

Planet Hunters: the first two planet candidates identified by the public using the *Kepler* public archive data[★]

Debra A. Fischer,¹† Megan E. Schwamb,^{2,3}‡ Kevin Schawinski,^{2,3}§ Chris Lintott,^{4,5} John Brewer,¹ Matt Giguere,¹ Stuart Lynn,⁴ Michael Parrish,⁵ Thibault Sartori,^{1,6} Robert Simpson,⁴ Arfon Smith,^{4,5} Julien Spronck,¹ Natalie Batalha,⁷ Jason Rowe,⁸ Jon Jenkins,⁹ Steve Bryson,⁸ Andrej Prsa,¹⁰ Peter Tenenbaum,⁹ Justin Crepp,¹¹ Tim Morton,¹¹ Andrew Howard,¹² Michele Belem,¹ Zachary Kaplan,¹ Nick vanNispen,¹ Charlie Sharzer,¹ Justin DeFouw,¹³ Agnieszka Hajduk,¹³ Joe P. Neal,¹³ Adam Nemeč,¹³ Nadine Schuepbach¹³ and Valerij Zimmermann¹³

¹Department of Astronomy, Yale University, New Haven, CT 06511, USA

²Department of Physics, Yale University, PO Box 208121, New Haven, CT 06520, USA

³Yale Center for Astronomy and Astrophysics, Yale University, PO Box 208121, New Haven, CT 06520, USA

⁴Oxford Astrophysics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH

⁵Adler Planetarium, 1300 South Lake Shore Drive, Chicago, IL 60605, USA

⁶Ecole normale supérieure, 45, rue d'Ulm / 29 rue d'Ulm, F-75230 Paris Cedex 05, France

⁷Department of Physics and Astronomy, San Jose State University, San Jose, CA 95192, USA

⁸NASA Ames Research Center, Moffett Field, CA 94035, USA

⁹SETI Institute/NASA Ames Research Center, Moffett Field, CA 94035, USA

¹⁰Department of Astronomy and Astrophysics, Villanova University, 800 East Lancaster Avenue, Villanova, PA 19085, USA

¹¹Department of Astrophysics, California Institute of Technology, MS 249-17, Pasadena, CA 91125, USA

¹²Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA

¹³Planet Hunter

Accepted 2011 September 30. Received 2011 September 30; in original form 2011 July 13

ABSTRACT

Planet Hunters is a new citizen science project designed to engage the public in an exoplanet search using NASA *Kepler* public release data. In the first month after launch, users identified two new planet candidates which survived our checks for false positives. The follow-up effort included analysis of Keck HIRES spectra of the host stars, analysis of pixel centroid offsets in the *Kepler* data and adaptive optics imaging at Keck using NIRC2. Spectral synthesis modelling coupled with stellar evolutionary models yields a stellar density distribution, which is used to model the transit orbit. The orbital periods of the planet candidates are 9.8844 ± 0.0087 d (KIC 10905746) and 49.7696 ± 0.00039 d (KIC 6185331), and the modelled planet radii are 2.65 and 8.05 R_{\oplus} . The involvement of citizen scientists as part of Planet Hunters is therefore shown to be a valuable and reliable tool in exoplanet detection.

Key words: stars: individual: KIC 10905746 – stars: individual: KIC 6185331 – planetary systems.

1 INTRODUCTION

The past decade has witnessed an explosion in the number of known planets beyond our Solar system. From the ground, planet searches using techniques that include Doppler observations, transit photometry, microlensing and direct imaging have identified more than 500 exoplanets (Schneider 2011; Wright et al. 2011). These observations have provided a wealth of information, including constraints on dynamical interactions in multiplanet systems, non-coplanar orbits of hot Jupiters and atmospheric properties of transiting gas giant

[★]This publication has been made possible by the participation of more than 40 000 volunteers in the Planet Hunters project. Their contributions are individually acknowledged at <http://www.planethunters.org/authors>.

†E-mail: debra.fischer@yale.edu

‡NSF Fellow.

§Einstein Fellow.

planets. The combination of Doppler and photometric measurements of transiting planets is particularly informative because it yields planet densities and enables theoretical modelling of the interior structure and composition of exoplanets.

The *Kepler* mission is monitoring more than 150 000 stars with unprecedented 29-min observing cadence (Jenkins et al. 2010) and a relative photometric precision approaching 20 ppm in 6.5 h for $K_p = 12$ mag stars to search for transiting planets. After just 1 yr of operation, Borucki et al. (2010a) announced the detection of 706 transiting planet candidates based on the first quarter (Q1) data. On 2011 February 1, one month before the 2-yr anniversary of launch, the total number of planet candidates increased to more than 1200 (Borucki et al. 2011). The Q1 data were released into the public archive in 2010 June, followed by a release of second quarter (Q2) data in 2011 February. The public archive is hosted by the Multi-mission Archive at STScI (MAST¹) and the NASA/IPAC/NEExSci Star and Exoplanet Database (NSTED²).

Although there are more than 1200 *Kepler* candidates, only 1–2 per cent of these are confirmed planets with measured masses from Doppler observations (Borucki et al. 2010b; Batalha et al. 2011). These are challenging confirmations. The *Kepler* stars are faint compared to stars in ground-based radial velocity surveys, and most of the *Kepler* candidates have radii consistent with Neptune-like planets, so most of the stellar reflex velocities are comparable to the formal measurement errors. Transit timing variations (Holman et al. 2010; Lissauer et al. 2011) offer a novel way to derive planet masses, but require multiplanet systems with measurable non-Keplerian orbital perturbations.

The *Kepler* team has developed sophisticated algorithms for detecting transits by fitting and removing periodic or quasi-periodic stellar variability (with low and high frequencies). In addition to modelling out background variability, the *Kepler* pipeline stitches together data from different observing quarters by determining the median flux from adjacent observing windows and using polynomial fits across the boundary. The *Kepler* team developed the transit planet search (TPS) algorithm, a wavelet-based adaptive filter to identify a periodic pulse train with temporal widths ranging from 1 to 16 h (Jenkins 2002; Jenkins et al. 2010). Photometric uncertainties are assessed to identify light curves with phase-folded detection statistics exceeding 7.1σ . This threshold was selected so that given the number of required independent statistical tests per star, 4 yr of data for the entire set of *Kepler* targets could be robustly searched for orbital periods up to 2 yr.

While the human brain is exceptionally good at detecting patterns, it is impractical for a single individual to review each of the $\sim 150\,000$ light curves in every quarterly release of the *Kepler* data base. However, crowd-sourcing this task has appeal because human classifiers have a remarkable ability to recognize archetypes and to assemble groups of similar objects while disregarding obvious glitches that can trip up computer algorithms. This skill has recently been put to use in a wide range of scientific fields, from galaxy morphology to protein folding. To engage these uniquely human talents, and to give the public the opportunity to participate in an exciting exoplanet search, we developed Planet Hunters³ to present *Kepler* light curves to the public.

Planet Hunters is a new addition to the successful Zooniverse network of Citizen Science Alliance projects (Lintott et al. 2008,

2011) and the first Zooniverse project to present time series data (rather than images) to the public. The site was launched on 2010 December 16, and after six months, more than 40 000 users have made more than 3 million light-curve classifications. Here we describe the layout of the site and two new planet candidates identified by the public using the PlanetHunter interface.

2 IDENTIFYING TRANSITS

The Planet Hunters website makes use of the Zooniverse⁴ tool set, which now supports a wide variety of citizen science projects. Its primary function is to serve up assets – in this case, ~ 33 d flux-corrected light curves derived from the *Kepler* data – to an interface and to collect user-generated interactions with these data.

Previous Zooniverse projects have included a separate tutorial to assist volunteers. While the Planet Hunters website includes such a tutorial, initial guidance is given within the interface, accessed via a single click from the site home page. Volunteers see a light curve with example transits and can then begin to classify data. Users who have not registered with the Zooniverse, or who are not logged in, can begin classifying but receive frequent reminders to log in. The site supports prioritization of the light curves; for logged-in users viewing the Q1 data discussed in this paper, simulated or already identified transits were shown 5 per cent of the time. A curve associated with a dwarf star was then shown 66 per cent of the remaining time and one associated with a giant star 33 per cent of the time. Once a category (i.e. simulated light curve, dwarf or giant star) has been selected, a light curve is chosen randomly from the top 10 scoring assets in that category. (The score is the number of transits marked on each curve.) Once curves have been classified by 10 volunteers, they are removed from the list. The results are made available to the science team immediately via a private website.

The actual classification proceeds via a decision tree. In the first step, users are asked whether the light curve is variable or quiet (icons and help buttons provide visual prompts). The user is then asked whether any transit features are present and has the option to zoom in and out of particular areas of the light curve. If transit features are found, the user can mark them with boxes as demonstrated in Fig. 1. In some cases, the transit features seen are synthetic transits of known period and radius, which are used to assess the completeness of the user classifications.

