A&A 620, A127 (2018)
https://doi.org/10.1051/0004-6361/201833514

Gaia Data Release 2

Astronomy & Astrophysics
Special issue

Gaia Data Release 2

Validating the classification of RR Lyrae and Cepheid variables with the *Kepler* and K2 missions⋆

László Molnár[1,2], Emese Plachy[1,2], Áron L. Juhász[1,3], and Lorenzo Rimoldini[4]

[1] Konkoly Observatory, MTA CSFK, 1121, Budapest, Konkoly Thege Miklós út 15-17, Hungary
[2] MTA CSFK Lendület Near-Field Cosmology Research Group, 1121, Budapest, Konkoly Thege Miklós út 15-17, Hungary
e-mail: molnar.laszlo@csfk.mta.hu
[3] Department of Astronomy, Eötvös Loránd University, 1117, Budapest, Pázmány Péter sétány 1/a, Hungary
[4] Department of Astronomy, University of Geneva, Chemin d'Ecogia 16, 1290, Versoix, Switzerland

Received 28 May 2018 / Accepted 22 October 2018

ABSTRACT

Context. The second data release of the *Gaia* mission (DR2) includes an advance catalogue of variable stars. The classifications of these stars are based on sparse photometry from the first 22 months of the mission.

Aims. We set out to investigate the purity and completeness of the all-sky *Gaia* classification results with the help of the continuous light curves of the observed targets from the *Kepler* and K2 missions, focusing specifically on RR Lyrae and Cepheid pulsators, outside the Galactic bulge region.

Methods. We cross-matched the *Gaia* identifications with the observations collected by the *Kepler* space telescope. We inspected the light curves visually, then calculated the relative Fourier coefficients and period ratios for the single- and double-mode K2 RR Lyrae stars to further classify them.

Results. We identified 1443 and 41 stars classified as RR Lyrae or Cepheid variables in *Gaia* DR2 in the targeted observations of the two missions and 263 more RR Lyre targets in the full-frame images (FFI) of the original mission. We provide the cross-match of these sources. We conclude that the RR Lyrae catalogue has a completeness between 70–78%, and provide a purity estimate of between 92 and 98% (targeted observations) with lower limits of 75% (FFI stars) and 51% (K2 worst-case scenario). The low number of Cepheids prevents us from drawing detailed conclusions, but the purity of the DR2 sample is estimated to be about 66%.

Key words. stars: variables: general – stars: variables: RR Lyrae – stars: variables: Cepheids

1. Introduction

RR Lyrae stars are large-amplitude pulsating stars with easily recognisable light-curve shapes. They can be used to measure distances, to trace structures in the Milky Way and other galaxies, and to investigate stellar pulsation and evolution. Large sky surveys have mapped the distribution of RR Lyrae stars in particular to large numbers and depths (see e.g. Drake et al. 2013; Sesar et al. 2013; Beaton et al. 2016; Hernitschek et al. 2016; Minniti et al. 2017, for some examples).

Cepheids, an umbrella term used to describe the classical δ Cephei stars, the Type II Cepheids, and the anomalous Cepheids, are even more important tools for mapping the structure of the cosmos and individual galaxies: although they are less frequent than RR Lyrae stars, they are more luminous, and hence more easily detectable. The OGLE survey has very thoroughly searched both the Galactic bulge and the Magellanic Clouds for all members of the Cepheid and RR Lyrae families. According to Soszyński et al. (2017), their collection of classical Cepheids is now nearly complete, concluding the work started a century ago by Leavitt (1908).

The *Gaia* Data Release 2 (DR2; Gaia Collaboration 2018) represents a new step forward in mapping the distribution of RR Lyrae and Cepheid stars, as well as other variable stars in the Milky Way and its vicinity. However, these classifications are based on sparse photometry, and those classifications can be affected by poor phase coverage and/or confusion between variable types.

Space photometric missions, such as MOST, CoRoT, *Kepler,* or BRITE, provide dense, continuous photometry of variable stars that might in principle be used to validate the results from other surveys. However, most missions have either small apertures or very limited sample sizes or both, therefore their use for validation purposes is very limited. The only exception is the *Kepler* space telescope, which observed hundreds of thousands of targets so far, including thousands of RR Lyrae and hundreds of Cepheid stars (Szabó et al. 2017; Molnár 2018). Our preliminary studies with the PanSTARRS PS1 survey by Hernitschek et al. (2016) indicate that a comparison with data from *Kepler* is feasible, but the first results indicate that eclipsing binary stars may contaminate the RR Lyrae sample in the Galactic disc in that survey (Juhász & Molnár 2018).

⋆ Full Tables A1, A4, and A5 are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/620/A127

These findings led us to use the observations of *Kepler* to provide an independent validation of the classifications of the various RR Lyrae and Cepheid-type variables in *Gaia* DR2. Conversely, input from *Gaia* DR2 will allow us in the future to correctly identify as RR Lyrae or Cepheid stars targets that hide in the *Kepler* and K2 databases.

