Dear Dr. Endl,

We thank you for organizing the review of our paper, and also thank the referee for constructive criticism. We have considered the remarks of the reviewer carefully and revised our manuscript accordingly.

Below, we have reproduced the relevant portions of the reviewer's report, alongside our responses. The changes are bolded in the attached ms_boldchanges.pdf file, using the trackchanges AASTeX macros. The revised text is in the ms.pdf file.

Sincerely,

Luke Bouma

****
Reviewer #1 MODERATE CONCERNS

REVIEWER COMMENT

Figure 1:
- It is interesting to note that all four planets are concentrated in X-Z, vb-vl*, and vl*-l. This could be because of the small fraction of space that the Kepler field surveyed. Is that accurate? Is there any way to indicated which stars were observed by Kepler in these figures?
- Note that in B&W only the + symbol has a different color than the rest, although since the symbols are different this wasn't much of an issue. The exception is that I can't tell which group is magenta and which lime green, so specifying which are the larger filled circles and which are the smaller would be helpful.
- Why is it important to show the projection of the solar velocity?
RESPONSE

> - The Kepler field is shown with the large gray squares in the top panel of the figure, between l of 70 to 85, and b of 5 to 20. This approximately corresponds to a cone in XYZ -- in case it helps intuition, (l,b)=(90,0) corresponds to the line of X=Z=0. The resulting concentration of the planets in v_l* follows from the narrow range of galactic longitudes that Kepler looked at, and the narrow spread in v_l* at a given galactic longitude.
>
> - We now specify in the caption which are the larger and smaller filled circles.
>
> - We have briefly summarized in the first paragraph of section 2.2 why we have shown the projection of the solar velocity: it is because we are interested in how much of the spread in v_l* it might explain. The short answer is "some of it, but not all of it". To give a more complete explanation: the Sun moves at ~17 km/s toward (l,b)=(53deg,23deg), relative to a hypothetical circular orbit around the center of the galaxy. This means that in our observed frame, nearby

stars on average appear to move opposite this direction. At l=53 and 53+180 degrees, the effect does not influence v_l*. At l=53+90 degrees, the effect on v_l* is maximized, with amplitude = v_LSR = 17 km/s. The minimum is at l=53+270 degrees. Figure 11 of Zari, Brown et al 2018 provides an example for how much of the variance in observed v_l* can be ascribed to this projection effect for stars near the galactic plane. The lower-right panel of our Figure 1 however shows that while the LSR projection accounts for some of the change in observed v_l* in Cep-Her, the data and the gray band do not match! There therefore needs to be some intrinsic differences in the UVW velocities of these stars (though v_r, v_phi, and v_z may perhaps the better coordinate system).

REVIEWER COMMENT
\----------
If at all possible, please assure in Kerr+in prep that the nomenclature does not change such that CH-2 remains CH-2, or update this paper prior to publication.
RESPONSE
\----------
> The sub-group identifier will remain the same.

REVIEWER COMMENT
\----------
Line 248: I am skeptical of the few % false positive rate (later referenced as 1%). Once stars get hotter than K2, the young and field sequences overlap significantly and so field stars wouldn't likely be outliers in the CAMD. I'm also wondering if not all of the stars are shown on the CAMD; the text states 173 stars in RSG-5; it doesn't quite look to me like there are 173 stars on the upper left panel, but I could be wrong.
RESPONSE
\----------
> Thank you for this comment -- it has helped us assess the false positive rates in RSG-5 more closely.
>
> In the color versus absolute magnitude diagram (CAMD), let us use (BP-RP)0 = 1.5 as the dividing line, red-ward of which stars can be easily seen to be on the pre-main-sequence (or at least elevated relative to typical field stars of the same color). For RSG-5, there are 89 stars with (BP-RP)0 > 1.5. One of them falls on the "field star" locus; the others fall on the pre-main-sequence locus. This implies that, photometrically, 88/89= 98.9% of the candidate RSG-5 members red-ward of (BP-RP)0=1.5 are photometrically consistent with being on the pre-main-sequence.
>
> Regarding more massive stars, yes, there are plenty: see hr_rsg5.png attached to the response. The caption of Figure 2 already explicitly states "The range of colors is truncated to emphasize the pre-main-sequence." The manuscript also has a sentence "Finally, for the plots we set the color [axis] to best visualize the region of maximal age information content: the pre-main-sequence." But yes, we cannot determine whether these stars more massive than ~K2V are young based on their location in the CAMD. We can and do use stellar rotation.