After all transits are marked, the user has the option to discuss this particular star on the Planet Hunters Talk site and connect with other citizen scientists. The user can also download the light curve data to analyse it independently or save the star to their ‘favourites’. The Discussion Board (‘Talk’⁵) is a critical component of the Planet Hunters project. Here, the science team interacts with the public, and experienced users establish collections of similar light curves (e.g. ‘Variables in a Hurry’, ‘Definite Transits’ and ‘Weird Stars’) and provide advice for new users. The integration of discussion into the work flow has been successful in encouraging greater participation than in previous Zooniverse projects; more than 60 per cent of registered Planet Hunters participants visit ‘Talk’, and more than 35 per cent make comments.

¹ <http://archive.stsci.edu/>

² <http://nsted.ipac.caltech.edu>

³ www.planethunters.org

⁴ www.zooniverse.org

⁵ <http://talk.planethunters.org/>

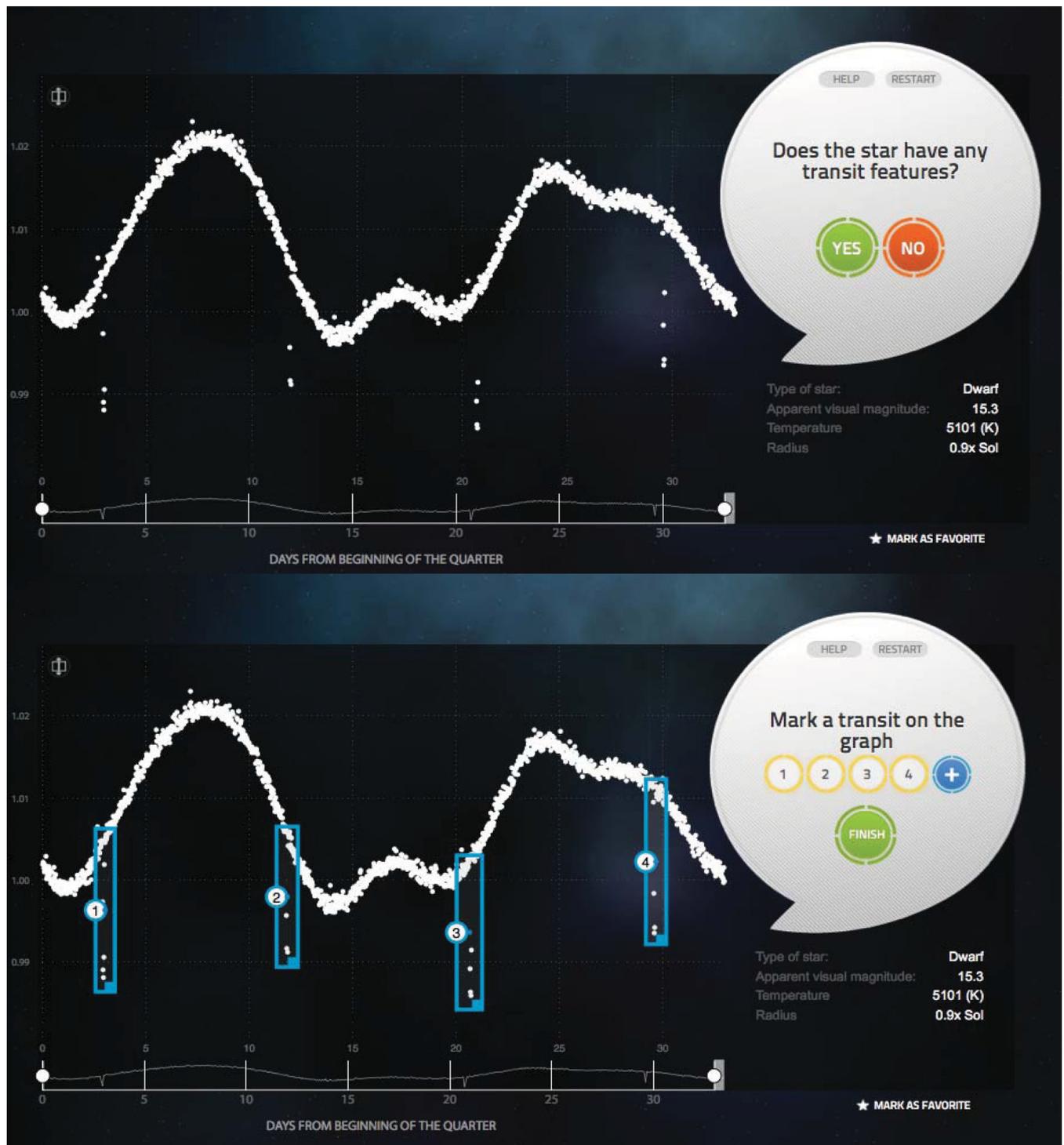


Figure 1. These slides from the Planet Hunters interface show the light curve for KOI 889.01 (top). Participants use a mouse drag to identify prospective transit features (bottom).

2.1 Planet Hunters detection efficiency

As a first check, we visually inspected all user assessments made in the first month after the site was launched for the first 306 *Kepler* planet candidates announced by Borucki et al. (2010a). This essentially provided a ‘head count’ or a rough estimate of how many transit events were being flagged by participants, and it provided feedback that was considered by the web development team for

upgrades to the site (e.g. streamlining the assessment questions and transit marking routines). Note that this is simply a tally of the fraction of transits that were marked; we are not calculating the percentage of planets detected. For example, if a sample of 10 stars had 100 transit events and 80 of them were marked by 50 per cent of classifiers, then the percentage of detected transits would be $0.8 \times 50 = 40$ per cent. The 306 *Kepler* planet candidates (Borucki et al. 2010a) exhibited 1371 transits with planet radii between 0.1 and

$1 R_{\text{Jup}}$. Overall, we found that two-thirds of the transits for candidates announced by Borucki et al. (2010a) were correctly flagged. Only 10 per cent of transit boxes were spurious (i.e. did not obviously correspond to a transit event).

3 Kepler PLANET HUNTERS CANDIDATES

We also visually inspected ~ 3500 transit flags marked by Planet Hunters in light curves where five or more people indicated that a transit had been found. We first eliminated the known false positives, typically grazing and eclipsing binaries (Batalha et al. 2010; Rowe et al. 2010; Prsa et al. 2011), and published *Kepler* candidates (Borucki et al. 2010a, 2011) from the set of light curves flagged by Planet Hunters. On our internal website, the team searched the extracted light curves, ran periodogram analyses, modelled light curves for prospective candidates and checked for correlated pixel brightness centroid shifts to try to eliminate additional false positives. After an extensive filtering process, we reduced the number of possible planet candidates down to a preliminary list of 45.

We ranked these candidates and sent the ‘top 10’ to our *Kepler* co-authors; they examined the light curves with their data verification pipeline and immediately found that six of the 10 were unlikely to be planet candidates. KIC 11904734 has a V-shaped transit and very large radius, suggesting an eclipsing binary (EB) star system. KIC 8043052 and KIC 12009347 have secondary occultations that are also consistent with EB systems. KIC 4913000 and KIC 9097892 showed changing transit depths from quarter to quarter. This can occur when a nearby star contributes an amount of flux that is quarter-dependent, changing as the instrumental point spread function changes. A more complete pixel centroid analysis showed that the transit signals for KIC 4913000, KIC 8242434 and KIC 9097892 were offset from the star by 4–6 arcsec. KIC 11820830 initially appeared to be a strong planet candidate; however, stellar modelling indicated that the most likely interpretation for this star was that it was an EB system with a large early type star as the primary and an M or K dwarf secondary. The six false positive candidates are listed in Table 1.

However, three candidates survived the *Kepler* data verification pipelines. One of these is a possible multiplanet candidate, and we are now obtaining Doppler follow-up. The remaining two candidates are presented here. Each of these candidates had in fact been flagged in Q1 by the *Kepler* TPS as threshold crossing events. However, for various reasons, these objects were not promoted to the status of a ‘Kepler Object of Interest’ (KOI).

3.1 KIC 10905746

KIC 10905746 has a *Kepler* magnitude of 13.496 and $g - r$ colour of 0.949. The Kepler Input Catalogue (KIC; Kepler Mission Team 2009) does not list T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$ or stellar radius for this star. The star was dropped from the *Kepler* target list after Q1

because variability characteristics (amplitude and frequency) indicated that the star could be a giant and was therefore less desirable for the exoplanet transit survey; planet transit signals are much shallower and more difficult to detect around stars with large radii. The photometry for this star shows low-frequency variability, with a period of ~ 16 d and an amplitude of more than 2 per cent, which could be caused by spots rotating on the surface of the star.