The paper is structured as follows: in Sect. 2 we introduce the variable star observations made by *Gaia* and *Kepler*, respectively; in Sect. 3 we describe the classification of the *Kepler* light curves into the various subclasses; in Sect. 4 we present the results of the validation, and in Sect. 5 we provide our conclusions.

2. Observations

2.1. Variable stars in Gaia DR2

Holl et al. (2018) summarized the results for the more than half a million variable stars published in *Gaia* DR2, which include candidates of RR Lyrae stars, Cepheids, long-period variables, rotation modulation (BY Draconis) stars, δ Scuti & SX Phoenicis stars, and short-timescale variables.

Different techniques were employed in the variability pipeline depending on the variability type. Here, we focus on the Cepheids and RR Lyrae stars identified by the all-sky classification (Rimoldini et al. 2018), available in the `gaiadr2.vari_classifier_result` table of the *Gaia* archive. Such classifications are normally verified by a subsequent pipeline module dedicated to the RR Lyrae and Cepheid variables (Clementini et al. 2018), but only for sources with at least 12 field-of-view (FoV) transits in the G band. The variable classes that we use are the four RR Lyrae subtypes, RRAB, RRC, and RRD/ARRD (corresponding to fundamental-mode, first-overtone, and double-mode stars with either normal or anomalous period ratios), and various Cepheid types, CEP, T2CEP, and ACEP (corresponding to classical or δ Cephei stars, Type II Cepheids, and anomalous Cepheids). Throughout the paper, we refer to the variable classifications in *Gaia* DR2 in all capitals (RRAB, RRC, RRD), and use the regular RRab, RRc, RRd notations for our classifications based on the *Kepler* and K2 light curves.

Our goal is to provide an independent assessment of the completeness and purity rates of the all-sky classification results of RR Lyrae and Cepheid candidates based on the fine light-curve sampling of *Kepler* and K2 targets, regardless of the number of observations and other features or limitations related to the *Gaia* data. We define completeness here as the fraction of sources with properly identified classes in *Gaia* DR2 against all RR Lyrae stars observed in the *Kepler* and K2 missions. Purity is defined here as the fraction of sources with properly identified variable classes compared to all sources in those classes in *Gaia* DR2.

Kepler is superior in terms of photometric precision and duty cycle to all ground-based surveys. The OGLE survey comes close as a result of the decade-long coverage, but it is limited to the Galactic bulge, disc, and the Magellanic Clouds (Udalski et al. 2015), whereas *Kepler* observed several halo fields. We note that Holl et al. (2018) presents some completeness and purity estimates based on the OGLE survey results for the RR Lyrae and Cepheid candidates. The only other program similar to *Kepler* is the Transiting Exoplanet Survey Satellite space telescope (*TESS*; Ricker et al. 2016): it has already started collecting continuous photometry from nearly the whole sky, but it will not reach the depth of either *Gaia* or *Kepler* or ground-based surveys.

In this study, we used the following information available in the *Gaia* DR2 archive (fields and table names specified in footnotes): celestial coordinates[1], median G brightnesses[2], the number of FoV transits in the G band[3], classification classes[4] and scores[5] ranging between 0 and 1, provided by the all-sky classification pipeline (Rimoldini et al. 2018).

2.2. Kepler and K2 missions

The *Kepler* space telescope was designed to detect transiting exoplanets and determine the occurrence of Earth-like, temperate rocky planets around Sun-like stars (Borucki et al. 2010; Borucki 2016). To achieve this, it collected quasi-continuous photometry of approximately 170 000 stars in a field between Lyra and Cygnus, spanning about 115 deg^2. The majority of stars were observed with a 29.4 min sampling, called long cadence (LC), and the telescope was rolled by 90 deg every quarter year. Not all stars were continuously observed: the target list was updated for every quarter. The prime mission lasted for four years, until the breakdown of two reaction wheels on board.

Afterwards, *Kepler* was reoriented to observe in shorter, 60- to 80-day-long campaigns along the ecliptic, using only the two remaining reaction wheels. The new mission, named K2, had no core science program, and the space telescope has been utilized as a space photometric observatory instead, carrying out a very broad science program that extends beyond exoplanets and stellar physics (Howell et al. 2014). Coordination and target selection is managed by the *Kepler* Guest Observer Office[6]. Stellar physics studies are coordinated by the *Kepler* Asteroseismic Science Consortium (KASC), which has a dedicated RR Lyrae and Cepheids Working Group.

During the original mission, *Kepler* observed about 50 RR Lyrae stars, one classical Cepheid, V1154 Cyg, and a Type-II Cepheid of the RVb subtype, DF Cyg (see e.g. Benkő et al. 2010; Nemec et al. 2013; Bódi et al. 2016; Derekas et al. 2017; Vega et al. 2017, and references therein). The K2 mission greatly expanded the spatial coverage of the telescope, and propelled the numbers of observed RR Lyrae and Cepheid stars into several thousands and hundreds, respectively, greatly exceeding the observations of previous space photometric missions (Szabó et al. 2017).