>
> Using stellar rotation, we previously wrote in Section 2.3.2 that in RSG-5 41/47 stars have rotation periods consistent with being at least as fast as the Pleiades (87%). Of the six inconsistent stars, two are the M-dwarfs previously discussed in the text. The new point that emerged from our reanalysis concerns the other four stars without any detectable rotation period: three of them were the bluest three stars in the sample! This implies that our blue cutoff was too lenient -- the amplitudes of spot-induced modulation plummet for young stars blue-ward of the Kraft break (eg. Figure 10 in Rebull+2020, 2020AJ....159..273R). We have propagated this through the manuscript: we now require (BP-RP)0 > 0.6, and G < 16; the previous cutoff was (BP-RP)0 > 0.5. The corresponding fraction of RSG-5 stars for which detected rotation periods is 39 detections / 42 expected, i.e. 93%.
>
> The manuscript has been adjusted following this discussion, including the quantitative statements regarding the pre-main-sequence K and M dwarfs, as well as the rapidly-rotating FGK stars.

REVIEWER COMMENT

Figure 2:
- I don't see any squares for the field sample, is this just the background of small indistinct points?
- Very hard to see Praesepe in black and white
- Praesepe isn't very useful as a reference since the figure is considering things younger than Pleiades. However, it would be useful to see a younger comparison cluster more similar in age to Cep-Her.

RESPONSE

> Yes, the field sample is the gray indistinct background.
>
> We have updated the Praesepe markers to be thicker, which improves their visibility.
>
> The attached figure rotn_ic2602_ic2391_reference.png shows a concatenation from S. Douglas (private communication) of TESS-derived rotation periods from IC-2602 and IC-2391, which are both similar in age to Cep-Her/RSG-5. These are preliminary rotation period measurements, from the same Douglas+2021 work referenced in the manuscript, which is not yet published. The IC-2602 and IC-2391 rotation period distributions look quite similar to those from RSG-5, though their work a) extends to lower masses, b) appears to show somewhat more super-Pleiades rotation periods from F5 to G0 spectral types, and c) appears to show more stars on the fast sequence, which are lacking in the RSG-5 sample for unknown reasons. Given the unpublished nature of this comparison IC-2602 & IC-2391 sample, and the fact that our aim is to simply check for rotational consistency with the stars being at least as young as the Pleiades, we prefer to leave the reference to Douglas+2021 as it is and to not include these data in the manuscript.

REVIEWER COMMENT

Line 280, "Generally speaking, this method is expected": It is not clear that this method of interpolation should be accurate, since it relies on changes occurring linearly in time. Since the ages are quite close to one of the anchors it is probably fine in any case. Please provide a citation or more discussion.
RESPONSE

> We have added the following text:
> \added{While in reality stellar populations do not evolve linearly in the dimensions of absolute magnitude versus color, in our case the Cep-Her loci are nearly indistinguishable from IC\,2602 (e.g., Figure~3 of \citealtt{bouma_kep1627_2022}). Systematic errors incurred in the age from the non-linear evolution are therefore likely much smaller than the $\approx$10\,Myr systematic uncertainty in the absolute reference age for IC\,2602 itself \citep{david_ages_2015,randich_gaiaeso_2018}.

REVIEWER COMMENT

Overall, the discussion of rotation discrepancies in Sec 2.3.2 is not clear and has the feel of the authors trying overly hard to justify their argument (I do agree with the argument, however).
- It would generally be much easier to follow this section if Fig 2 contained a younger cluster for comparison, such as the two cited at the end of this section.
- After some reflection, I think the argument about the M dwarfs is that both very young M dwarfs still in the process of contracting, and moderately young M dwarfs that have just begun to spin down, can have periods of 10-12 days. I suggest simply stating this up front, and then only consider FGK stars in the subsequent discussion.
- Line 328: "no field star outliers" - but the authors have previously stated that there is one outlier.
- 13% seems like a lot of pole-on stars for random chance, and at these young ages the amplitudes should almost always be quite large. Rampalli+2021 detects a higher fraction of rotators amongst older Praesepe stars. It seems more likely that the false positive rate is similar to the other sub-associations considered.
RESPONSE

> - Regarding younger clusters for comparison, please see our response regarding Figure 2, which contains the comparison against IC-2391 and IC-2602, and our reasons for not including it in the manuscript.
>
> - Regarding M dwarfs, the suggestion to exclude them by cutting at say (BP-RP)0=1.84 (Mamajek's M0V main sequence boundary) would have the following effect. In RSG-5, it would omit 13 stars currently in the 42 star sample, including the one non-detection, and the one 12.7 day rotator. The net result would be that 28 stars in RSG-5 would have rotation period detections consistent with being younger than the Pleaides, out of 29 expected. In CH-2, this