The Planet Hunters participants were able to look past the large-scale structure in the light curve, and they identified possible transit events with a depth of about 0.2 per cent that repeated on ~ 10 d intervals in the Q1 data. The shape and depth of the light curve seemed consistent with a planet, and we did not detect photocentre offsets in the pixel arrays in our initial screening, which would have indicated a blended background EB system.

To better understand the host star, we obtained a spectrum of this star at Keck with a resolution of $R \sim 55\,000$, using HIRES (Vogt et al. 1994) on 2011 April 12. A faint companion was observed at a separation of about 5 arcsec on the guide camera, and the image rotator was used to ensure that the light from the companion did not enter the slit. With the excellent seeing and the greater than 1 mag difference between KIC 10905746 and the companion star, the scattered light contamination would have been less than one part in a thousand. The spectrum had a signal-to-noise ratio (S/N) of about 140, and we used the Spectroscopy Made Easy (SME) code (Valenti & Piskunov 1996; Valenti & Fischer 2005) to model the stellar parameters: $T_{\text{eff}} = 4237 \pm 114$ K, $\log g = 4.73 \pm 0.1$, $v \sin i = 1.1 \pm 1$ km s $^{-1}$ and $[\text{Fe}/\text{H}] = -0.23 \pm 0.1$. The surface gravity that we measure with our local thermodynamic equilibrium spectroscopic analysis is consistent with a main-sequence star, rather than an evolved giant. Fig. 2 (left, top row) shows a wavelength segment that includes the Mg I b triplet lines from the Keck spectrum. The wings of these lines are sensitive gravity indicators. However, in this case, the star is cool with significant line blanketing, which suppresses the continuum and makes it difficult to model the line wings. We tested the hypothesis that this star was a giant by running a grid of synthetic models and fixing the gravity between $\log g$ of 2.0 and 3.5. The chi-squared fit for our models improved with decreasing surface gravity over this range, but all fits were significantly worse than our model with $\log g = 4.73$.

The Ca II H&K lines provide additional support of main-sequence status for this star. Late type main-sequence stars often have significant emission in the spectral line cores as a result of dynamo-driven magnetic activity in the star, like the strong emission in the Ca II H&K line cores, shown in Fig. 2 for KIC 10905746. However, it is far less common for evolved stars to show emission unless the stars are rapidly rotating or members of close spectroscopic binary systems (Gray & Nagar 1985; Gunn, Mitrou & Doyle 1998; GizisE, Reid & Hawley 2002; Isaacson & Fischer 2010), and we see no evidence for either of these attributes in KIC 10905746. The combination of emission in the cores of the Ca II H&K and pressure-broadened wings in the Mg I b lines, together with the spectroscopic T_{eff} , suggests that the star has a spectral type of roughly M0V. The stellar parameters are summarized in Table 2.

Our *Kepler* co-authors found that the *Kepler* TPS algorithm had flagged the light curve for KIC 10905746 in Q1 with a multiple event statistic (MES) of 9σ , greater than the 7.1σ threshold. However, the fit failed to converge during the next stage of data verification. As a result, the star was dropped, the full pipeline analysis was never carried out until it was flagged by the Planet Hunters.

The *Kepler* time series photometry for Q1 is shown in the top panel of Fig. 3 (after removing the large amplitude, low-frequency

Table 1. False positive planet candidates.

Star name	Comments
KIC 11904734	V-shaped transit and very large radius (EB)
KIC 8043052	Secondary occultations (EB)
KIC 12009347	Secondary occultations (EB)
KIC 4913000	Astrometric motion in pixel centroids (BGEB)
KIC 9097892	Astrometric motion in pixel centroids (BGEB)
KIC 11820830	Eclipsing binary (based on model fits)

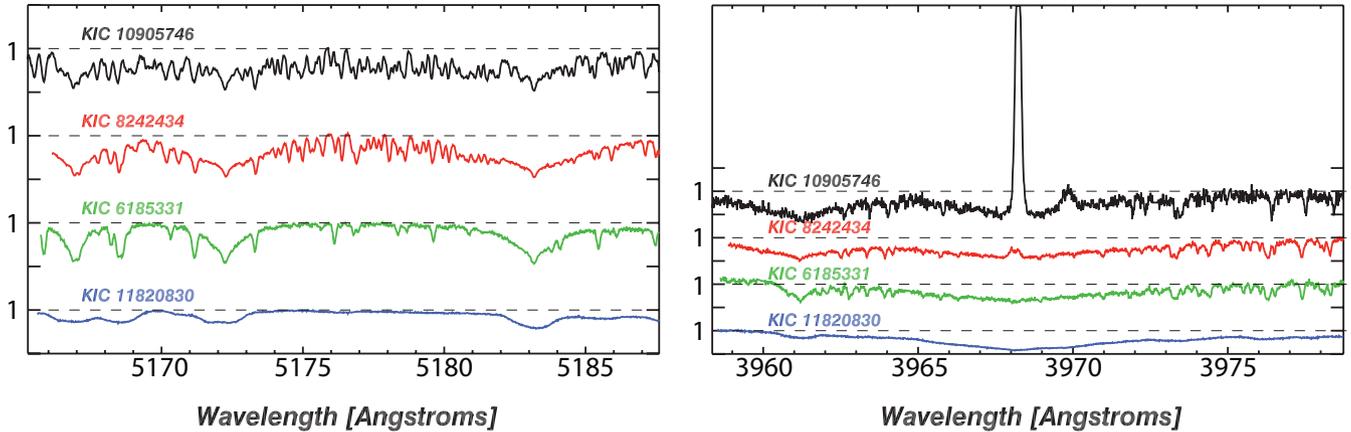


Figure 2. Left: the wings of the Mg *Ib* triplet lines are sensitive to pressure broadening, making these lines useful diagnostics of the surface gravity or luminosity class of stars. The spectra above were obtained at Keck and the stars are ordered from high to low surface gravity based on our spectral synthesis models. Right: emission in the cores of the Ca II H&K line is an activity indicator for main-sequence stars. The spectra above show the Ca II H&K line for each of the planet candidate hosts presented here. The strong emission for KIC 10905746 is typical for a late type main-sequence star.

Table 2. Stellar parameters.

Parameter	KIC 10905746	KIC 6185331	KIC 8242434	KIC 11820830
Right ascension	18 54 30.92	18 57 05.75	19 39 49.22	19 40 51.98
Declination	48 23 27.6	41 32 06.1	44 08 59.3	50 05 03.58
<i>Kepler</i> mag	13.49	15.64	13.05	12.09
<i>g</i> − <i>r</i>	0.949	0.556	0.937	0.198
M^* (M_{\odot})	0.578 (0.032)	1.027 (0.042)	0.761 (0.028)	2.25 (0.3)
R^* (R_{\odot})	0.548 (0.026)	1.27 (0.17)	0.719 (0.031)	4.1 (0.3)
<i>Z</i>	0.0119 (0.003)	0.0261 (0.0032)	0.0234 (0.003)	
Age (Gyr)	–	8.7 (1.5)	–	
L^* (L_{\odot})	0.086 (0.081)	1.02 (0.03)	0.77 (0.04)	2.25 (0.3)
ρ^* (g cm^{-3})	4.97 (0.54)	0.70 (0.26)	2.9 (0.38)	
T_{eff} (K)	4240 (112)	5619 (80)	4757 (60)	6300 (250)
[Fe/H]	−0.23 (0.1)	+0.11 (0.15)	+0.07 (0.08)	+0.26 (0.2)
$v \sin i$ (km s^{-1})	1.1 (0.50)	0.5 (0.50)	0.4 (0.50)	52 (5)
$\log g$	4.724 (0.028)	4.239 (0.098)	4.608 (0.041)	3.6 (0.2)

variability). The bottom panel of Fig. 3 shows the data folded at the prospective orbital period, and the red curve is the best-fitting theoretical curve with a period of 9.8844 ± 0.0087 d, an orbital inclination of $88^{\circ}42$ and an inferred planet radius of $2.65 \pm 0.67 R_{\oplus}$. Just above the transit curve, we show the photometry from the opposite phase, where a putative secondary occultation might be observed. The search for secondary occultations allowed for eccentric orbits that were consistent with the data and stellar parameters. The antitransit data are folded at a phase of 0.5 since no eccentricity or secondary occultations were detected when modelling the light curves.

A Monte Carlo analysis (Jenkins et al. 2008) iterates between a family of evolutionary models in the Yale–Yonsei isochrones (Demarque et al. 2004; Yi, Demarque & Kim 2008) and the spectroscopic and orbital parameters (orbital period, transit depth and duration) to provide self-consistent estimates for uncertainties and stellar parameters, including *Z* (total heavy element abundance), age, density, luminosity, mass and radius. For KIC 10905746, age is not listed in Table 2 since there was almost no constraint from the evolutionary tracks. Since the transit depth is a function of the ratio of the planet to star radius, an accurate assessment of the stellar radius is critical for deriving the planet radius. The characteristics of this planet candidate are summarized in Table 3.