An important aspect of the *Kepler* space telescope that sets it apart from other space photometric missions is the large optical aperture (95 cm), allowing us to reach targets fainter than Kp ~20 mag (where Kp refers to brightness in the wide passband of *Kepler*). In case of RR Lyrae stars, this translates into the distance of the nearest dwarf galaxies (Molnár et al. 2015). *Kepler* is the only space telescope that is able to deliver continuous time-series photometry while matching the depth of *Gaia* DR2. Ground-based, deep, but sparse surveys such as Catalina (Drake et al. 2009) and PanSTARRS PS1 (e.g. Chambers et al. 2016) can also be useful to identify faint RR Lyrae candidates and have been used to select targets for the K2 mission. These surveys lack the potential of continuous light curves to unambiguously distinguish pulsating stars from other sources, however, like eclipsing binaries that could be confused with them in sparse data sets

[1] `ra` and `dec` from `gaiadr2.gaia_source`
[2] `median_mag_g_fov` from `gaiadr2.vari_time_series_statistics`
[3] `num_selected_g_fov` from `gaiadr2.vari_time_series_statistics`
[4] `best_class_name` from `gaiadr2.vari_classifier_result`
[5] `best_class_score` from `gaiadr2.vari_classifier_result`
[6] http://keplerscience.arc.nasa.gov

(Juhász & Molnár 2018). We exploit the properties of the *Kepler* sample mentioned above (depth, coverage, and FoV size) to verify and validate the classification of RR Lyrae and Cepheid variables in *Gaia* DR2.

It is important to note that only pre-selected targets have been observed in the *Kepler* and K2 missions. Data storage and download bandwidth limitations meant that only a low percentage of pixels were used to gather data in any given observing quarter or campaign. Full-frame images (FFI) were rarely stored, but a sparse sample is available for the original mission[7]. In the K2 mission, targets were exclusively proposed through the Guest Observer program. RR Lyrae candidates from large sky-surveys were revised and the weakest candidates were not proposed (Plachy et al. 2016). The target list was then cut by the Guest Observer Office, especially in the early campaigns. Therefore, the observed sample is neither a complete sample of the potential RR Lyrae stars nor fully representative of the populations of stars within a FoV. However, these data still represent the largest sample of space-based photometry available to us. At the time of this study, data from Campaigns 0 to 13 were processed and released.

The resolution of *Kepler* at 4"/px is much poorer than that of *Gaia*. This could potentially lead to source confusion. However, RR Lyrae and Cepheid light curves are easily recognisable even if a target is blended with another star nearby. Therefore, if photometric variation was detected at the target coordinates, we accepted it as a confirmed variable. Multiple stars of the same variable type could still be confused, but we found no such examples. Given the limited angular resolution, we decided to avoid the Galactic bulge areas. The OGLE survey provides superior coverage for classification and validation purposes in the bulge, and was already used in this manner by Holl et al. (2018). The K2 Campaigns 9 and 11 (C9 and C11) targeted the Galactic bulge: we omitted C9 entirely, and only used stars on C11 below –6 deg Galactic latitude, that is, excluding the region covered by the OGLE survey. Only a small fraction of the OGLE RR Lyrae targets were included in the K2 observations in C11, and their inclusion would not have been representative for the bulge population. We note that even with these cuts, blending and confusion hindered the detection and classification of some targets in C11 and C7; the latter targeted the Sagittarius (Sgr) stream and the outskirts of the bulge.

3. Target identification and classification

We cross-matched the *Gaia* DR2 sources that were classified as RR Lyrae or Cepheid variables around the original *Kepler* Lyra-Cygnus field with the *Kepler* Input Catalog (KIC, Brown et al. 2011). The results of the cross-match are presented in Appendix A. We did not use the DR2 parallax and $G_{BP} - G_{RP}$ colour data for classification and relied only on the light-curve shape information provided by *Kepler*. For the original mission, we searched for observed targets and classifications within the literature (Benkő et al. 2010, 2014; Nemec et al. 2013; Moskalik et al. 2015) or visually inspected the light curves available at MAST[8].

During the original *Kepler* mission, 52 FFIs were recorded: 8 during commissioning, and a further 44 during the mission, each taken before the monthly data downlink period. The integration

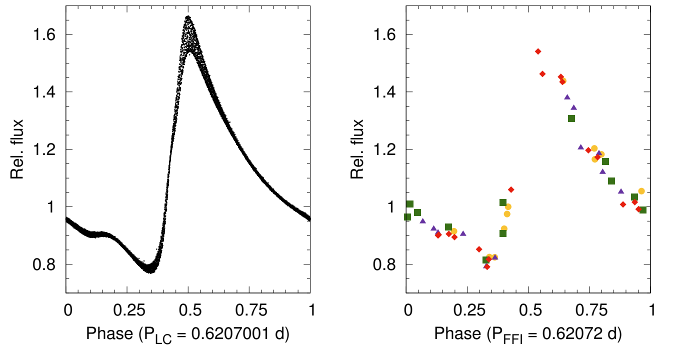

Fig. 1. *Left panel*: phased long-cadence light curve of a modulated RRab star, KIC 5559631, from the original *Kepler* mission, using the customised light curve of Benkő et al. (2014). *Right panel*: light curve from the full-frame image of the same star, extracted by the `f3` code (Montet et al. 2017). Pulsation periods were determined independently for the two data sets. Different colours and symbols denote data from different CCD modules.

