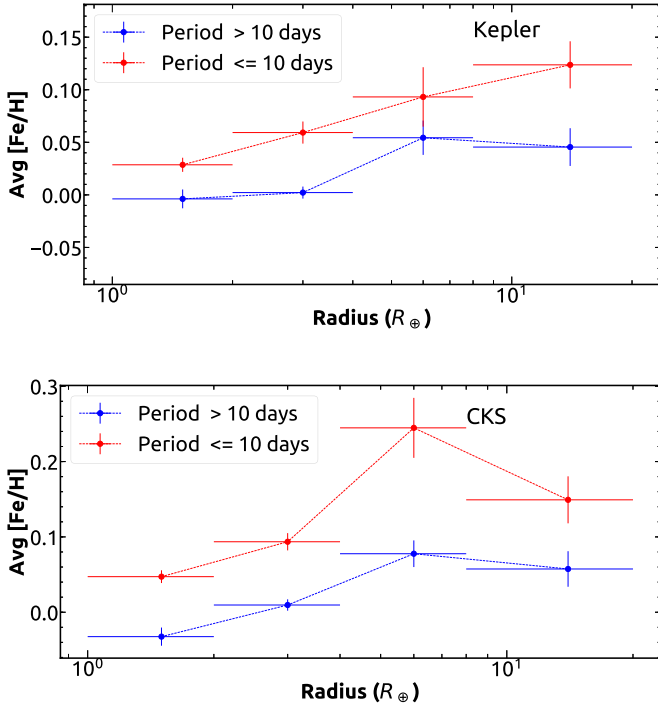## 6. Host Star Metallicity and the Planet's Orbital Period

Previous studies on the orbital period of exoplanets and host star metallicity have revealed a dearth of planets with orbital periods less than 10 days around metal-poor stars (e.g., Adibekyan et al. 2013; Beaugé & Nesvorný 2013; Mulders et al. 2016; Maldonado et al. 2018; Wilson et al. 2018; Petigura et al. 2018). Host stars of planets with orbital periods greater than 10 days were more metal-poor than their short-period counterparts. Mulders et al. (2016) and Petigura et al. (2018) showed that the occurrence rate of planets with periods less than 10 days was higher around metal-rich host stars than metal-poor host stars.

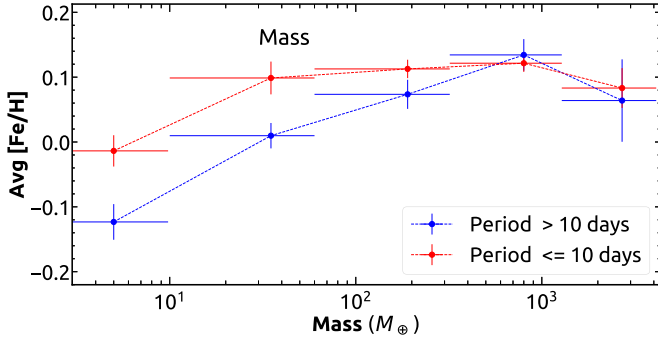### 6.1. Average Host Star Metallicity and Orbital Period

In Figure 10, we show the average host star metallicity as a function of planetary radius for two bins of the orbital period. In Figure 10(a), we divide the *Kepler* planet candidate sample into two subgroups: one having a period less than 10 days (red) and one having a period greater than 10 days (blue). Just like Figure 4, we bin the planet candidates into various radius bins following Section 3.1. In Figure 10(b), we follow the same prescription for the CKS sample. From Figure 10, we find that for planets with periods less than 10 days, the host stars are usually more metal-rich than their longer-period counterparts.