Because we do not have an independent measurement of the mass of the transiting object, KIC 10905746 is a planet candidate rather than a confirmed planet. Photometrically diluted background eclipsing binaries (BGEBS) can have transit depths similar to planets. The depth of an EB system will normally be 10 per cent or more (depending on the ratio of the stellar radii to the impact parameter), but if the EB light curve is blended with a brighter foreground star, the composite light curve will have a shallower depth during the eclipse and can masquerade as a transiting planet candidate. However, other signatures of the BGEBS can sometimes be found in the light curve: unequal primary eclipse and secondary occultation or V-shaped light curves (Batalha et al. 2010). Three tests were carried out to search for a BGEBS. First, the light curve was examined for deviations from a planet model (e.g. variations in the depths of alternating transits or evidence for secondary occultations). In Fig. 3, the even and odd transits are indicated with plus symbols and asterisks, respectively, and show that the alternating transit events do not have significant variations in depth and are well-fitted with a transiting planet model, which is overplotted as a solid line. The photometric data plotted just above the transit curve are phase folded at the predicted time of secondary occultation for a BGEBS and fitted with a theoretical (green) line that solves for an occultation with zero depth. We note that the search for occultations does not assume

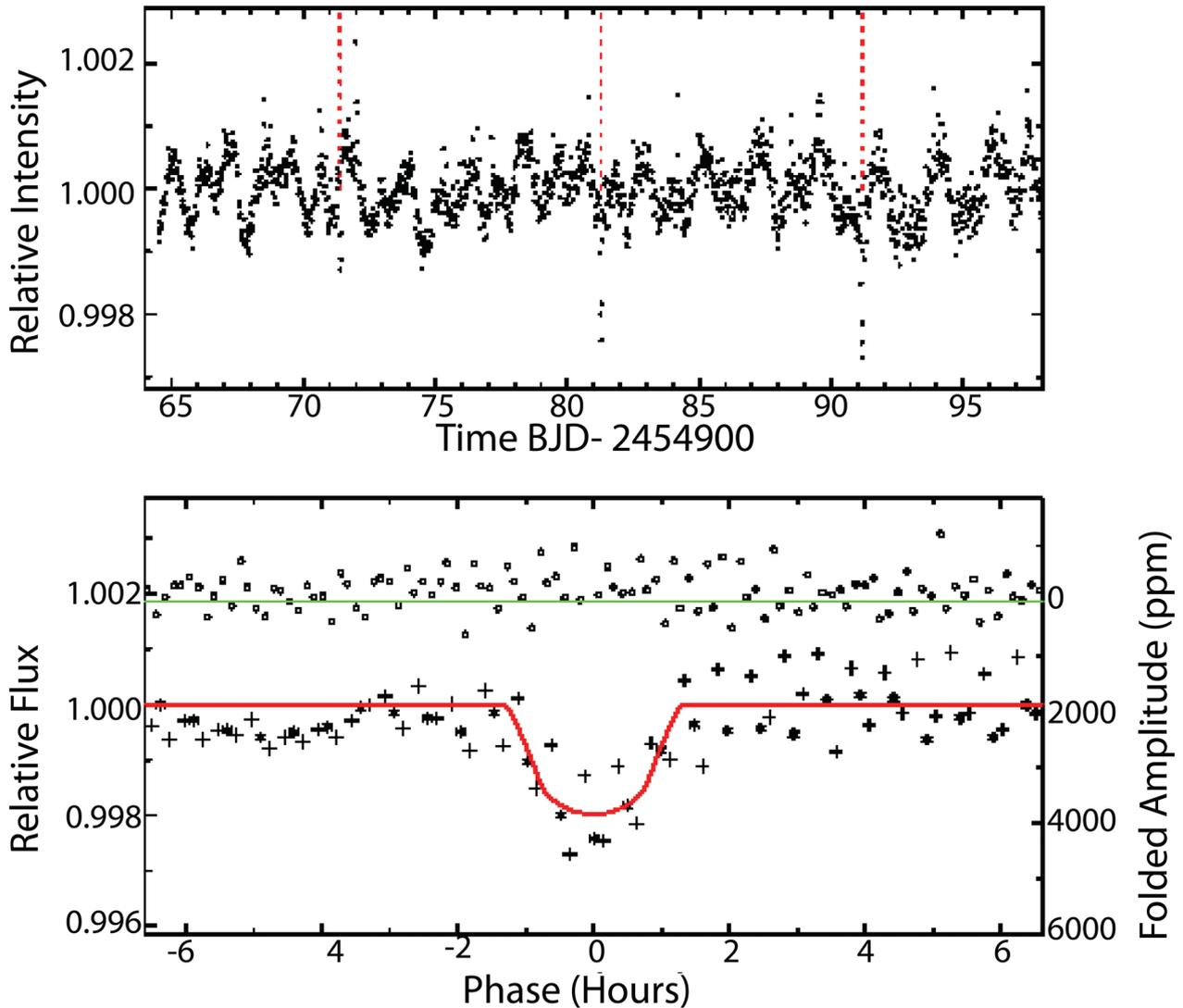


Figure 3. The top panel shows the time series data for KIC 10905746 between 2009 May 2 and June 15 after removing a large amplitude periodic signal. Planet Hunters flagged the three transit events indicated with a vertical dashed red line in the Q1 data. In the bottom panel, the light curve is phase-folded at the prospective orbital period $P = 9.8846$ d after removing the baseline variability. The fitted transit model is overplotted with a red curve. Just above the transit light curve, the antitransit photometry is plotted and fitted with a green curve showing zero depth for the occultation.

Table 3. Characteristics of planet candidates.

Parameter	KIC 10905746	KIC 6185331
T_0 (BJD - 245 4900)	71.4045 (0.0102)	92.9877 (0.0028)
Orb. per. (d)	9.8844 (0.0087)	49.769 71 (0.000 39)
Impact parameter b	0.82 (0.21)	0.642 (0.142)
R_{PL}/R_*	0.0442 (0.0110)	0.0581 (0.0018)
$e \sin \omega$	0.08 (0.42)	0.10 (0.32)
$e \cos \omega$	0.00 (0.43)	0.00 (0.34)
R_{PL} (R_{\oplus})	2.65 (0.67)	8.05 (1.08)
Incl ($^{\circ}$)	88.42 (0.42)	89.20 (0.21)
a/R_*	29.4 (1.1)	38.1 (8.4)
a (au)	0.0751 (0.0014)	0.2672 (0.0036)
T depth (ppm)	1881 (343)	3633 (59)

zero eccentricity; however, zero eccentricity is used to generate the antitransit phased plot. For many BGEs, some dimming would be observed. The lack of a detected occultation is a necessary, but still not exclusive, condition for a planet origin of the transit event.

To place stronger limits on the presence of a blended BGE, adaptive optics (AO) observations were obtained on 2011 June 23 UT using NIRC2 at Keck. The conditions were excellent with ~ 0.5 arcsec seeing and very little cirrus. The spatial resolution of the K -band AO images is about 45 mas. Fig. 4 (top panels) compares a K -band image of KIC 10905746 from Two Micron All Sky Survey (2MASS)⁶ (left-hand panels) with our diffraction-limited K -band AO images (right-hand panels) with square root scaling for the brightness. The 2MASS image is unresolved, but reveals a faint source ~ 4.2 arcsec east of KIC 10905746, identified as KIC 10905748. The high-resolution K -band AO images clearly resolves these two sources. Our ability to rule out other close companions depends on the brightness contrast of the stars in the K band and their angular separation. The 3σ magnitude differences for excluding other sources are listed in Table 4 for separations ranging from 0.25 out to 4.0 arcsec. We also obtained J -band images to better