time was the same as for LC exposures. Although these images only provide sparse photometry, they were obtained at completely different epochs with respect to the *Gaia* observations, and are numerous enough to provide acceptable phase coverage and a more representative stellar sample that is not affected by selection bias. We used the `f3` code developed by Montet et al. (2017) to extract the photometry of the stars, then folded the light curves with the most likely period based on the FFI data. A comparison of LC and FFI light curves of an RRab star is illustrated in Fig. 1, which also shows the agreement of the periods recovered from the two light curves. We compared the FFI periods to the periods determined from normal light curves for all RR Lyrae-type stars that *Kepler* observed, and they agree in all cases. The low number of data points in FFI light curves effectively reduces the faint limit of this data set to $G \sim 20$ mag.

For the K2 observations, we selected the sources near the campaign fields from the *Gaia* DR2 RR Lyrae and Cepheid classifications. The selection of RR Lyrae candidates is described in this section, while the one related to Cepheids is discussed in Sect. 4.4. Cross-match results for both groups are listed in Appendix A. We used the `K2FoV` tool to determine which stars fell on the CCD modules of *Kepler* in each campaign (Mullally et al. 2016). Overall, 11 361 stars were potentially observable during the selected campaigns, that is, about 5% of the RR Lyrae variables identified in *Gaia* DR2. Most of these are near the bulge or the Sgr stream (9843 stars for Campaigns 2, 7, and 11), and only 1518 fell into the halo FoVs. We cross-matched these sources with the K2 Ecliptic Plane Input Catalog (EPIC) and with the list of targets selected for observation in the mission (Huber et al. 2016).

The K2 observations contain various systematics caused by the excess motion of the space telescope, and multiple solutions were developed to correct for these. We generated an initial classification list by selecting the best light curves created by the various photometric pipelines. For the majority of the stars, we used the official pipeline-produced pre-search data conditioned simple aperture photometry (PDCSAP) products (Stumpe et al. 2012). Where PDCSAP was not available or was of inferior quality, we used data from the K2 extracted light-curves (or K2SFF), the EPIC Variability Extraction and Removal for Exoplanet Science Targets (EVEREST), and, in a few cases, the K2 Planet candidates from OptimaL Aperture Reduction (POLAR) light curves (Vanderburg & Johnson 2014; Luger et al. 2016; Barros et al. 2016). For a few stars, none of the light curve

[7] 52 FFIs were recorded in the original *Kepler* mission, and one or two per campaign in the K2 mission.
[8] Mikulski Archive for Space Telescopes, `http://archive.stsci.edu.`