In Figure 11, we use the sample from Section 4 and show the average host star metallicity from SWEET-Cat as a function of the period for planets in various mass bins. We again divide the planet sample into two subgroups: one having periods less than 10 days (red) and one having periods greater than 10 days (blue)

Figure 10. Average host star metallicity as a function of the planetary radius for two different period bins. (a) Planetary radius and host star metallicity from the *Kepler* DR25 sample. (b) Planetary radius and host star metallicity from the CKS sample. The error bars for the *X*-axis are the width of the radius bins, and for the *Y*-axis, we show the standard error in the mean of [Fe/H] as the error bar.
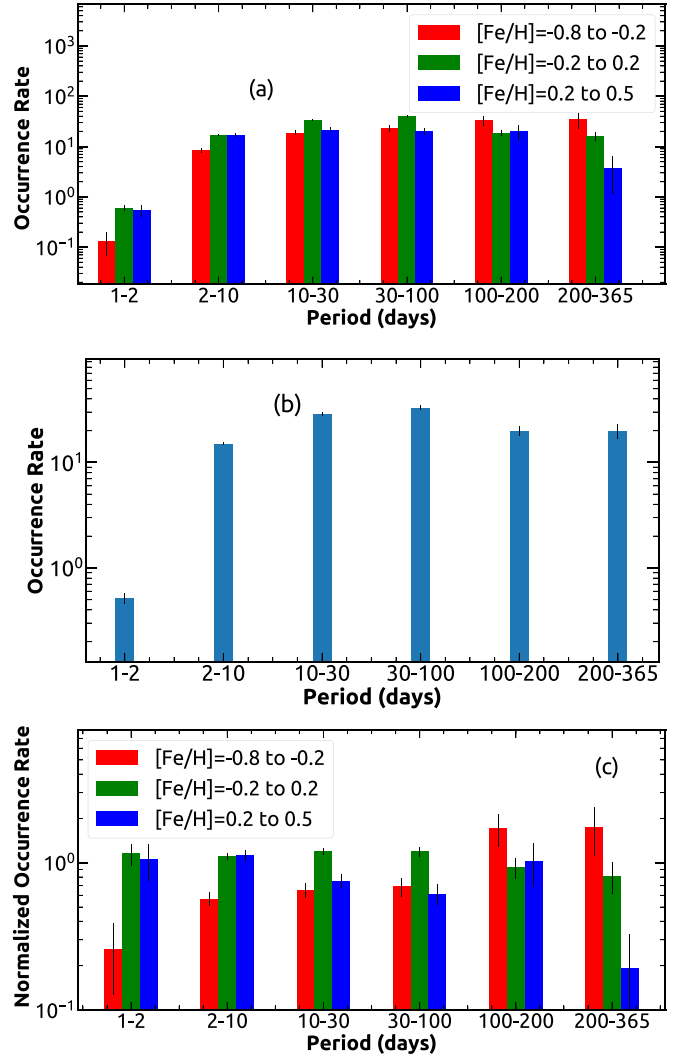

Figure 11. Average host star metallicity as a function of planetary mass. The planetary mass is from the NASA Exoplanet Catalog, and the host star metallicity is from the SWEET-Cat catalog. The error bars for the *X*-axis are the width of the mass bins, and for the *Y*-axis, we show the standard error in the mean of [Fe/H] as the error bar.

(blue). We find that for planets with masses up to $50\,M_\oplus$ and periods less than 10 days, the host stars are usually more metal-rich than their longer-period counterparts, but no such trend is found for giant planets.

6.2. Occurrence Rate of Exoplanets as a Function of Orbital Period and Host Star Metallicity

In this section, we investigate the occurrence rate of exoplanets as a function of host star metallicity and orbital period for planets in two radius regimes: planets with $R_P \leqslant 4\,R_\oplus$ (Figure 12) and planets with R_P between 4 and $20\,R_\oplus$ (Figure 13). The occurrence rate as a function of host star metallicity and the orbital period of the planet is shown in Figures 12(a) and 13(a), and the total occurrence rate as a

Figure 12. (a) Occurrence rate of exoplanets as a function of orbital period and host star metallicity for planets having radii less than $4\,R_\oplus$. (b) Total occurrence rate of the sample without subdividing it into different metallicity bins. (c) Normalized occurrence rate of exoplanets as a function of orbital period and host star metallicity. The error bars in these plots are the Poissonian errors from the counting of planets.

function of orbital period is shown in Figures 12(b) and 13(b). From Figures 12(a) and 13(a), we can infer that the occurrence rate is a much stronger function of the orbital period than the host star metallicity. The occurrence rate as a function of period increases as the orbital period increases and plateaus around an orbital period of 10–30 days. To remove the underlying effect of orbital period on the occurrence rate, we normalize the occurrence rate in Figures 12(a) and 13(a) with the total occurrence rate as a function of period (Figures 12(b) and 13(b)) to calculate the normalized occurrence rate (Figures 12(c) and 13(c)).