cut would omit 10 stars currently in the 19 star sample. The net result would be that 8 stars in CH-2 would have rotation period detections consistent with being younger than the Pleiades, out of 9 stars expected. We would prefer to not make this cut, chiefly because of the factor of 2x reduction in the CH-2 rotation sample, which would weaken its statistical power more than necessary. In our view, the rotation periods of the early M-dwarfs are still informative, particularly in the sense that non-detections would imply inconsistency with the expected >1% amplitude spot-induced modulations (see again Figure 10 of Rebull 2020 2020AJ....159..273R). We have clarified the wording in the relevant paragraphs of Section 2.3.2 to emphasize the main points, which concern the range of plausible false positive rates based on the rotation period analysis. The contamination rates seem to be at most one in three for CH-2 candidate members, and one in fifteen for candidate RSG-5 members.
>
> - Regarding "nearly no field star outliers", which was the relevant phrase, the wording has been adjusted to "only one field star outlier".
>
> - Regarding 87/13%, please see the earlier discussion regarding the sensitivity of this fraction to our blue cutoff boundary, and the adjustment to (BP-RP)0 > 0.6, which yielded a 93% detection rate. Given the small amplitudes of rotation-induced signals in the bluer regime, this seems like the wiser approach. Based on the color absolute magnitude diagrams, the false positive rate in the candidate RSG-5 members is quite likely lower than in the candidate δ Lyr or CH-2 members.

REVIEWER COMMENT

The discussion of vsini measurements did not provide adequate details, such as macroturbulence and potential impact of the targets being very young.
RESPONSE

> We have added the following footnote to the manuscript:
>
> \footnote{The broadening is calculated using the joint rotational and macroturbulent broadening kernel from \citet{hirano_2011}, assuming that the macroturbulent velocity scales with effective temperature similar to the prescription from \citet{2014MNRAS.444.3592D}. The latter assumption could be a source of systematic uncertainty in our equatorial velocity measurements, since the macroturbulent velocity could be systematically higher (or lower) on the pre-main-sequence than it is for more slowly rotating field stars.}
>
> Relatedly, we have corrected the number of significant figures in the KOI-7368 vsini, and imposed an uncertainty floor of 1 km/s.

REVIEWER COMMENT

The discussion of Li could use further clarity.

- Identification of which stars have Li is inconsistent. Line 456 says neither component of 7913 has strong Li, but at line 477 Li is obvious in all but one star, and at line 487 Li is noted as secure in 7913A.
- Line 497: "Lack of Li" implies there's no Li, not that there's not as much as expected
- As I mentioned previously, I am skeptical of the 1% false positive rate for RSG-5
- The authors point out an interesting possibility about why the Li in 1643 could be low, namely that the dispersion could be higher than expected. Though I am not an expert in Li, this seems reasonable. However, I think it should be qualified a bit more: how many stars were previously studied; few enough that scatter could be missed? Why might RSG-5 be different?
- Did the authors consider initial Li abundance, or differences in Fe abundance, as a possible reason for differences in current Li? 1643 certainly stands out the most, but I might be able to argue from Fig 4 that on average, the observed planet hosts in Cep-Her are all a bit lower than expected from the reference sequences.

RESPONSE

> - The wording at L456 has been updated as follows: "The primary showed lithium in absorption with ${\rm EW}_{\rm Li} = 65^{+8}_{-6}$\,m\AA, while the secondary had a marginal detection of ${\rm EW}_{\rm Li} = 42^{+12}_{-19}$\,m\AA."
>
> - The wording at L497 has similarly been adjusted.
>
> - Regarding the 1% false positive rate, we have previously addressed that photometrically, 88/89= 98.9% of the candidate RSG-5 members red-ward of (BP-RP)0=1.5 are consistent with being on the pre-main-sequence. Based on rotation periods, the relevant fraction was 39/42 (93%). The detection process for stellar rotation periods is more complicated than that for measuring a star's brightness and color, so our initial wording of "$\approx 1\%$ false positive probability" seems accurate. We have nonetheless updated the text with the term "few-percent false positive probability". This does not change the conclusion, which is based on both the FPP from the spatio-kinematic selection, as well as the chance of randomly selecting an interloping field star at this color with a sub-Pleiades rotation period.
>
> - We have added a sentence quantifying the discrepancy: "Quantitatively, there are 14 reference stars within $\pm$150\,K of Kepler-1643. The mean and standard deviation of their lithium EWs is 255 ± 31\,m\AA, which implies that Kepler-1643 is 4.0-σ discrepant from expectations." We have also added sentences indicating that given these ~dozen stars, an anomalous Li EW scatter would indeed be surprising. However, in our view it remains the most plausible interpretation. The suggested alternative of differences in overall metallicity seems unlikely because the [Fe/H] abundances measured for the Kepler Objects of Interest are all near-solar.
>
> - In constructing this response, a typo was noticed and corrected: the manuscript previously referenced Randich+2018 for the lithium sample in IC-2602. The correct reference is Randich+2001. We did consider the Randich+2018 study, but the results were more

challenging to interpret because that study had less clearly stated uncertainties on the equivalent width measurements, and a selection function (their Figure 1b) that we did not want to impose.