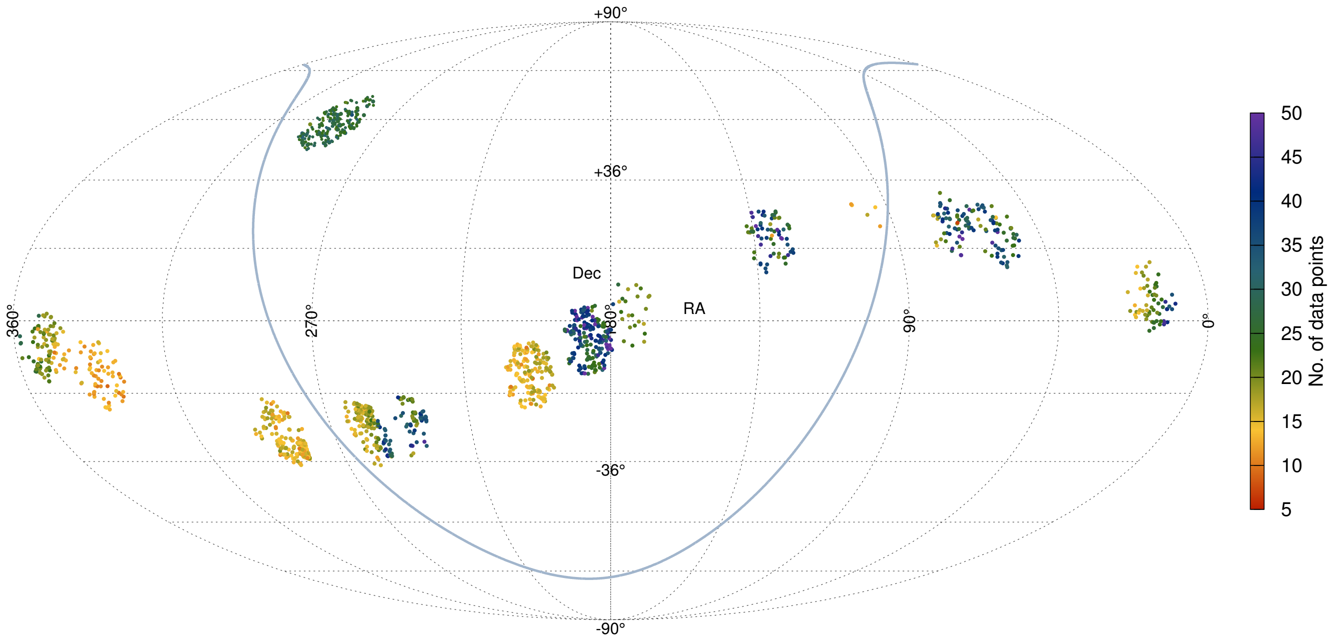

Fig. 2. Distribution of the *Gaia* DR2 RR Lyrae candidates confirmed by the *Kepler* (targeted and FFI) and K2 measurements in the sky, mapped in a Mollweide projection of equatorial coordinates. The thick line in light blue marks the Galactic equator, and the colour-coding shows the number of *Gaia* FoV transits in the G band per star. The original *Kepler* field is the northernmost group of stars at RA ≈ 300 deg.

solutions provided by MAST were useful, and we applied the PyKE software to generate customised extended aperture photometry to obtain better data (Still & Barclay 2012; Plachy et al. 2017; Vinícius et al. 2017). The distribution of the RR Lyrae stars from the *Kepler* and K2 missions cross-matched with the *Gaia* DR2 candidates are plotted in Fig. 2 in equatorial coordinates.

For the K2 observations, our validation procedure was based on visual inspection and quantitative properties of the light-curve shapes. Our criteria for the visual inspection were the following. Fundamental-mode RR Lyrae and classical Cepheid stars have very distinct, almost sawtooth-like light curves with short, sharp rising branches and long descending branches that repeat (almost) regularly, or vary smoothly if they are modulated. Characteristic bumps or humps, generated by shockwaves appearing at the surface of the star, also show up at distinct pulsation phases. Hump- or bump-like features can also appear in rotating variables, but in these cases, they drift in phase because of differential rotation.

Double-mode stars can be harder to identify as they have less asymmetric light curves. They usually still have steeper rising branches, often with prominent humps before maximum light. Scatter in the light curve extrema may indicate beating with additional modes. In contrast, regular or smoothly varying alternations in the minima or maxima can be attributed to differences in the components and/or appearance and evolution of spots in eclipsing binaries. Nevertheless, classification of nearly sinusoidal light curves can remain ambiguous.

Overtone stars show distinct beating patterns in the light curve, which unlike beating caused by spot patterns in rotating stars, repeat very regularly. Given the small number of double-mode stars, we examined all candidates in more detail before accepting them (see below).

3.1. Fourier parameters

For a quantitative analysis, we calculated the Fourier fits of the five strongest frequency components in the light curve using the LCFit code (Sódor 2012). Then we calculated the relative Fourier parameters, comparing the ith and the first harmonics with amplitude ratios $R_{i1} = A_i/A_1$ and phase differences $\Phi^c_{i1} = \Phi^c_i - i\,\Phi^c_1$ (where c refers to cosine-based Fourier fits), as defined by Simon & Lee (1981) and Simon & Teays (1982). These terms are frequently used to classify pulsating variables and to separate various subclasses. We plot them as a function of period for $i = 2, 3$ in Fig. 3. The stars clearly separate into the RRab and RRc loci, with very few outliers (for comparison, we refer to Soszyński et al. 2009). We discuss some of the outliers in Sect. 4.2.2. The low quality of light curves in the dense stellar field of the Sgr stream prevented us from calculating Fourier parameters in only six cases. Because hints of RRab variability can be visually confirmed in them, however, we included these stars as positive detections.

To validate double-mode candidates, we also computed the period ratios of the two main modes of the RRd stars: these give strong constraints if the star is a normal or anomalous RRd star (Soszyński et al. 2016). These values, along with the analysis of all RRd stars observed in the K2 mission, will be published in a separate paper (Nemec, priv. comm.).

3.2. Brightness comparison

We compared the *Gaia* DR2 median G brightnesses with the values obtained from the K2 data. The passbands of the two missions are similar, with *Kepler* spanning a 420–900 nm and *Gaia* a slightly wider 350–1000 nm wavelength range, both peaking around 600–700 nm (Van Cleve & Caldwell 2016; Evans et al. 2018).