Figures 12(c) and 13(c) show that the normalized occurrence rate is highest for the super-solar host stars (blue bin) for planets with orbital periods less than 10 days. This indicates that the host star metallicity is higher for super-Earths, Neptunes, sub-Saturns, and Jupiters at orbital periods of less than 10 days. At orbital periods longer than 10–30 days for planets with $R_P \leqslant 4\,R_\oplus$, the normalized occurrence rates do not show any clear trend.
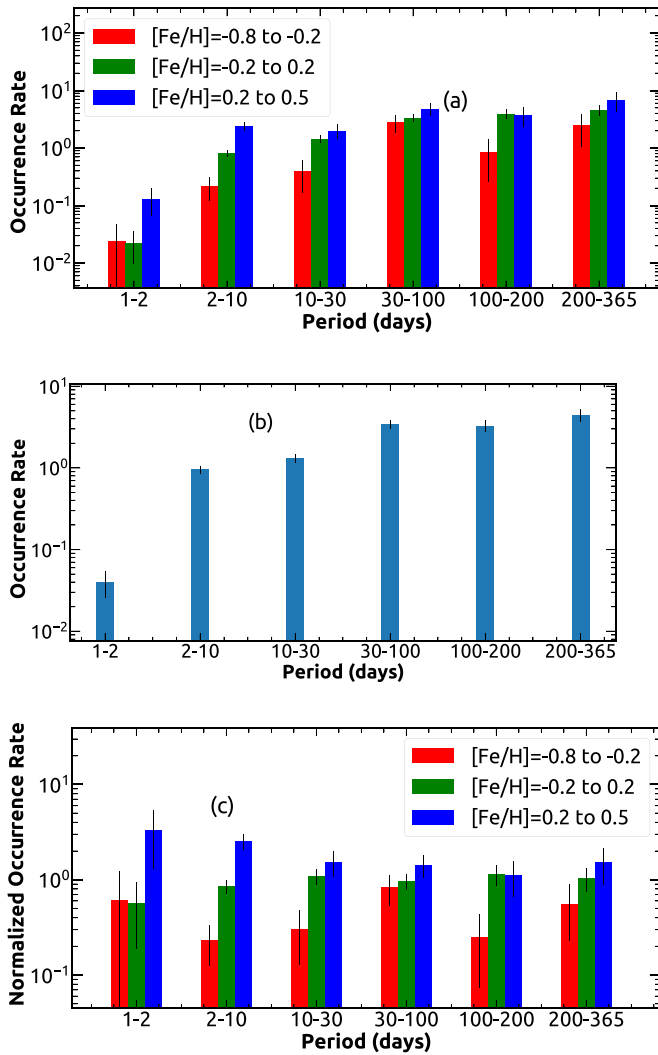
Figure 13. (a) Occurrence rate of exoplanets as a function of orbital period and host star metallicity for planets having radii between 4 and 20 $R_\oplus$. (b) Total occurrence rate of the sample without subdividing it into different metallicity bins. (c) Normalized occurrence rate of exoplanets as a function of orbital period and host star metallicity. The error bars in these plots are the Poissonian errors from the counting of planets.

The occurrence rates (Figure 13(a)) and normalized occurrence rates (Figure 13(c)) of sub-Saturns and Jupiters are always highest for the super-solar (blue) bin. This shows that sub-Saturns and Jupiters are preferentially formed around metal-rich host stars. These results are consistent with the findings of Figures 10 and 11. Similar results have been derived by Maldonado et al. (2018), Wilson et al. (2018), and Petigura et al. (2018).

7. Discussion

7.1. Comparing Our Results from DR25 with Petigura et al. (2018)

Making use of the host star and planetary properties from the CKS and the LAMOST spectra of the *Kepler* field to derive properties of non-planet-hosting stars, Petigura et al. (2018) estimated the occurrence rate of exoplanets as a function of host star metallicity, planetary radius, and orbital period.