⁶ <http://irsa.ipac.caltech.edu/applications/FinderChart>

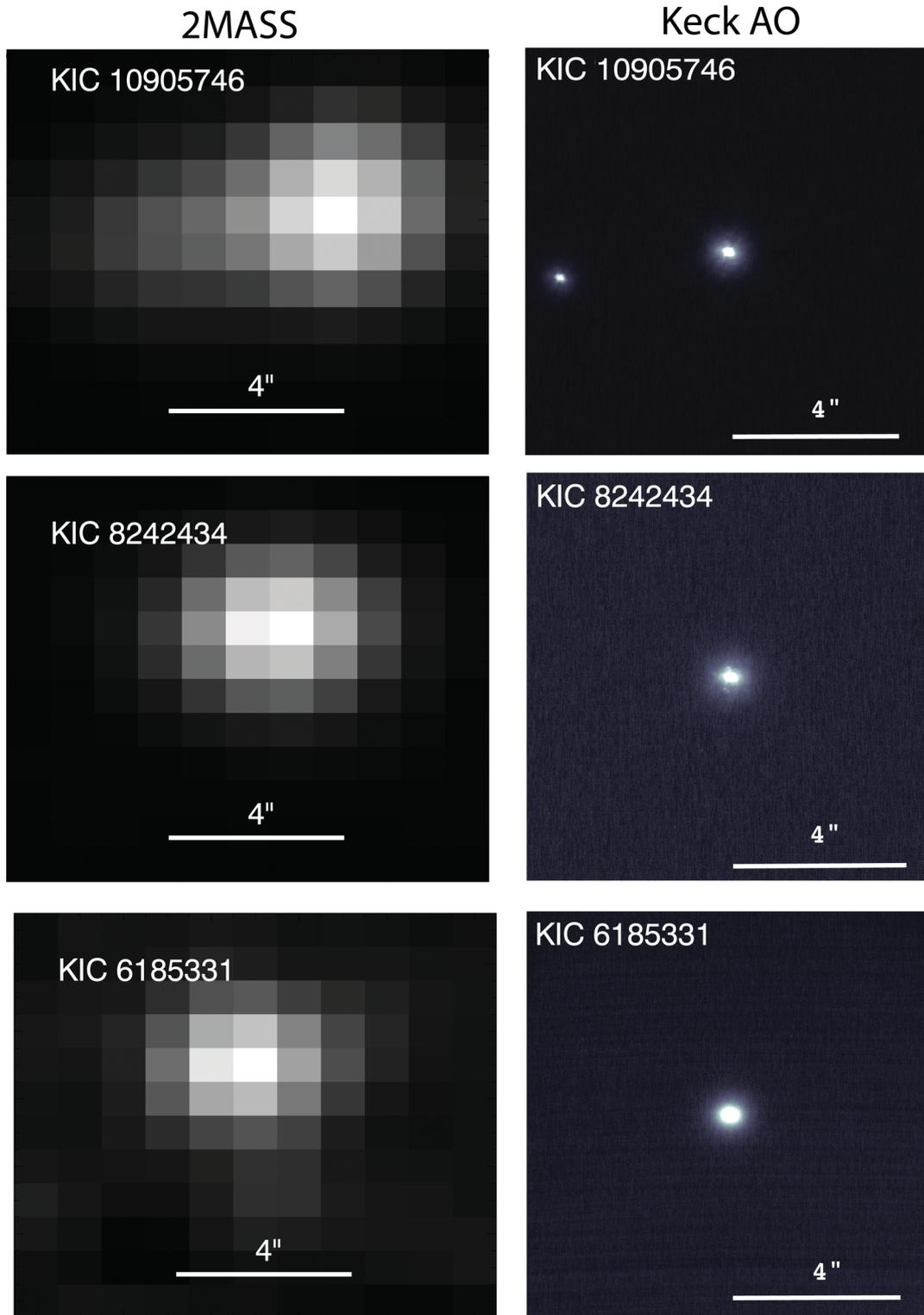


Figure 4. The 2MASS K -band images (left) and AO images (right) for the two planet candidate hosts, KIC 10905746 and KIC 6185331 and for a star where a background EB was found, KIC 8242434. The horizontal line indicates the image scale in arcseconds. North is up in these images and east is to the left. KIC 10905746 is shown in the top panel; the 2MASS image shows distortion from a nearby star at about 4 arcsec due East, which is completely resolved by the AO K -band image (right). KIC 8242434 is shown in the middle row. No additional sources were observed down to the magnitude limits listed in Table 4. However, the photocentre was observed to shift during the prospective transit, indicating that a nearby background EB star producing the transit signal. The bottom panel shows images for KIC 6315331 with weaker limits on excluded background sources because of the intrinsic faintness of this star.

Table 4. ΔK magnitude AO exclusion limits.

Star name	0.25 arcsec	0.5 arcsec	1.0 arcsec	2.0 arcsec	4.0 arcsec
KIC 10905746	3.7	5.6	7.8	8.6	8.6
KIC 6185331	1.1	2.7	4.8	5.2	5.3
KIC 8242434	3.5	5.1	7.2	7.8	7.9
KIC 11820830	4.2	5.7	7.2	7.6	7.6

characterize the neighbouring source. The magnitude difference between KIC 10905746 and KIC 10905748 is $\Delta K = 1.42$ mag and $\Delta J = 1.38$ mag. These images did not reveal any additional prospective contaminating sources.

The file headers of the *Kepler* data contain information about the pixel centroid at the time of every photometric measurement. If the transit is occurring on the source, then the brightness of the star will decrease, but the image centroid position will be unchanged. However, if we are really observing a blended system with a background EB that is offset from the source, then the image centroid will shift during the eclipse. Centroids for the pixel images in the *Kepler* data were examined for this astrometric motion. The pixel centroid analysis yielded a high S/N detection for KIC 10905746, and no sign of astrometric motion was detected beyond the error circle of beyond 0.08 pixel. While these results do not rule out a background binary close to KIC 10905746, they do eliminate the nearby star, KIC 10905748, which is 46σ away in the model fit, as the source of the transit.

3.2 KIC 6185331

The Planet Hunters identified a single transit event for KIC 6185331 in the Q1 data and one additional transit was found in the Q2 *Kepler* light curve. The *Kepler* team notes that the TPS code also identified this as a prospective candidate with a MES of about 10σ in Q1 and 20σ in Q2. However, the data verification pipeline did not trigger to process these curves.

According to the Kepler Input Catalogue, KIC 6185331 has a *Kepler* magnitude of 15.64, $g - r$ colour of 0.556, stellar radius of $0.664 R_{\odot}$, $T_{\text{eff}} = 5578$, $\log g = 4.786$ and $[\text{Fe}/\text{H}] = -0.287$. We obtained a spectrum of this star with S/N ~ 30 using HIRES at Keck Observatory. Our spectral synthesis modelling with SME yields an effective temperature of 5615 ± 80 K, consistent with the KIC value. However, our analysis yields a lower gravity of $\log g = 4.19 \pm 0.15$. Comparing the Mg *1b* triplet lines (Fig. 2), there is indeed less pressure broadening than for KIC 8242434, which had a $\log g$ of 4.608. We also derive a slightly less metal rich composition than the KIC, with $[\text{Fe}/\text{H}] = +0.11 \pm 0.1$, and we obtain a best-fitting model for the lines with $v \sin i = 0.5 \text{ km s}^{-1}$. No emission is seen in the core of the Ca II H&K lines (Fig. 2), indicating that this sun-like star has low chromospheric activity.

Fig. 5 shows the time series data (top) and the phase-folded data (bottom), modelled with a 49.769 71-d period using the Q1–Q7 data. We carried out the Monte Carlo analysis described in Section 3.1 for KIC 10905746 with the Y2 isochrones, orbital parameters and the spectral synthesis results to obtain self-consistent stellar parameters (listed in Table 2). With the derived stellar radius of $1.27 R_{\odot}$, the planet is modelled with a best-fitting radius of $8.05 R_{\oplus}$. There is some evidence in the model fit for an eccentric orbit or stellar radius as large as $1.4 R_{\odot}$. We did not detect a contaminating BGEB: alternating transit events have the same depth, no decrease in brightness is observed at the predicted occultation time and the pixel centroid

analysis yielded a clean result for a transit on KIC 6315331 without any detected astrometric motion. The 2MASS and AO images are shown in Fig. 4 (bottom, left and right). Because this is the faintest of the stars (*Kepler* magnitude of 15.64), the AO images can only rule out contaminating background stars within $\Delta M_V < 2.7$ mag at separations larger than 0.5 arcsec. The AO contrast sensitivities are listed in Table 4.

3.3 KIC 8242434

Planet Hunters identified a single transit event in the Q1 data for KIC 8242434. When the Q2 data were released, two additional transit events were identified that were separated by 44 d. In consultation with the *Kepler* team, we learned that the TPS had flagged this star with a MES of about 10σ . Because this was a single event, the data verification was not processed until Q2 and was not classified as a KOI.

The KIC lists a *Kepler* magnitude of 13.054 and $g - r$ colour of 0.937, $T_{\text{eff}} = 4665$ K, $\log g = 4.176$, a high metallicity of $+0.437$ and a stellar radius of $1.337 R_{\odot}$ for KIC 8242434. We analysed a Keck HIRES spectrum with S/N of about 55 and derive a similar temperature, $T_{\text{eff}} = 4757 \pm 60$ K. However, we find a higher surface gravity, $\log g = 4.608 \pm 0.1$, consistent with a main-sequence luminosity class. The wings of the Mg *1b* triplet lines (Fig. 2) are broad and by eye are consistent with the higher surface gravity. Our analysis also yields a lower metallicity, $[\text{Fe}/\text{H}] = 0.07$ and $v \sin i = 0.4 \text{ km s}^{-1}$. The Ca II H&K lines (Fig. 2) have emission in the line core; this emission would be typical for a low-mass main-sequence star, but less common for a subgiant. The stellar parameters are summarized in Table 2.