We computed the flux-calibrated brightnesses ($Kp_{\rm fc}$) of the K2 stars from the PDCSAP light curves, using the zero-point of 25.3 mag, as determined by Lund et al. (2015). The two measurements generally agree well (see Fig. 4), with 60% of the stars below 0.1 mag difference, although it exceeds 1 mag for about 10% of the targets. The flux-calibrated magnitudes provide a better agreement with *Gaia* than the values found in EPIC ($Kp_{\rm EPIC}$)

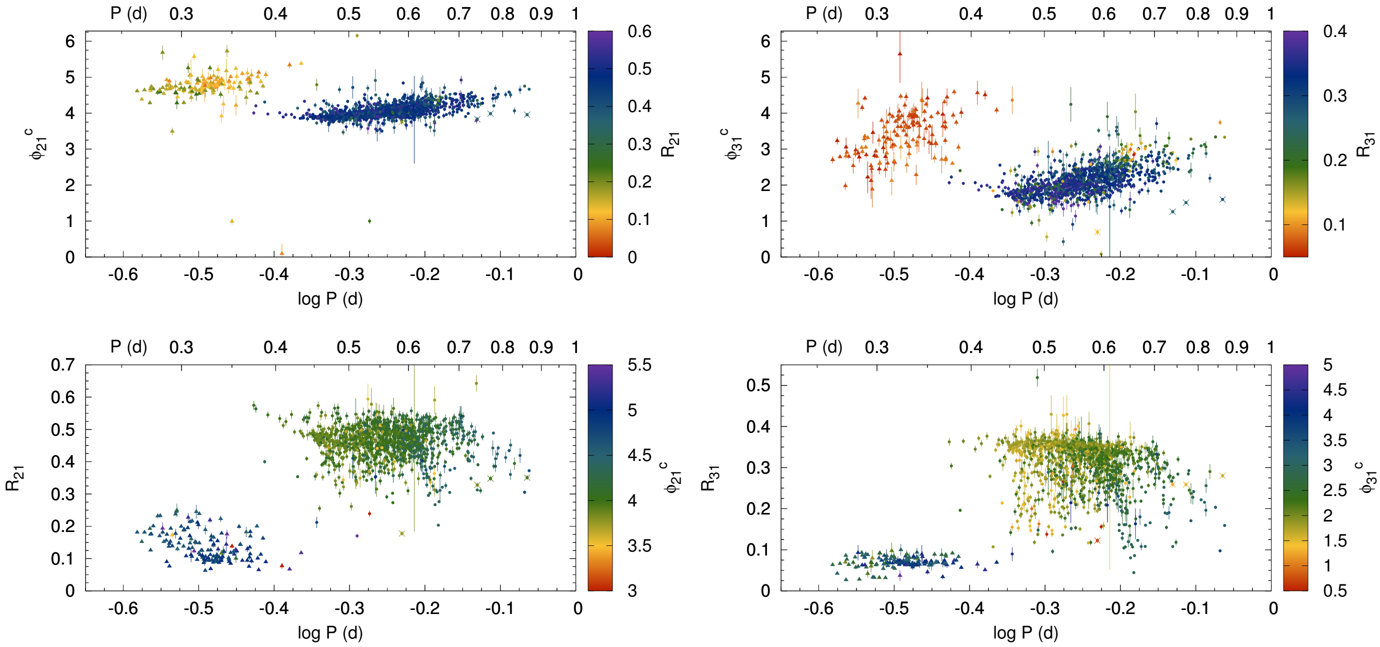

Fig. 3. Fourier parameters of the light curves of cross-matched K2–*Gaia* RR Lyrae-type sources. The *i*th and the first Fourier harmonics are compared by the Φ_{21}^c, Φ_{31}^c relative phase differences (in radians, using cosine-based Fourier fits) and the R_{21}, R_{31} amplitude ratios, as a function of the periods (P). The colour-coding shows the respective $R_{21}-\Phi_{21}^c$ and $R_{31}-\Phi_{31}^c$ pairs. Triangles show RRc stars, dots represent RRab stars, and crosses show potential anomalous Cepheid stars (see Sect. 4.2.2).

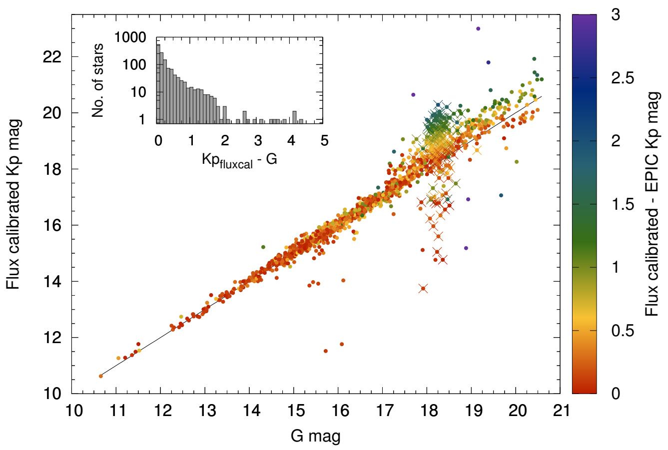

Fig. 4. Comparisons of the *Gaia* DR2 median G magnitudes and the flux-calibrated Kp magnitudes from the K2 mission, with the *inset* showing the distribution of the absolute differences. Colour-coding marks the absolute differences between the flux-calibrated Kp values and those found in EPIC: there is a clear systematic difference between the two Kp values towards, the faint end. Red outliers indicate agreement between the Kp values, but difference from the *Gaia* brightnesses. Crosses mark the Sgr stream stars in Field 7. The black line marks equality.

