We thank the referee for the very careful and considered report. The comments were all very helpful and will lead to significant improvement in the paper. We apologize for the length of time it has taken us to reply to this report.

☑1. The light curve analysis is based on PDC-msMAP data. The paper states that this preserves stellar variability but that is not true. PDC-msMAP has been shown to suppress variability on timescales longer than 20 days entirely and partially for timescales between 1 and 20 days. This is shown quantitatively in Figure 1 of the Kepler Data Release notes 21. Specifically, it shows that the amplitude of signals on timescales of ~6 days (the orbital period of the planet) is corrupted by about 5%. The authors should discuss how this might affect their conclusions in terms of the out-of-eclipse variability model and the albedo measurement given in Section 5.

The is an excellent point brought up here, something we should have considered given the lead author worked on PDC. The astrophysical signals with timescales in the range 3-6 days and will be damped by a few percent are (a) the reflection/thermal emission from the planet and (b) the ellipsoidal variations from the star.

For case (a), reflection/emission from the planet is not significantly detected. Even were the reflection not to be damped by a few percent it would still not be detected.
For case (b) we measure an amplitude of 52 +/-5 ppm - i.e. an uncertainty of 10%. Therefore, a damping of a signal by a few percent would not significantly affect our measurement this signal. However, it is a systematic offset. We added a discussion of the damping for the signal and modify the text where we introduce PDC-msMAP.

The geometric albedo we derive for the planet is not affected by this affect because it relies on the secondary eclipse depth which only last a few hours.

☑ 2. In section 2.3 the exposure times used for the spectra should be given. Could the stellar oscillations affect the RV measurements? This needs quantifying.

We have included this information and discussed how oscillation affect RV measurements - they don't significantly change things.

☑ 3. In section 3.2, it might be a good idea, when explaining why the squared exponential kernel was chosen, to state that it is adequate for modelling smooth functions, and to discuss why that might be appropriate for the sort of correlated noise expected here.

We have added a discussion to why we used a squared exponential kernel. In essence we argue that this kernel is well matched to the noise expected from red giant granulation.

☑ 4. There is an apparent inconsistency between table 3 and the text of section 3.3: the text says 1/e priors were used but the table shows uniform priors on ecosw and esinw. This needs to be resolved, and I hope it was the text that was correct!

We have updated the table. The text was correct.

☑ 5a. Strictly speaking, one should not use uniform priors on scale quantities such as the photometric and RV measurement uncertainties and the amplitude and timescale of the GP kernel, but rather modified Jeffreys priors. Given the fairly small uncertainties reported for these parameters in Table 4, I doubt that the choice of prior had a very strong impact, but it would be better to be in a position to state this unambiguously.

We have resampled the data to use a Jeffery's prior. This did not change out results significantly. We have updated the text and table to make it clear we used a Jeffery's prior on the modeled uncertainties.

☑ 5b. I strongly recommend including (probably online only) plots of the MCMC chains, posterior distributions for all parameters, and 2-D correlation plots for any pairs of parameters that show evidence of correlation, so the reader can ascertain the convergence of the chains, the asymmetry or otherwise of the distributions, and the degeneracies between parameters.

We have included an appendix with additional plots showing the results of our MCMC sampling.

☑ 6. On Figure 2, the GP component of the model should be shown together with a 68% or 95% confidence interval (e.g. using a shaded area). This interval can be estimated with or without the white noise contribution - just state which is the case.

The extent of the confidence interval is now shown in the figure. We opted to show the mean noise model but show the combined noise+transit model as shaded. This seemed the best way to convey the information without overcomplicating the plot.

☑ 7. In section 4, at the beginning of page 4, the authors state that they were able to produce a self consistent model that fits the data well. This needs to be quantified for example by quoting the reduced chi-squared of the best-fit and, ideally, the distribution and power spectrum of the light curve residuals (after subtracting the best-fit transit + GP model).

We have included plots showing a histogram of residuals and the power spectral density of the data before/after the GP model is removed. Low frequency noise is heavily suppressed by the GP model.

☑ 8. The statement that the highest probability model is "the one to use if you require a single model" is correct but somewhat confusing. It would be good to clarify that this is the case if, for example, one wants to incorporate this model in additional modelling, but if one wants to draw conclusions about the physics properties of the planet, it is the median that one should use.

We included the self-consistent solution after repeated complaints from colleagues that our previous work did not allow them to, for example, predict planet formation parameters using an N-body code.

We have clarified the text to make it more obvious to the reader which estimates to use and gently chastised them for reading too much into point estimates.

☑ 9. The authors state their estimate of Rp/Rstar is inconsistent with that of Sliski & Kipping, by how many sigma?

We have included in the text an estimate of the disagreement with Sliski & Kipping. It is 2.5 sigma.

☑ 10. On figures 1, 4 and 5 there is no need to show two orbits. Showing two orbits exaggerates the periodicity of the data. One orbit only - with a little bit either side if desired - is plenty.

The figures 4 and 5 have been modified as requested (and combined into a single figure). We opted to keep figure 1 as is because we are emphasizing that no correlation with orbital period is observed. Showing two orbits helps to demonstrate this.

☑ 11. There should really be a figure - online only if needed - showing the RV data as a function of time as well as phase-folded. I presume there is already a plot of the unfolded light curve in one of the previous papers on this object, if not (or if it doesn't cover the full Q1-Q17) this should be included as well.

We have included both the additional requested figures as a single figure - now figure 1.

☑ 12. When comparing your posterior for the stellar density to the estimate of Sliski and Kipping, you really should have a plot of this posterior (comparing your results with the Sliski & Kipping version, and with the asteroseismic constraints).

We have included the plot. Our results mirror the asteroseismic density strongly.

☑ 13. I don't see the relevance of the paragraph where the authors repeated the analysis without the RV data. The radius ratio constraint is not driven by the RV data. The only circumstances in which it might be would be if the RV data contained strong evidence for high eccentricity but this is not the case here.

We have removed this discussion and replaced it with the discussion below.

If you want to conclusively show that your suspected explanation for the discrepancy between the Sliski and Kipping results is the right one, the thing to do is to repeat your analysis without the GP term of your model. Presumably you would get a different distribution for Rp/Rstar to what you got with the GP (perhaps more consistent with that obtained by Sliski & Kipping, but perhaps not, since the correlated noise is a stochastic process), and that would indeed demonstrate that it is the inclusion of this term that made the difference. Then one has to ask if including this term is justified, and this can be done using the residuals of the best-fit transit + OOT model (no GP): if it shows excess noise for the star's magnitude and the power spectrum or the auto-correlation function shows evidence for correlated noise, as seems very likely just from looking at Figure 2, then the inclusion of the GP term is justified and it is reasonable to conclude that your results are more robust than those of Sliski & Kipping.

An excellent suggestion! We performed this analysis and found a higher stellar density than the asteroseismic density. We were able to reproduce the results of Sliski and Kipping. The power spectrum of the star shows significant correlated noise (indeed it looks near identical to the plot PSD plot shown in Figure 7 in grey).