Fig. 6 shows the time series and phase-folded Q1–Q7 photometry for KIC 8242434. The light curve does not show evidence for a BGEB: the transit depth is constant for alternating transits and no dimming occurs at the predicted time of occultation in the phase-folded data just above the transit curve. The orbital period is modelled as 44.963 888 d. A Monte Carlo analysis was used to iterate to the self-consistent stellar parameters listed in Table 2 (again, there was no good constraint for the stellar age). The stellar radius is estimated to be $0.719 R_{\odot}$, and together with the transit depth, this implies a planet radius of $2.32 R_{\oplus}$. The parameters for the planet candidate are summarized in Table 3.

The measured position of the transit source shows a statistically significant (5.7σ) 0.6 arcsec offset from KIC 8242434, indicating that the transit signal is likely due to a dim background binary. The source position is measured by taking robust weighted average of the observed transit source position in Q1–Q8, as determined by centroiding the difference between average in-transit and out-of-transit pixels (Bryson et al., in preparation). Modelling indicates that this offset is not due to systemic centroid biases due, for example, to crowding. The *K*-band 2MASS image is shown in Fig. 4 (middle, left) and the AO image (Fig. 4, right) shows some unusually bright speckles within an arcsecond, with the most prominent one in the south-east. The AO images and pixel centroid analysis casts doubt on the planet interpretation and suggests the presence of a confusing background source, likely a BGEB.

3.4 KIC 11820830

KIC 11820830 exhibits significant oscillations; however, participants readily identified several transit events in the Q1 light curve. The *Kepler* TPS had also flagged this star with a MES of 46σ , the highest S/N threshold of any of the candidates presented in this

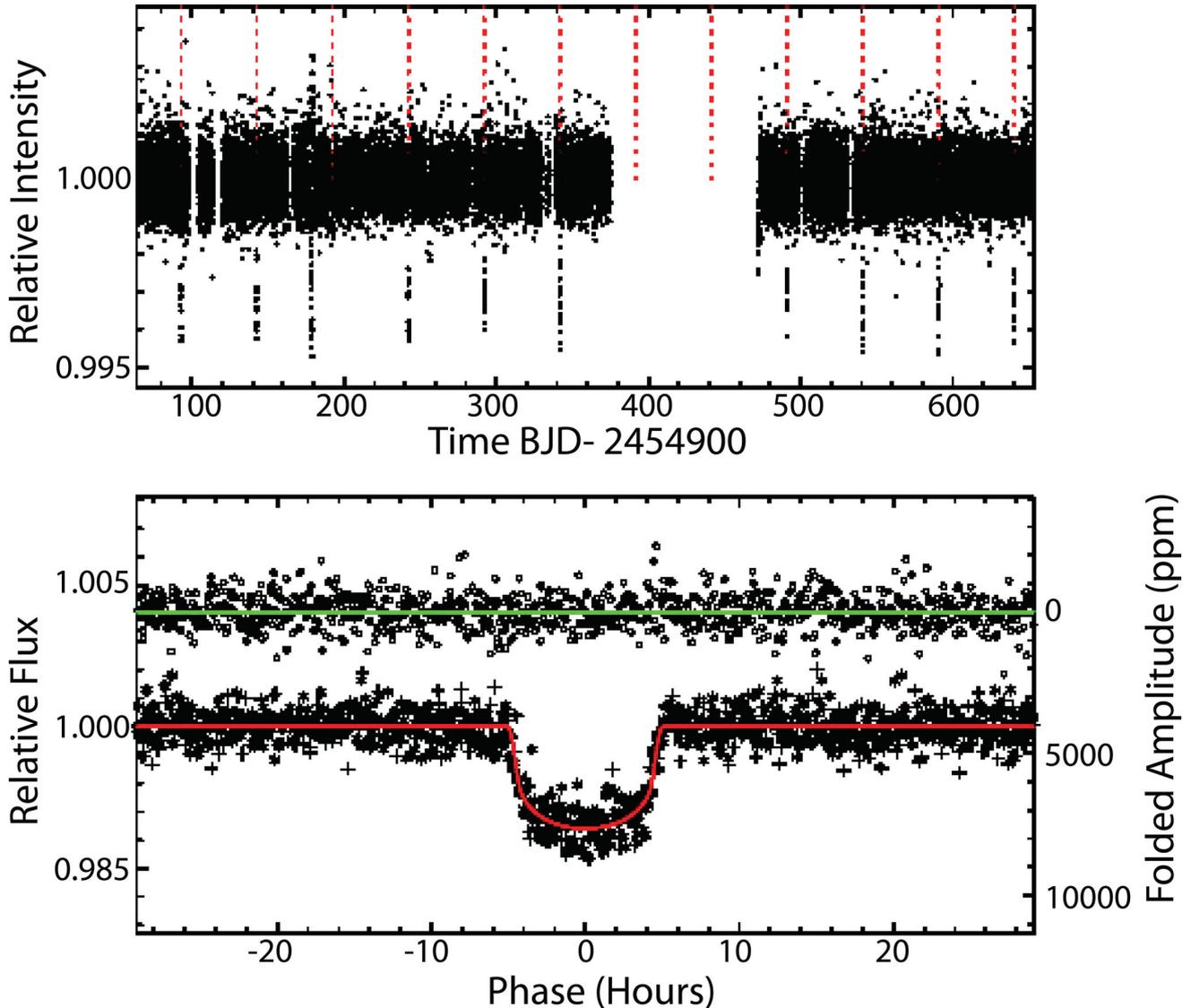


Figure 5. The time series data for KIC 6185331 (top) include Q1–Q7 data. Planet Hunters flagged a single transit in the Q1 data and one additional transit was seen in the Q2 data. The bottom panel shows the data folded at the prospective orbital period, 49.7700 d.

paper. However, the light curve failed additional tests and was not processed by the data verification pipeline. Fig. 7 shows the remarkable time series (top) and phase-folded (bottom) light curves for Q1–Q7 observations of this star.

The KIC lists stellar parameters for KIC 11820830, including *Kepler* magnitude of 12.087, $g - r$ colour of 0.198, stellar radius of $1.428 R_{\odot}$, $T_{\text{eff}} = 7007$ K, $\log g = 4.224$ and $[\text{Fe}/\text{H}] = -0.009$. We obtained a spectrum of this star using HIRES on Keck with S/N of 90. We carried out spectral synthesis modelling and derive spectroscopic properties of the star.

This is the brightest of our initial Planet Hunters candidates, and normally it would have been possible to follow up on this star with Doppler measurements to confirm the mass of the transiting object. However, our spectroscopic analysis revealed a high rotational velocity, $v \sin i = 52 \pm 5 \text{ km s}^{-1}$ which significantly reduces the intrinsic radial velocity precision. Fig. 2 shows the Keck wavelength segments for the Mg *1 b* triplet and Ca *II* H&K lines, respectively, and the high rotational velocity is apparent from the broad stellar lines in these figures. The broad spectral lines also reduce the precision

of our derived spectral parameters. With this caveat, we report the results of our analysis: $T_{\text{eff}} = 6300 \pm 250$ K, $\log g = 3.6 \pm 0.2$ and $[\text{Fe}/\text{H}] = +0.26 \pm 0.2$.

Unfortunately, the self-consistent Monte Carlo analysis indicates that KIC 11820830 is likely to be an EB system, with an early type primary star eclipsed by a K or M dwarf in an eccentric orbit. No astrometric motion was detected in the pixel centroid analysis, and the AO images did not detect an additional source with a $\Delta M_V < 4$ mag at separations of 0.25 arcsec. The AO contrast sensitivities are summarized in Table 4.