up to about $G \approx 18$ mag. For $G > 18$, as the colour-coding in Fig. 4 illustrates, the flux-calibrated values appear to be systematically fainter than the EPIC values. These issues probably come from the properties of the PDCSAP pipeline (poor background correction for dense fields and small pixel apertures for faint targets). This is especially pronounced for the population of the Sgr stream stars, marked with crosses in Fig. 4.

For about 1% of the stars, the *Gaia* brightness differs significantly from the two nearly identical Kp values ($|Kp_{\rm fc} - Kp_{\rm EPIC}| < 0.1$ mag, while $|G - Kp_{\rm fc}| > 1.0$ mag). The cause of discrepancy in most of these cases is that the cross-matched EPIC refers to a close-by (within 1–3 *Kepler* pixels) brighter star that is blended with the faint RR Lyrae variable. In these cases, the light curve also consists of the combined flux of the two stars. Many of these *G-Kp* discrepant stars are also members of the Sgr stream.

4. Results

The two quantities we are most interested in are the purity, defined as the fraction of bona fide variable stars of the appropriate class in the sample, and the completeness, defined as the fraction of the sources that are (properly) identified in *Gaia* DR2 compared to all (known) RR Lyrae stars within the FoVs. We investigated the distribution of stars according to the classification score, the number of *Gaia* FoV transits in the G band, and the median G brightness.

4.1. RR Lyrae stars in the original Kepler field

We identified 48 *Gaia* DR2 targets that were observed by *Kepler,* and we were able to confirm 44 of them as RR Lyrae variables, suggesting a purity of 92%. Beyond these 44 stars, 12 additional RR Lyrae stars have been identified by the KASC Working Group in the Lyra-Cygnus field, not all of which are published yet, indicating a completeness of 78% (44/56). One of the stars that is missing from the *Gaia* DR2 RR Lyrae sample is RR Lyr itself (as explained in Rimoldini et al. 2018), which was observed in the original FoV of *Kepler* (Kolenberg et al. 2011). (RR Lyr is present in DR2, but with an erroneous mean G brightness and parallax values (Gaia Collaboration 2018).)

After checking the data from the targeted observations, we then generated FFI-based light curves for a further 267 stars. We were able to classify 147 and 38 as (potential) RRab or RRc variables. The photometry of a further 10 targets was not successful as they were either near the edges of a CCD module or were blended with bright stars: we did not include them in our

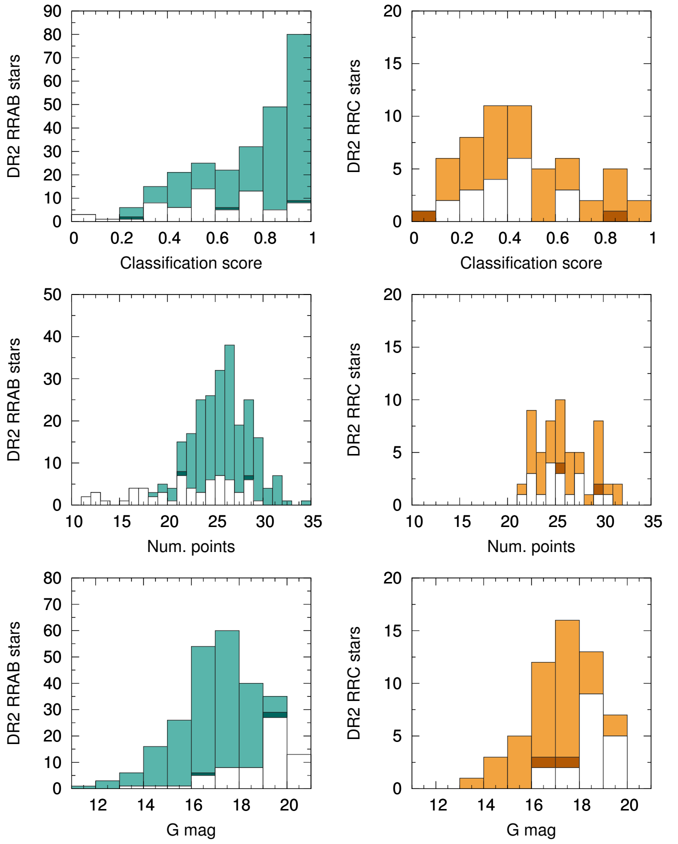

Fig. 5. Histograms for the *Gaia* DR2 RRAB- and RRC-classified objects in the original *Kepler* field. The light hues (light green and orange) mark stars that were confirmed as RRab or RRc variables based on the *Kepler* light curves. Dark hues are RR Lyrae stars that we reclassified into the other group (dark green: RRc variables found in the RRAB class; brown: RRab stars in the RRC class). White indicates stars that we could not confirm as RR Lyrae variables.

statistics. Combined with the LC targets discussed above, we can provide a lower-limit estimate of at least 75% for the purity of the *Gaia* DR2 RR Lyrae candidates within the original *Kepler* field, although the low value can partially be attributed to the sparse FFI data we used. The completeness of the sample rises to 96% when we include the stars confirmed by the FFI light curves. A more detailed analysis of the FFI sample will be published elsewhere (Molnár & Hanyecz, in prep.).