They applied several filters to the sample, including a magnitude limit for the host star, with a much smaller spectral range. They also excluded planets with a grazing transit and planets around host stars that have a nearby stellar companion. After applying these filters, they had only 970 planets for which they computed the occurrence rate as a function of host metallicity, planet radius, and orbital period. In contrast, we are using the *Kepler* DR25 to calculate the occurrence rate as a function of host metallicity, planet radius, and orbital period.

On comparing our results with those of Petigura et al. (2018), we find similar trends between host star metallicity and planetary radius and orbital period, though for a much larger sample of 2864 planets. Both studies show that the average host star metallicity increases as the planetary radius increases (see Figures 4 and 7). Similarly, both studies show a higher occurrence rate for giant planets around metal-rich host stars, confirming the early claims in the literature (e.g., Santos et al. 2001, 2004; Fischer & Valenti 2005; Johnson et al. 2007, 2010; Sozzetti et al. 2009; Ghezzi et al. 2010; Sousa et al. 2011; Buchhave et al. 2012, 2014; Mortier et al. 2012; Mann et al. 2013; Buchhave & Latham 2015; see Figure 10). The average host star metallicity is also shown to be higher for planets with orbital periods of 10 days or less. The trends in the occurrence rate as a function of host star metallicity and period are similar in both studies (see Figures 12 and 13). For small planets ($R_P \leqslant 4\,R_\oplus$) with periods of about 10 days or less, Petigura et al. (2018) showed that the occurrence rate for metal-rich ([Fe/H] > 0) host stars was higher than that for metal-poor ([Fe/H] < 0) host stars. We also find similar results in Figure 12, where we find a higher occurrence rate for the solar- and super-solar-metallicity bins for small planets ($R_P \leqslant 4\,R_\oplus$) with periods less than 10 days.

The preceding discussion shows that although we use stellar parameter values from the DR25 stellar catalog (Mathur et al. 2017), the planet occurrence rates and their overall behavior as functions of planet radius, orbital period, and host star metallicity that we derive are consistent with those obtained by Petigura et al. (2018) using higher-precision stellar parameters from high-resolution spectroscopy.

Ongoing and upcoming surveys such as *K2* and *TESS* will observe a significantly larger number of stars than *Kepler*. High-resolution spectroscopy follow-up to estimate stellar parameters for such a large sample is likely to be extremely resource-intensive and time-consuming. Therefore, it is encouraging to know that stellar parameters that are less precise, such as the ones derived from broadband photometry (as was done for the KIC and still constitutes more than half of all the values of KOI host stars in the DR25 stellar table), can be useful to reveal trends between planetary and stellar parameters.

7.2. Disk Instability as the Formation Mechanism for Super-Jupiter Planets

It is clear from the analysis so far that the metallicity of the host star is strongly correlated with the radius and mass of the planet. We find that as the radius of the planet increases, the host star metallicity increases. A similar trend is seen for planetary mass as well: as the mass of the planet increases, the average host star metallicity also increases. This trend, however, reverses at about $4\,M_J$, after which, as the mass of the planet increases, the host star metallicity decreases. This trend of decreasing host star metallicity as the mass of the secondary increases extends even into the brown dwarf regime and low-mass stars (Schlaufman 2018).

The observed correlation of host star metallicity increasing with the increase in radius/mass of the planet is consistent with the predictions of the core accretion model of planet formation, where a 10–$15\,M_{\oplus}$ core needs to form before the planet can accrete gas and grow (Mizuno 1980; Pollack et al. 1996; Rice & Armitage 2003; Alibert et al. 2004; Ida & Lin 2004, 2005; Laughlin et al. 2004; Kornet et al. 2005; Johnson & Li 2012; Mordasini et al. 2012a). Both the star and the protoplanetary disk (from which the planets form) form out of the same cloud material. A metal-rich star indicates that the cloud from which the star formed was also metal-rich. This means that the protoplanetary disk would also be metal-rich. Since a metal-rich disk means more solid material available to form the planetesimals, the core can form faster and grow before the disk dissipates.