REVIEWER COMMENT

Figure 4, H-alpha plot:

- Only the part of the plot showing M0 and earlier is really useful here, perhaps limit the axis range.
- It is difficult to distinguish the field stars and TucHor stars in black and white
- There must be more field K stars with H-a measurements for comparison; the two showing in the late K dwarf regime don't do a good job of illustrating how H-a inactive field K dwarfs are. It can be inferred from Rauscher & Marcy (2006), although I couldn't really find an appropriate reference for field K dwarfs since 1990 (https://iopscience.iop.org/article/10.1086/132624). However, https://academic.oup.com/mnras/article/476/1/908/4828379#112872390 includes hotter stars in numerous clusters.

RESPONSE

> - We have limited the axis range.
>
> - We have updated the markers to be more distinguishable.
>
> - Good point regarding the field comparison. We have replaced it for the Pleiades using the data from Fang+2018, as suggested.

REVIEWER COMMENT

The discussion of centroid offsets could use further elaboration.

- Line 664: how many sigma is 0.4"? Is this generally considered a good value?
- 1643 and 7368 both have centroid shifts that are larger than normal for Kepler, and it is stated that stellar variability is the likely cause. But not all young stars show centroid shifts, and presumably some level of centroid shift would still be concerning. Is the impact of stellar variability well established in the literature? Can the amplitude of stellar variability be used to inform one about the centroid offsets?
- The discussion for 7368 indicates that the centroid shift can't be due to another star due to the high contrast imaging constraints. Is there an established method for this? Can the same analysis be done for 1643?

RESPONSE

> For reference, the relevant Data Validation Summaries for the KOIs are available at the following links.

> Kepler-1643:
https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/008/008653/008653134/dv/kplr008653134-001-20160209194854_dvs.pdf
> KOI-7368:
https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/010/010736/010736489/dv/kplr010736489-001-20160209194854_dvs.pdf
> KOI-7913:
https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/008/008873/008873450/dv/kplr008873450-006-20160209194854_dvs.pdf
>
> - Re: 0.4" and "good values", \citet{bryson_2013} showed that for typical field star KOIs without centroid offsets, the mean offset distribution peaks at 0.3$''$ (their Figure 23). By comparison, stars with centroid offsets that can be localized to nearby stars have a distribution that peaks at 7$''$ (their Figure 32). The stellar variability in Kepler-1643 complicates the centroid-based vetting tests, because the shifts measured by these tests are determined from the in- and out-of-transit flux-weighted centroids. For stars with significant spot-induced variability there is no static baseline in either the in- or out-of-transit , and so the centroid location may shift depending on the rotational phase combined with the local scene. Nonetheless, it appears that the centroid-level diagnostics for Kepler-1643 are consistent with the transit signal being localized to the target star. We have modified the text to include these points.
>
> - Re: what level of centroid shift is concerning -- if the centroid shifts across quarters were consistently >~ 1 arcsecond (and measured at ~0.1 arcsecond precision), and in the same direction across many Kepler quarters, it would be concerning. The measured centroid shifts in Kepler-1643 and KOI-7368 are not that large, and appear to be in effectively random directions, which suggests that the stellar rotation and systematic errors in the NASA SPOC pipeline's whitening are more likely to be the root cause. A few sentences in the manuscript have been modified to this effect.
>
> - Re: methods for establishing whether other stars can create the centroid shift, and Kepler-1643, yes, the same logic applies. We have added a comment to the manuscript to this effect. The "established method" for this is the logic laid out by Bryson+2013 (which unfortunately excludes considerations of intrinsic stellar variability). For instance, given the target star T and a neighbor star N separated by distance D with known flux ratio, one can calculate the expected centroid shifts for the various scenarios where the target T hosts an eclipse signal of depth δ, and also when the neighbor N hosts the signal. We do not present these calculations in the text, because they would not improve the clarity, and there is no evidence that the offsets themselves are statistically significant. Nonetheless, the absence of directly resolved companions for Kepler-1643 and KOI-7368 does help bolster the interpretation.

****

Reviewer #1 CLARITY, CITATIONS, ETC.

REVIEWER COMMENT

Line 54, "To date...": Sentence meaning unclear. These analyses have focused on positions & velocities? Or have used clustering based on?
RESPONSE

> Sentence has been reworded: \replaced{To date these analyses have mostly clustered on stellar
positions and 2D velocities}{So far, these analyses have mostly leveraged 3D stellar positions and 2D on-sky tangential velocities.}.

REVIEWER COMMENT

Line 63: Unclear whether "larger" refers to size or kinematic spread, perhaps "more diffuse kinematically" if referring to kinematic spread.
RESPONSE

> Sentence has been reworded as suggested.

REVIEWER COMMENT

Line 70: Consider also citing Arancibia-Silva et al. (2020) which also looked at Li in Psc-Eri
RESPONSE

> Done.