4 DISCUSSION

The Planet Hunters website was launched to engage the public in front-line research by presenting light-curve data from the *Kepler* mission. This project joins a growing list of citizen science Zooniverse projects and is the first to present time series data, rather than images. We debated whether the unique pattern recognition skills of the human brain would be able to compete with the efficient

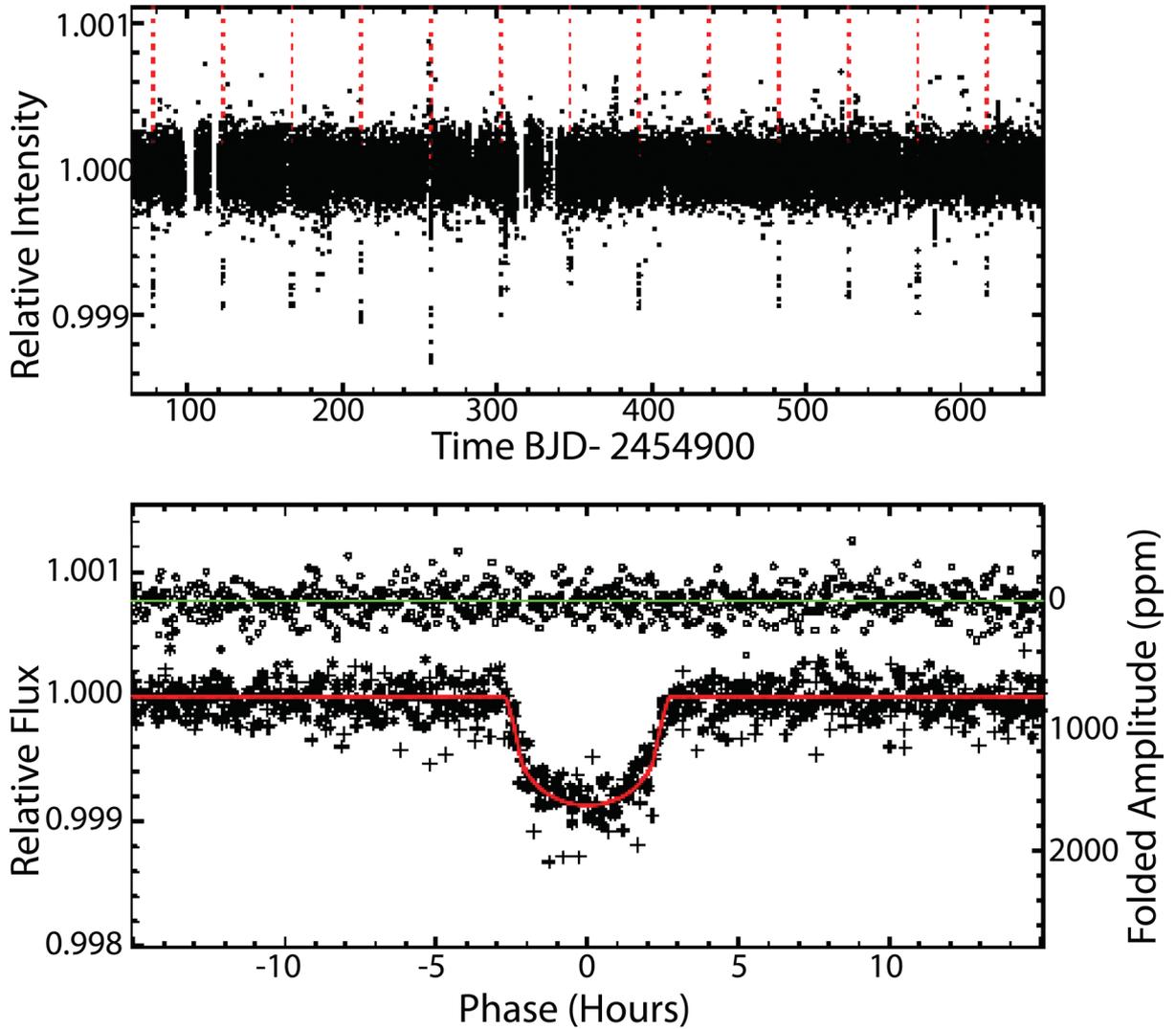


Figure 6. The time series data for KIC 8242434 (top) include photometry for Q1–Q7, provided by the *Kepler* team. Planet Hunters flagged a single transit in the Q1 data and two additional transits were found in the Q2 data. The bottom panel shows the data folded at the prospective orbital period, 44.9634 d. Unfortunately, the pixel centroid check shows that this is likely a background EB system.

computer algorithms. However, we expected that citizen scientists might discover unexpected patterns in the data or unusual types of transits, which could then be used as feedback to further improve the *Kepler* transit search algorithms. Citizen scientists identified some unusual objects in the Galaxy Zoo programme, and we expected that some unpredictable and unanticipated discoveries and correlations might also emerge from Planet Hunters. Automated algorithms and citizen science are complementary techniques and both are important to make the best use of the *Kepler* data.

An initial assessment was made of the performance and efficiency of the Planet Hunters participants by counting the number of transit events detected among the 306 candidates announced for Q1 data by Borucki et al. (2010a). We found that Planet Hunters flagged about two-thirds of those transit events. The deeper transits were found more often than the shallow transits.

In the first month after the launch of the Planet Hunters website, more than 40 stars were flagged as possible planet transits that were not known false positives (grazing binaries or blended BGEs) or published *Kepler* candidates. Because we felt it was important to preserve the integrity of the *Kepler* planet candidates, we contacted

members of the *Kepler* team who provided important data verification for our top 10 candidates. More than half of these were found to be false positives.

We present the first two planet candidates, discovered by Planet Hunters using Q1 data: KIC 10905746 and KIC 6315331, with orbital periods that range from 9.88 to 49.96 d and radii ranging from 2.32 to 8.0 R_{\oplus} . We have carried out a Monte Carlo analysis for a self-consistent set of stellar parameters and analysed the pixel centroid’s to check for astrometric motion. We also obtained AO observations to eliminate BGEs with separations wider than ~ 0.5 arcsec and $\Delta M_V < 5$ in the infrared *K*-band data. However, the pixel centroid analysis and AO observations cannot exclude eclipsing binaries that are closer than 0.5 arcsec or those with wider separations that are more than about 5 mag fainter than the tentative planet host stars. Because such systems could still produce the observed light curves, these two candidates are not confirmed planets.

We estimate false positive probabilities (FPP) for the two candidates presented here following the framework presented in Morton & Johnson (2011), which relies on Galactic structure and stellar population synthesis models. We consider two possible false

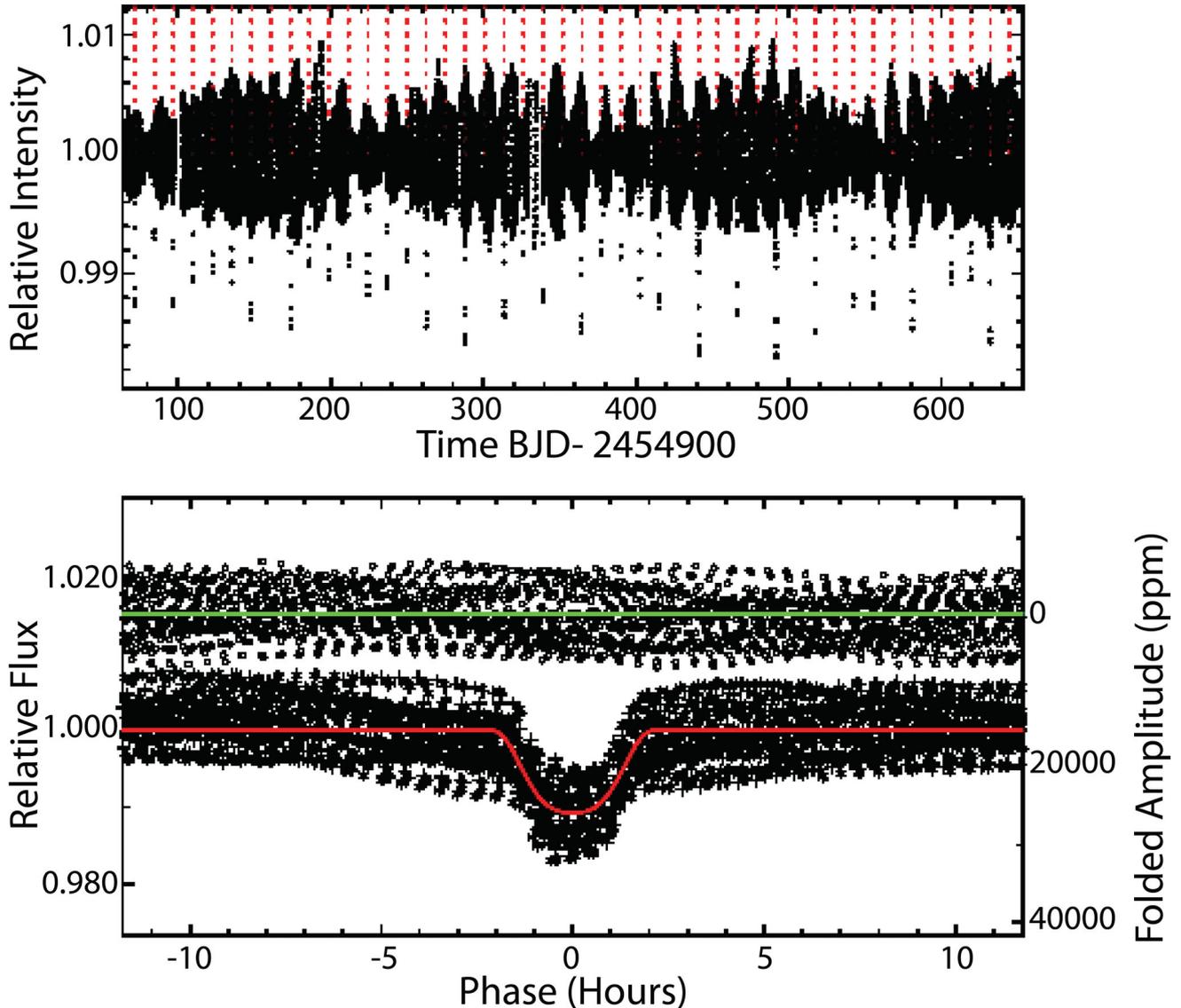


Figure 7. The time series data for KIC 11820830 (top) include Q1–Q7 data. This star has a remarkably variable background. However, Planet Hunters were able to see past that structure and flagged several transits in the Q1 data. The bottom panel shows the phase-folded data with the prospective orbital period, 12.7319 d. Unfortunately, the best model for this star suggests that the primary is an early type star with an eclipsing M or K dwarf companion.