Our results seem to indicate that super-Jupiters ($M_{\mathrm{P}} > 4\,M_{\mathrm{J}}$) are preferentially found around metal-poor stars compared to stars that host Jupiters (1–$4\,M_{\mathrm{J}}$; see also Santos et al. 2017). An explanation for this observed bimodal population of giant planets is that Jupiters ($M_{\mathrm{P}} \leqslant 4\,M_{\mathrm{J}}$) are formed via core accretion, whereas super-Jupiters ($M_{\mathrm{P}} > 4\,M_{\mathrm{J}}$) are formed via the disk instability or the gravitational instability (e.g., Boss 1997; Rafikov 2005; Stamatellos & Whitworth 2008; Boley 2009; Boss 2010, 2011; Cai et al. 2010). The disk instability model assumes that the protoplanetary disk was massive enough to be unstable due to its own self-gravity. The major difference between this scenario and the core accretion model is the fact that in the disk instability model, the solid components of the disk do not play a direct role in the process of planet formation. Therefore, the properties of planets formed via disk instability will not be as strongly correlated with host star metallicity as is the case for planets formed via the core accretion model. Simulations of disk fragmentation have been shown to produce giant planets with a wide range in host star metallicity (Boss 2002).

We also find that the metallicity of host stars with brown dwarf companions is lower compared to that of host stars of Jupiters (1–$4\,M_{\mathrm{J}}$). The fact that masses of super-Jupiters ($M_{\mathrm{P}} > 4\,M_{\mathrm{J}}$) and of low-mass brown dwarfs are comparable and that their host star metallicities are similar possibly suggests a common formation mechanism for both. Brown dwarfs in orbit around a more massive companion are thought to form via disk instability (e.g., Bate et al. 2003; Bate 2012; Kratter & Lodato 2016), lending further credence to the claim that super-Jupiters could form via disk instability. Similar results have recently been reported by Maldonado & Villaver (2017), Santos et al. (2017), and Schlaufman (2018) as well.

7.3. Relationship between Host Star Metallicity and Planet Orbital Period

Our analysis shows that planets smaller (and less massive) than Jupiter that have orbital periods less than 10 days are preferentially found around higher-metallicity host stars compared to the ones with longer periods. A similar result was found by Mulders et al. (2016) and Petigura et al. (2018) but for a smaller sample.

One of the suggested explanations for this is that the hot super-Earths and Neptunes are the remnants of gas giants whose atmospheres have been eroded due to the extreme environment (Lopez et al. 2012; Lundkvist et al. 2016; Mazeh et al. 2016). The other possible explanation is that these hot planets might have started off like hot Jupiters but were not

able to accrete the gas quickly before the disk dissipated (Mulders et al. 2016).

Alternatively, this result can be explained by examining where the planets halt their migration (Mulders et al. 2016). Planet-forming disks around metal-rich stars will have more solid material in them. Such metal-rich disks can support planet migration much closer to the host star. This leads to planets migrating and being found much closer in for metal-rich stars (Mulders et al. 2016; Wilson et al. 2018; Petigura et al. 2018).

8. Summary

In this paper, we have studied the dependence of the observed properties (radius, mass, and orbital period) of exoplanets on their host star metallicity based on an analysis of more than 2800 *Kepler* exoplanet candidates using the *Kepler* DR25. Both stellar and planetary properties were taken from the DR25, except for stellar and planetary radii, for which we used improved estimates based on *Gaia* DR2 from Berger et al. (2018). This is the largest sample for which such a study has been carried out so far. The results presented in this paper are consistent with all of the previous work in the literature (Gonzalez 1997; Santos et al. 2001, 2004, 2017; Fischer & Valenti 2005; Johnson et al. 2007, 2010; Sozzetti et al. 2009; Buchhave et al. 2012, 2014; Mann et al. 2013; Buchhave & Latham 2015; Schlaufman 2015, 2018; Wang & Fischer 2015; Mulders et al. 2016; Maldonado & Villaver 2017; Wilson et al. 2018; Petigura et al. 2018), though for a much larger sample of exoplanets.