REVIEWER COMMENT

Line 114: Consider adding citations for ages for delta Lyr and RSG-5.
RESPONSE

> The end of Section 2.1 already cites Kerr+21's ages for some of the subgroups. The relevant citations other than Bouma+2022 (Kepler-1627 / delta Lyr cluster) would be Kounkel & Covey 2019 and Roser+2016. The Roser+2016 reference is already made in the preceding sentence. The KC2019 reference we prefer to avoid in reference to the age based on the manual stitching step they used which yielded a very different "delta Lyra cluster" (Stephenson-1) than that originally described by Stephenson (1959). Given the existing citations to the relevant studies, these age citations are not needed.

REVIEWER COMMENT

Line 118: From the information given so far about delta Lyr and RSG-5, I do not know where to look in Zari+2018 figures 11 and 13, so this is not a very useful reference without additional information on the velocities and XY positions.

RESPONSE

> This Zari+2018 reference has been updated to also refer to our Figure 1, to help the reader know where to look.

REVIEWER COMMENT

Line 122: It is worth noting that Zucker+2022 found that many of the "strings" from Kounkel & Covey are not likely real associations. Obviously, some of the associations K&C identify are real, but the extended groups that are more similar to Cep-Her may more often be artifacts.

RESPONSE

> A reference has been added to Zucker et al's rebuttal.
>
> In more detail, the Kounkel & Covey 2019 study indeed included many false positives, both in the sense of "clusters" that do not correspond to stars that were ever physically associated, and also in the sense of clusters that do have some "true" core population, but many erroneous field star interlopers. Zucker et al's argument however was not particularly helpful for our purposes of identifying which are which. One reason is that for the youngest (<100 Myr) cluster candidates from KC2019, most of the HR diagrams independently show the expected pre-main-sequence M-dwarfs as well as massive stars (http://mkounkel.com/mw3d/hr.html). It is hard to fake such stellar populations. Nonetheless, Zucker+2022 did nicely demonstrate the extreme nature of some of the groups in e.g., XY space (https://faun.rc.fas.harvard.edu/czucker/Paper_Figures/String_Gallery_Interactive.html). And yes, for the extended groups older than a few hundred million years, there is almost no independent evidence for their coevality. Despite these points, for the youngest groups, the KC2019 study to us seems to be more likely an example of why the manual stitching must be treated cautiously, and also for why there are better dimensions to cluster in than those selected in the original KC2019 study.

REVIEWER COMMENT

Line 130, "identified three of the nearest subpopulations": subpopulations of Cep-Her?

RESPONSE

> Yes, clarification made.

REVIEWER COMMENT

Line 140: Figure 1 reference seems to be specifically to the lower right panel, not both lower panels.

RESPONSE

> The wording in this paragraph has been clarified.

REVIEWER COMMENT

Line 162: vb and vl not defined
RESPONSE

> Thank you for this catch -- v_b is the distance-corrected proper motion in the direction of increasing galactic latitude, and $v_{l^*}=v_l \cos b$ is the same in the direction of increasing galactic longitude after accounting for the local tangent plane correction. We have added verbiage to this effect.

REVIEWER COMMENT
----
*Line 171: core distance metric is not defined, but is quite important per the subsequent paragraph and so should be explained. It would also be helpful to state that larger values of D are better.
RESPONSE

> We have added the following clarifying sentence to the manuscript:
> \added{Core distance is the distance to the $k^{\rm th}$ nearest star, and therefore k acts as a smoothing parameter, where a larger value reduces the influence of local overdensities smaller than the scale that interests us.}

REVIEWER COMMENT

Line 174-175, "at least as close to the 10th nearest": Not clear what this means.
RESPONSE

> The relevant sentence has been reworded for clarity.

REVIEWER COMMENT

Line 188: I understand the notation for a normal distribution, but perhaps worth just writing in words instead. I also don't know why D is Gaussian in the first place, but this would presumably be explained by defining the core distance.
RESPONSE

> Thank you for this catch -- the wording in the original text had a subtle error. The revised wording is as follows.
>
> "To enable a selection cut that filters out field-star contaminants, we also compute a weight metric, D, defined such that the group member with the smallest core distance has $D=1$, the group member with the greatest core distance has $D=0$, and the metric D scales linearly between the two extremes. After applying a set of quality cuts on the astrometry and

photometry [...] this procedure yields a distribution of weights D that is well described by a log-normal distribution with \log_{10}\mathcal{N}(-1.55,0.61)$."
>
> The original wording incorrectly stated that the "weights for the other group members are log-normally distributed" between the two extremes, while instead the scaling is linear. The fact that the resulting distribution of weights is well-fit by a log-normal is an emergent property of the HDBScan clustering and our definition of D -- the attached hist_weight.png substantiates the claim that the distribution is "well-described" by such a functional form.