positive scenarios: chance-alignment blended eclipsing binaries and hierarchical triple eclipsing systems, both of which can produce signals that mimic transiting planets. However, these transits are not V-shaped, we do not observe a secondary eclipse, and pixel offset calculations and AO observations indicate that any possible blending systems can only reside within a fraction of an arcsecond of the target stars. Thus, we are able to put strong statistical constraints on the likelihood of false positive scenarios. Assuming an overall 20 per cent occurrence rate for planets, a planet radius function $dN/dR \sim R^{-2}$, and the binary and multiple system properties according to Raghavan et al. (2010), as discussed in more detail in Morton and Johnson (2011), we derive an FPP of only 0.3 per cent for KIC 10905746 and an FPP of 5.0 per cent for KIC 6185331. The higher FPP for KIC 6185331 is set primarily by the fact that it has a deeper transit and thus is more susceptible to the hierarchical blend false positive scenario, which is not significantly constrained by the AO observations or centroid analysis.

An obvious question is why these candidates were not identified by the *Kepler* team. One motivation for the Planet Hunters project was that there might be odd cases that computer algorithms might miss, but that the human brain would adeptly identify. In fact, we learned that all of the planet candidates presented here had previously been flagged by the TPS algorithm. However, two of the candidates presented here had multi-quarter light curves that did not converge and the third candidate was dropped after Q1 because it was thought to be an evolved star. Therefore, these stars were not promoted to the status of a KOI, which would have triggered extensive follow-up. It is not really surprising that a few candidates failed to converge in the analysis pipelines and remained behind to be gleaned by Planet Hunters. The discoveries presented in this paper show that citizen scientists can make important contributions. However, extensive follow-up was still required to eliminate false positives.

Planet Hunters is a novel and complementary technique to the *Kepler* team's detection algorithms. Algorithms are now being

developed to process Planet Hunters classifications and to assess the capabilities of individual volunteers based on light curves injected with synthetic short-period planet transits. Weightings will be assigned to individuals, and an iterative process will be used to converge on final classifications for each star. These algorithms will extract transit candidates automatically, and this analysis will be presented in a future paper.

ACKNOWLEDGMENTS

The data presented in this paper are the result of the efforts of the Planet Hunters volunteers, without whom this work would not have been possible. The following list of people flagged transit events for the light curves discussed in this paper: Juan Camilo Arango Alvarez, Mary-Helen Armour, Ferdinand de Antoni, Frank Barnett, Carolyn Bol, Les Bruffell, Dr. David M. Bundy, Troy Campbell, Elisabeth Chaghafi, Amirouche Chahour, Arunangshu Chakraborty, Mathias Chapuis, Fabrice Cordary, Daniel, ClmentDoyen, Graham Dungworth, Michael Richard Eaton, David Evans, Raymond Ashley Evans, Evgeniy, Enrique Ferreyra, Marc Fiedler, Dave Fischer, Fin J. R. FitzPatrick, Dr Ed Foley, Lorenzo Fortunato, Sebastian Frehmel, Robert Gary Gagliano, Fabio Cesar Gozzo, Howard Hallmark, Dave Henderson, Brent Hugh, Inizan, irishcoffee, Thomas Lee Jacobs, Marta Kałużna, Bill Kandiliotis, Rafal Kurianowicz, Lukasz Kurzysz, Piotr Laczny, David M. Lindberg, Janet Lomas, Luis Miguel Moreira Calado Lopes, Michael Lujanac, Lukasz, Cssia Solange Lyra, Jacek M., Riccardo Marzi, Karen McAuley, Tomasz Miller, Cedric MOULIS, Adrian Nicolae,Njaal, Michael W. Novak, Osciboj, Kai Pietarinen, Anna Podjaska, Marc A. Powell (Omeganon), Gerry A. Prentice, John M. Raser, René-Pierre BUIGUES, Andres Eloy Martinez Rojas, Rouz, Andrey Saponov, Matt Schickele, Terrance B. Schmidt, David Smith, Paulina Sowick, Lubomír Štiak, Charles H. Tidwell, III, tuckdydes, v5anw, Joop Vanderheiden, VIATG, vovcik91, Sbastien Wertz, Bohdan Widla, Steven C. Wooding, Charles Yule.

DAF acknowledges funding support from Yale University and support from the NASA Supplemental Outreach Award, 10-OUTRCH.210-0001. DAF thanks the Yale Keck TAC for telescope time used to obtain data for this paper. MES is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-100325. Support for the work of KS was provided by NASA through Einstein Postdoctoral Fellowship grant number PF9-00069, issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060. The Zooniverse is supported by The Leverhulme Trust. We gratefully acknowledge the dedication and achievements of Kepler Science Team and all those who contributed to the success of the mission. We acknowledge

use of public release data served by the NASA/IPAC/NEExSci Star and Exoplanet Database, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. We particularly thank the organizers (Charles Beichman, Dawn Gelino and Carolyn Brinkman) and lecturers (David Ciardi, Stephen Kane and Kaspar von Braun) at the 2010 July Sagan Summer Workshop for providing information and guidance that led to the inspiration for the Planet Hunters site. The Kepler public release data is primarily hosted by the Multimission Archive (MAST) at the Space Telescope Science Institute (STScI) operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-*HST* data is provided by the NASA Office of Space Science via grant NNX09AF08G. This research has made use of NASA's Astrophysics Data System Bibliographic Services.

REFERENCES

- Batalha N. M. et al., 2010, *ApJ*, 713, L103
 Batalha N. et al., 2011, *ApJ*, 729, 27
 Borucki W. et al., 2010a, *Sci*, 327, 977
 Borucki W. J. et al., 2010b, *ApJ*, 713, L126
 Borucki W. J. et al., 2011, *ApJ*, 736, 19
 Demarque P., Woo J.-H., Kim Y.-C., Yi S. K., 2004, *ApJS*, 155, 667
 Gizis J. E., Reid I. N., Hawley S. L., 2002, *AJ*, 123, 3356
 Gray D., Nagar P., 1985, *ApJ*, 298, 756
 Gunn A. G., Mitrou C. K., Doyle J. G., 1998, *MNRAS*, 296, 150
 Holman M. et al., 2010, *Sci*, 330, 51
 Isaacson H., Fischer D. A., 2010, *ApJ*, 725, 875
 Jenkins J. M., 2002, *ApJ*, 575, 493
 Jenkins J. M. et al., 2008, *ApJ*, 724, 1108
 Jenkins J. M. et al., 2010, *Proc. SPIE*, 7740, 10
 Kepler Mission Team, 2009, The Kepler Input Catalog, *VizieR On-line Data Catalog: V/133*
 Lintott C. et al., 2008, *MNRAS*, 389, 1179
 Lintott C. et al., 2011, *MNRAS*, 410, 116
 Lissauer et al., 2011, *Nat*, 470, 53
 Morton T. D., Johnson J. A., 2011, *ApJ*, 738, 170
 Prsa A. et al., 2011, *AJ*, 141, 83
 Raghavan D. et al., 2010, *ApJS*, 190, 1
 Rowe J. F. et al., 2010, *ApJ*, 713, L150
 Schneider J., 2011, <http://www.encyclopedia.eu>
 Valenti J. A., Piskunov N., 1996, *A&AS*, 118, 595
 Valenti J. A., Fischer D. A., 2005, *ApJS*, 159, 141
 Vogt S., S. et al., 1994, *Proc. SPIE*, 2198, 362.
 Wright J. T. et al., 2011, *PASP*, 123, 412
 Yi S. K., Demarque P., Kim Y.-C., 2004, *Ap&SS*, 291, 261

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.