Most of the planet host stars in the DR25 sample have their properties determined from spectroscopy (about 60% of the total sample), whereas most of the non-planet hosts have their properties, particularly the metallicities, derived from photometry. Although we have applied corrections in order to make the spectroscopic and photometric metallicities consistent with each other, the metallicities derived from photometry have relatively larger uncertainties and are less accurate compared to spectroscopically determined metallicities. The occurrence rate calculated from such a mixed sample may not be as robust as that calculated with a pure spectroscopic sample for both host stars and non-host stars. But as stated earlier, the upcoming exoplanet surveys will be targeting several hundreds of thousands of stars; obtaining high-resolution spectra for all of the host stars would not be possible. Hence, it is reassuring that even with a mixed sample of host star properties, we are able to obtain results that have been derived previously using a much more homogeneous sample.

Our main results are summarized below.

1. We investigated the correlation between the radius or mass of the planet and the host star metallicity. We find that the host star metallicity, on average, increases with increasing planet radius/mass up to about $10\,R_{\oplus}$ or $1\,M_{\mathrm{J}}$. This is consistent with the predictions of the core accretion model of planet formation. In the case of planetary mass, we further show that as the mass of the planet increases above $4\,M_{\mathrm{J}}$, the average host star metallicity decreases. This indicates that super-Jupiters possibly have a different formation mechanism than the Jupiter-size giant planets.

2. We calculated the occurrence rate of planets as a function of planetary radius and host star metallicity to account for known observational biases and selection effects. We find

that the occurrence rate is a much stronger function of radius than host star metallicity. Therefore, in order to study the occurrence rate as a function of host star metallicity alone, we computed the normalized occurrence rate by removing the effect of planetary radius on the occurrence rate. We find that the normalized occurrence rate for super-solar-metallicity host stars increases as a function of planetary radius. This is consistent with the host star metallicity rising with increasing planet radius.

3. We further investigated the correlation between the host star metallicity and the orbital period of the planet. Planets with orbital periods less than 10 days appear to be more frequent around higher-metallicity stars compared to planets with orbital periods longer than 10 days. For sub-Saturn and Jupiter-size planets (R_P between 4 and $20 R_\oplus$), the occurrence rate is the highest for host stars with super-solar metallicity ([Fe/H] > 0.2) for all orbital periods. This further supports the idea that these planets are preferentially formed around metal-rich stars.

ORCID iDs

P. Manoj https://orcid.org/0000-0002-3530-304X
E. Furlan https://orcid.org/0000-0001-9800-6248
C. Mordasini https://orcid.org/0000-0002-1013-2811
Blesson Mathew https://orcid.org/0000-0002-7254-191X