REVIEWER COMMENT
----
*Line 190-191: Please explain how thresholds of 0.02 and 0.10 were selected.
RESPONSE

> Done; these thresholds were selected visually based on the apparent purity with which they yielded pre-main-sequence stars on a color--absolute magnitude diagram.

REVIEWER COMMENT
----
*Line 192: vl' not defined
RESPONSE

> Thank you for the typo catch: v_l' was equivalent to v_l*, and has been relabelled as such. v_l* is also now defined (inre: Line 162 comment).

REVIEWER COMMENT

Line 209: The expected period and amplitude based on stars of what age?
RESPONSE

> For stars at least as young as the Pleiades -- words added.

REVIEWER COMMENT

Line 225: Phrasing makes it sound like Cep-Her candidates observed by Kepler are distinct from those in RSG-5 and CH-2, but I assume they overlap?
RESPONSE

> Yes, there is overlap. But the set of Cep-Her stars that Kepler observed is a distinct set from the described sets of candidate stars in RSG-5 and CH-2. The wording has been adjusted.

REVIEWER COMMENT

Line 235: May be useful to readers to specify which part of Gaia Collaboration + 2018 the data filters are listed.
RESPONSE

> We now include a reference to appendix B of Gaia Collaboration+2018.

REVIEWER COMMENT

Line 246: "set the color range" had me looking for a color map, suggest saying "color axis" instead
RESPONSE

> Done

REVIEWER COMMENT

Line 251: This struck me as an odd use of "control sample" since delta Lyra seems like just another example equivalent to RSG-5 and CH-2
RESPONSE

> We now call it a "reference sample". It is indeed another example, equivalent to RSG-5 and CH-2. The only difference is that it was studied in our earlier Kepler-1627 paper, where the explicit comparison against IC-2602 was made (Figure 3 of that work).

REVIEWER COMMENT

Line 253-254: If photometrically overlapping with the field, by definition a star must be an outlier from the cluster sequence, correct? Current wording suggests otherwise, suggest rephrasing.
RESPONSE

> Wording simplified, thank you.

REVIEWER COMMENT

Line 260: Specify overlap is on CAMD.
RESPONSE

> Done, thank you.

REVIEWER COMMENT

Line 359: Stars are referred to by spectral types, but then define ZAMS in terms of mass so it is not possible to compare without additional information about the spectral type of the sun on the ZAMS.

RESPONSE

> We do not entirely understand what, if any, change is requested: the purpose of these sentences is to give a broad-brush overview of the stellar properties. The implicit understanding that the Sun would be "near" G2V at ZAMS (while in reality perhaps being a little less bright, and a little lower surface gravity) provides the necessary context that the stars being discussed spanning "G8V to K6V" are less massive than the Sun, and given their ~40 Myr ages, on the late stages of the pre-main-sequence.

REVIEWER COMMENT
----
*Table 1: Please specify which component of the binary KOI 7913 orbits
RESPONSE

> A table note has been added.

REVIEWER COMMENT

Line 364: From later comparisons with the spectroscopy, it seems that the MIST analysis also returns Teff.
RESPONSE

> Yes -- please see the response to Line 449 below.

REVIEWER COMMENT

Line 365: Clarify what is interpolating against the MIST isochrones.
RESPONSE

> The interpolation is performed in the dimensions of reddening-corrected absolute G-band magnitude versus (BP-RP)0 color. This clarification has been added.

REVIEWER COMMENT

Line 372: Should be stated Teff is also adopted from spectra, not just used to verify previous values.
RESPONSE

> Done.

REVIEWER COMMENT

Line 373, "we acquired spectra": might as well specify they are high resolution optical spectra!
RESPONSE

> Done.

REVIEWER COMMENT

Line 392, citation to Chubak et al.: were these values determined following the methods of those authors, or by comparison to templates from that work?
RESPONSE

> Following their methods; the manuscript has been updated accordingly.

REVIEWER COMMENT

Line 394: Would be helpful to include a one-sentence description of Li measurement.
RESPONSE

> Done.

REVIEWER COMMENT

Line 413: Just wanted to confirm the 2015 date, presumably this was an archival spectrum from Kepler days.
RESPONSE

> Confirmed.