References

Adams, E. R., Jackson, B., & Endl, M. 2016, AJ, 152, 47
Adibekyan, V. Z., Delgado Mena, E., Sousa, S. G., et al. 2012, A&A, 547, A36
Adibekyan, V. Z., Figueira, P., Santos, N. C., et al. 2013, A&A, 560, A51
Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989
Alibert, Y., Mordasini, C., & Benz, W. 2004, A&A, 417, L25
Andrews, S. M., Rosenfeld, K. A., Kraus, A. L., & Wilner, D. J. 2013, ApJ, 771, 129
Andrews, S. M., Wilner, D. J., Espaillat, C., et al. 2011, ApJ, 732, 42
Barros, S. C. C., Brown, D. J. A., Hébrard, G., et al. 2016, A&A, 593, A113
Batalha, N. M., Borucki, W. J., Bryson, S. T., et al. 2011, ApJ, 729, 27
Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, ApJS, 204, 24
Bate, M. R. 2012, MNRAS, 419, 3115
Bate, M. R., Lubow, S. H., Ogilvie, G. I., & Miller, K. A. 2003, MNRAS, 341, 213
Beaugé, C., & Nesvorný, D. 2013, ApJ, 763, 12
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, arXiv:1805.00231
Bodenheimer, P., Laughlin, G., & Lin, D. N. C. 2003, ApJ, 592, 555
Boley, A. C. 2009, ApJL, 695, L53
Bonfils, X., Lo Curto, G., Correia, A. C. M., et al. 2013, A&A, 556, A110
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, ApJ, 728, 117
Boss, A. P. 1997, Sci, 276, 1836
Boss, A. P. 2002, ApJL, 567, L149
Boss, A. P. 2010, ApJL, 725, L145
Boss, A. P. 2011, ApJ, 731, 74
Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112
Buchhave, L. A., Bizzarro, M., Latham, D. W., et al. 2014, Natur, 509, 593
Buchhave, L. A., & Latham, D. W. 2015, ApJ, 808, 187
Buchhave, L. A., Latham, D. W., Johansen, A., et al. 2012, Natur, 486, 375
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
Cai, K., Pickett, M. K., Durisen, R. H., & Milne, A. M. 2010, ApJL, 716, L176
Chabrier, G., Baraffe, I., Leconte, J., Gallardo, J., & Barman, T. 2009, in AIP Conf. Ser. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. E. Stempels (Melville, NY: AIP), 102
Christiansen, J. L. 2017, Planet Detection Metrics: Pixel-Level Transit Injection Tests of Pipeline Detection Efficiency for Data Release 25, KSCI-19110-001
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, ApJ, 810, 95
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279
Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, ApJS, 226, 7
Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, PASP, 120, 531
Demory, B.-O., & Seager, S. 2011, ApJS, 197, 12
Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
Dressing, C. D., Vanderburg, A., Schlieder, J. E., et al. 2017, AJ, 154, 207
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Fischer, D. A., & Valenti, J. 2005, ApJ, 622, 1102
Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Furlan, E., Luhman, K. L., Espaillat, C., et al. 2011, ApJS, 195, 3
Gaidos, E., Fischer, D. A., Mann, A. W., & Howard, A. W. 2013, ApJ, 771, 18
Gettel, S., Charbonneau, D., Dressing, C. D., et al. 2016, ApJ, 816, 95
Ghezzi, L., Cunha, K., Smith, V. V., et al. 2010, ApJ, 720, 1290
Gonzalez, G. 1997, MNRAS, 285, 403
Hadden, S., & Lithwick, Y. 2014, ApJ, 787, 80
Hartman, J. D., Bakos, G. Á., Bhatti, W., et al. 2016, AJ, 152, 182
Hartman, J. D., Bakos, G. Á., Torres, G., et al. 2011, ApJ, 742, 59
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15
Hsu, D. C., Ford, E. B., Ragozzine, D., & Morehead, R. C. 2018, AJ, 155, 205
Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, ApJS, 211, 2
Ida, S., & Lin, D. N. C. 2004, ApJ, 616, 567
Ida, S., & Lin, D. N. C. 2005, ApJ, 626, 1045
Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, PASP, 122, 905
Johnson, J. A., Fischer, D. A., Marcy, G. W., et al. 2007, ApJ, 665, 785
Johnson, J. A., Petigura, E. A., Fulton, B. J., et al. 2017, AJ, 154, 108
Johnson, J. L., & Li, H. 2012, ApJ, 751, 81
Kim, K. H., Watson, D. M., Manoj, P., et al. 2013, ApJ, 769, 149
Kim, K. H., Watson, D. M., Manoj, P., et al. 2016, ApJS, 226, 8
Kornet, K., Bodenheimer, P., Różyczka, M., & Stepinski, T. F. 2005, A&A, 430, 1133
Kratter, K., & Lodato, G. 2016, ARA&A, 54, 271
Laughlin, G., Bodenheimer, P., & Adams, F. C. 2004, ApJL, 612, L73
Lillo-Box, J., Demangeon, O., Santerne, A., et al. 2016, A&A, 594, A50
Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, ApJS, 197, 8
Lopez, E. D., Fortney, J. J., & Miller, N. 2012, ApJ, 761, 59
Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, NatCo, 7, 11201
Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, RAA, 15, 1095
Ma, B., & Ge, J. 2014, MNRAS, 439, 2781
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2015, AJ, 154, 94
Maldonado, J., & Villaver, E. 2017, A&A, 602, A38
Maldonado, J., Villaver, E., & Eiroa, C. 2018, A&A, 612, A93
Mann, A. W., Gaidos, E., Kraus, A., & Hilton, E. J. 2013, ApJ, 770, 43
Manoj, P., Kim, K. H., Furlan, E., et al. 2011, ApJS, 193, 11
Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, ApJS, 210, 20
Mathur, S., Huber, D., Batalha, N. M., et al. 2017, ApJS, 229, 30
Mathur, S., Huber, D., Batalha, N. M., et al. 2018, ApJS, 234, 43
Mayo, A. W., Vanderburg, A., Latham, D. W., et al. 2018, AJ, 155, 136
Mazeh, T., Holczer, T., & Faigler, S. 2016, A&A, 589, A75
McClure, M. K., Furlan, E., Manoj, P., et al. 2010, ApJS, 188, 75
Miller, N., Fortney, J. J., & Jackson, B. 2009, ApJ, 702, 1413
Mizuno, H. 1980, PThPh, 64, 544
Mohanty, S., Greaves, J., Mortlock, D., et al. 2013, ApJ, 773, 168
Mordasini, C., Alibert, Y., Benz, W., Klahr, H., & Henning, T. 2012a, A&A, 541, A97
Mordasini, C., Alibert, Y., Georgy, C., et al. 2012b, A&A, 547, A112

Mordasini, C., Mollière, P., Dittkrist, K.-M., Jin, S., & Alibert, Y. 2015, IJAsB, 14, 201
Mortier, A., Santos, N. C., Sozzetti, A., et al. 2012, A&A, 543, A45
Morton, T. D., & Swift, J. 2014, ApJ, 791, 10
Mulders, G. D. 2018, arXiv:1805.00023
Mulders, G. D., Pascucci, I., & Apai, D. 2015a, ApJ, 798, 112
Mulders, G. D., Pascucci, I., & Apai, D. 2015b, ApJ, 814, 130
Mulders, G. D., Pascucci, I., Apai, D., Frasca, A., & Molenda-Żakowicz, J. 2016, AJ, 152, 187
Muzerolle, J., Allen, L. E., Megeath, S. T., Hernández, J., & Gutermuth, R. A. 2010, ApJ, 708, 1107
Obermeier, C., Koppenhoefer, J., Saglia, R. P., et al. 2016, A&A, 587, A49
Pascucci, I., Mulders, G. D., Gould, A., & Fernandes, R. 2018, ApJL, 856, L28
Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, ApJ, 831, 125
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, PNAS, 110, 19273
Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, AJ, 154, 107
Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, AJ, 155, 89
Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icar, 124, 62
Pope, B. J. S., Parviainen, H., & Aigrain, S. 2016, MNRAS, 461, 3399
Rafikov, R. R. 2005, ApJL, 621, L69
Rice, W. K. M., & Armitage, P. J. 2003, ApJL, 598, L55
Santos, N. C., Adibekyan, V., Figueira, P., et al. 2017, A&A, 603, A30
Santos, N. C., Israelian, G., & Mayor, M. 2001, A&A, 373, 1019
Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153
Santos, N. C., Sousa, S. G., Mortier, A., et al. 2013, A&A, 556, A150
Schlaufman, K. C. 2015, ApJL, 799, L26
Schlaufman, K. C. 2018, ApJ, 853, 37
Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, ApJ, 669, 1279
Serenelli, A. M., Bergemann, M., Ruchti, G., & Casagrande, L. 2013, MNRAS, 429, 3645
Sinukoff, E., Howard, A. W., Petigura, E. A., et al. 2016, ApJ, 827, 78
Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M., & Udry, S. 2011, A&A, 533, A141
Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373
Sozzetti, A., Torres, G., Latham, D. W., et al. 2009, ApJ, 697, 544
Spake, J. J., Brown, D. J. A., Doyle, A. P., et al. 2016, PASP, 128, 024401
Stamatellos, D., & Whitworth, A. P. 2008, A&A, 480, 879
Swift, D. C., Eggert, J. H., Hicks, D. G., et al. 2012, ApJ, 744, 59
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Thorngren, D. P., & Fortney, J. J. 2018, AJ, 155, 214
Vanderburg, A., Latham, D. W., Buchhave, L. A., et al. 2016, ApJS, 222, 14
Wang, J., & Fischer, D. A. 2015, AJ, 149, 14
Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 768, 14
Wilson, R. F., Teske, J., Majewski, S. R., et al. 2018, AJ, 155, 68
Winn, J. N. 2018, arXiv:1801.08543
Winn, J. N., & Fabrycky, D. C. 2015, ARA&A, 53, 409
Youdin, A. N. 2011, ApJ, 742, 38