REVIEWER COMMENT
----
*Line 449: Teff values are only quoted for these two stars, and were not in previous sections, and the subsequent text also implies that for these stars, the spectra-based Teff values disagreed with the isochrone-based values. Either re-phrase, or discuss disagreement.
RESPONSE

> Thank you for this wording catch -- the text has been updated to clarify that the temperatures as well as the other spectroscopic parameters agreed with the isochrone parameters within 1-sigma.
>
> This is a rather brief way of hiding what in our view is more complicated than what needs to be discussed in the manuscript. Table 1 states the adopted (spectroscopic) effective temperatures for all the stars. We reported the spectroscopic effective temperatures in all cases because their formal uncertainties were smaller than those from the isochrone method, and they seemed less model-dependent. The isochrone method yielded the following effective temperatures for the stars in Table 1:
> Kepler-1643: 4878 +/- 140 K

> KOI-7368: 5135 +/- 162 K
> KOI-7913 A: 4286 +/- 144 K
> KOI-7913 B 4006 +/- 335 K.
> As discussed in the second paragraph of Section 3, the quoted values are from MIST, while the uncertainties are a quadrature sum of a statistical and systematic component (where the latter is from the absolute difference between the PARSEC and MIST isochrones). For KOI-7913 B, this systematic component is higher than for any other stars, because MIST gives Teff_MIST = 4005 +75 -73 K, while PARSEC gives Teff_PARSEC = 3679 +95 -60 K. This is most likely due to different assumptions in the MIST and PARSEC models for how to treat young M-dwarfs. The MIST models have known discrepancies between their synthetic photometry and observed open cluster data (e.g., Choi+2016 Section 8). Possible explanations include starspots, or incomplete molecular line lists (e.g., Mann+2013, Rajpurohit+2013). The PARSEC isochrones might be a better choice for the lowest mass stars (KOI-7913A and KOI-7913B), since Chen et al (2014) implemented an empirical temperature-opacity calibration into the PARSEC models to create synthetic photometry that matches available data. See for instance "hr_parsec_7913b.png" and "hr_mist_7913b.png" attached -- the 39.8 Myr MIST model diverges from the IC2602 and delta-Lyr loci at BP-RP>~1.6. In comparison, the 39.8 Myr PARSEC model fits the observed locus down to BP-RP~=2.0.
> Three additional plots are also included in the response: smemp_koi7913B.png, smemp_koi7913A.png, and smemp_kepler1643.png. These plots give some diagnostic understanding of how SpecMatch-Empirical performs for these stars. Note that a chi^2 of 5 corresponds to a reduced chi^2 of 1 in this model. While the KOI-7913B spectrum is still the worst fit(!), our preference is for calibrating our reported effective temperatures to the HIRES spectral library, rather than to the isochronal method. Any suggestions on possible methodological improvements would be appreciated.

REVIEWER COMMENT

Line 471: "amusing" ?
RESPONSE

> Depending on one's sense of humor.

REVIEWER COMMENT

Line 472: mass of X - at the assumed age of 40Myr?
RESPONSE

> Yes -- we've added this to the manuscript to make it explicit.

REVIEWER COMMENT

Section 3.4.2.: This section makes a big deal out of H-a self-reversals, but to my knowledge these are well-known and expected since the 80s. Admittedly the references I can think of are

all for M dwarfs, but I wouldn't expect a late K dwarf to be fundamentally different than an early M. The section also implies it's a possible (though ruled out) concern for binarity, but since no other spectral lines in Fig 3 are doubled this does not need to be a concern.
RESPONSE

> Enough of our team did not know about these H-alpha self-reversals that it merited some internal discussion, which yielded the second paragraph of text, including the comparison to proxima Cen and the references to NLTE models (Short & Doyle 1998) explaining the effect. Sometimes things that are "well-known" are worth repeating! Regarding binarity, yes, the first paragraph of Section 3.4.2 makes the point about double-lined binarity being ruled out more directly through analysis of the CCFs between KOI-7913A/B and the HIRES template matches. This is a stronger test than the visual inspection of the Li and Fe lines in the upper-right panels of Figure 3.

REVIEWER COMMENT

Line 618: for a GP regression, it is necessary to remove outliers, so this outcome is expected.
RESPONSE

> Thank you for this point -- it is true that we did not pre-mask the flares in our exploratory GP analysis, and so they may have contributed to the overfitting that we encountered. This does seem qualitatively surprising, since "flare-contaminated" data comprise ~0.1% of the 30-minute cadence data points (e.g., Appendix D of Bouma+22, AJ, 163, 121 found a total duration of 6.5 hours of flares over the single 90 day quarter of available 1-minute cadence data). Nonetheless, we've adjusted the discussion to omit the claim that fine-tuning of the hyperparameter priors would be strictly necessary, and instead focus on the reasons that we ultimately adopted our simpler model.

REVIEWER COMMENT

Figure 5: "planetary transit times are indicated" add "by dashed vertical lines"
RESPONSE

> Done, thank you.

REVIEWER COMMENT

Line 678: "clears the usual threshold" - what is the usual threshold?
RESPONSE

> We've updated the sentence as follows:
> "Though not as convincing as Kepler-1643, this clears the \replaced{usual threshold for calling the planet statistically validated \citep{morton_efficient_2012}}{threshold probability of 1 in 100 suggested by \citet{morton_false_2016} for calling a planet statistically validated}."

> Note the updated citation is to Morton+2016, rather than Morton 2012. The relevant portion from Morton 2012 that we had previously considered (Section 3) has important discussion about this boundary, but the Morton+2016 citation is more relevant as the practical adoption of the 2012 method.

REVIEWER COMMENT

Line 682: Please define positional probability.
RESPONSE

> We have added a footnote pointing to the relevant columns at the NASA Exoplanet Archive, which contains more extensive documentation on this calculation (as does the referenced memo by Bryson & Morton 2017). We have also reworded the text in favor of a plain-english interpretation, which is that the probability they calculated indicates that the transit signal shares its position with the target star.

REVIEWER COMMENT

Line 707: The FPP is only barely lower than 7368.
RESPONSE

> Good point -- the relevant text has been reworded to omit the comparison against 7368.

REVIEWER COMMENT

Line 710: Please define probability usability score.
RESPONSE

> We have reworded the relevant sentence to emphasize the interpretation: the positional probability calculation performed by Bryson & Morton yielded a near-unity probability that the transit event is at the same location as the host star. We have omitted discussion of the "usability score", since the discussion in the final sentences of Section 4 provide more salient information than this metric -- the pixel level data show two distinct stellar rotation signals from the two marginally resolved stars, and no evidence for the transiting planet signal in the pixels that show the KOI-7913B rotation signal.

REVIEWER COMMENT

In the discussion in 5.2, it may be worth noting that in recent years difuse associations have become more common. For example, what is the spatial density of Psc-Eri?
RESPONSE

> We have added the following sentences in paragraph 3: \added{For instance, the Psc-Eri stream, which has a shape that can be approximated as a 600 parsec-long cylinder with a

radius of 30 parsecs, has a number density roughly a factor of three times lower than even CH-2 \citep{roser_psceri_2020}. However its existence is discernible because of the $\lesssim 2.5$\kms\ scatter in its cylindrical velocities. }

> This estimate imposed a G<16.5 magnitude cut to the 1387 Psc-Eri members from Roser & Schilbach 2020, which yielded 858 stars. The shape of the stream was approximated as a 600 pc long cylinder, with radius 30 pc, and tangential velocity radius 2.5km/s, based on the figures and discussion in that analysis. The resulting relative number density is as quoted.

REVIEWER COMMENT

Line 804: Would "association" be a useful term here?
RESPONSE

> The term is used at a few points throughout Section 5.2 already, and it is not clear that its usage at this exact line would necessarily add anything that it hasn't already (e.g., in the first paragraph, or second-to-last paragraph, of Section 5.2).

REVIEWER COMMENT

Line 847: I am not aware of precise RVs coming from low resolution optical spectra for low mass stars. Perhaps for higher mass stars they would be useful, but in these cases I imagine DR3 will provide.
RESPONSE

> We have added the qualification that LAMOST and SDSS-V spectra would likely only be useful for velocities for the brightest (higher mass) members.

REVIEWER COMMENT

Line 849: I realize this is my fault since this report is quite delayed, but DR3 now is released and it should be possibly to determine (broadly, I am certainly not asking this paper to look at the DR3 data) whether DR3 RVs are going to be important.
RESPONSE

> We have reworded these sentences to point out that DR3 contains reported velocities down to G_RVS of 14, or spectral types of ~K5V.

REVIEWER COMMENT
----
*Line 925: "come from its binary" add "companion" but it would also be helpful to remind readers which component the KOI orbits.
RESPONSE

> Thank you for this catch. "binary" was a typo, and has been replaced by "binarity". We have added the phrase "transit-hosting" when describing the primary for the suggested reminder.

REVIEWER COMMENT

Table 4: is the median value the adopted value?
RESPONSE

> Posterior values quoted in the text are means and standard deviations for symmetric distributions, and are otherwise medians bracketed by the upper and lower 84.1 and 15.9 percentile deviations. We have added a table note to this effect.

REVIEWER COMMENT

Line 1182: I'm confused by the period(s) for this star. "The L-S favored 6.67 days..." L-S as applied to what data? Both the ZTF period and the TESS period are quoted subsequently. Perhaps I am confused by the concept of "double dipper" but it seems unlikely for spots to make a signal appear at 2P in the TESS light curve.
RESPONSE

> Thank you for this suggestion -- we've reworded this sentence for clarity.

****

REVIEWER COMMENT

Apologies to the authors for the very tardy report, I was sick for several weeks.
RESPONSE

Thank you for the helpful report! Health is more important than science. I hope everything is OK.

****

Other minor changes made to the manuscript include:

* The radius of Kepler-1643b was previously stated in the abstract to be 2.32 +/- 0.14 R_earth. This was a typo, it is +/- 0.13 R_earth.

* We added the following text as a note to table 2: "\added{Users who wish to minimize field star contamination should apply more restrictive weight cuts, {\it e.g.}, $\texttt{weight}>0.1$.}"

* We reduced the axis width on all figures for visual clarity.