The distribution of the stars in the original *Kepler* field against the classification scores, the number of *Gaia* *G*-band FoV transits, and the median *G* brightnesses are shown in Fig. 5. There is very little cross-contamination between the RRAB and RRC classes (stars that we classified into a different type). Contamination from sources that we could not confirm as RR Lyrae variables is significant among stars that have low coverage and/or are faint ($G > 18$ mag). However, this can be partially attributed to the FFI photometry pipeline that was not developed to handle very faint *Kepler* targets. The purity of the FFI sample is 90% or 85% when we limit the targets to $G < 18$ or $G < 19$ mag, respectively.

4.2. RR Lyrae stars in the K2 fields

Overall, we were able to inspect the light curves of 1395 cross-matched K2 targets from Campaigns 0–8 and 10–13. The

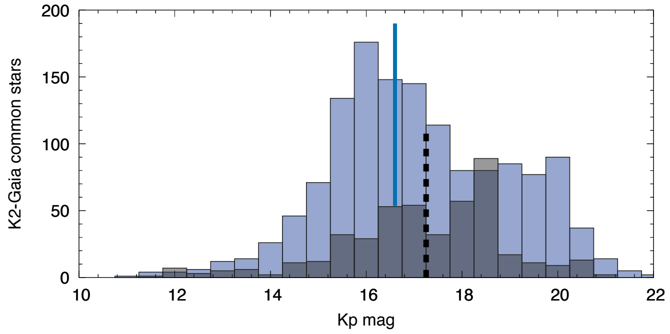

Fig. 6. Brightness distributions of the K2 targets stars that are classified as RR Lyrae variables both in *Gaia* DR2 and based on their K2 light curves (blue), vs. those that were missed as variable candidates in *Gaia* DR2, but their K2 light curves show RR Lyrae variation (grey). The blue solid and black dashed lines indicate the median values. These histograms are overlaid, not stacked.

distribution of the *Gaia* DR2 candidates observed in the K2 mission is not uniform: the 601 stars from Campaigns 2, 7, and 11 represent only 6% of the 9843 observable targets in those fields, whereas for the rest, the halo fields the ratio is 52% (787/1518).

Of these 1395 stars, we confirmed the RR Lyrae-type variability of 1371 stars (1243 RRAB, 141 RRC, and 10 RRD sources). No data were available for the five ARRD-type stars that were within the K2 FoVs. The remaining 24 observed stars turned out to be different types of variables. These numbers lead to a *Gaia* DR2 purity of 98%, in agreement with those of the targeted observations of the original Lyra-Cygnus field, and the findings of Holl et al. (2018). However, we emphasize again that the observed stellar samples of the K2 campaigns are not necessarily representative of the true stellar populations within those fields, therefore we treat the purity value as an upper limit. We discuss the range of possible purity values in Sect. 4.3 in more detail.

Although the K2 photometric data are superior to data from other surveys, the observed sample is not exhaustive, so we can only provide an estimate for the completeness of the *Gaia* DR2 classifications. Targets proposed for observation in the K2 campaigns were cross-matched from various surveys, and vetted based on the available photometry in the literature, and they are therefore more complete than any single catalogue (Plachy et al. 2016). Light curves gathered by *Kepler* based on these proposals were also checked visually, and they revealed very low level of contamination by other variables. Therefore, we considered the number of proposed and observed stars to be a good estimate for the number of true RR Lyrae stars in the FoVs. We then collected all stars that were proposed and observed during the mission, but had no counterparts in *Gaia* DR2 RR Lyrae classifications. We ended up with 445 targets, leading to a completeness estimate of 75% for the K2 fields. This is somewhat higher than the values computed for the OGLE fields (Holl et al. 2018), but agrees with our estimate for targeted observations in the original *Kepler* field. We plot the brightness distribution of the confirmed and missed RR Lyrae stars from the K2 mission in Fig. 6. The two distributions are fairly similar, but the maximum is shifted towards fainter magnitudes for the stars that are not classified as RR Lyrae in the *Gaia* DR2 classification table, by about 0.6 mag (the medians of the two groups are 16.6 and 17.2 mag).

We also compared the brightness distributions of all *Gaia* DR2 candidates falling into the K2 fields to all confirmed RR Lyrae stars therein (including the missed 445 stars) in Fig. 7

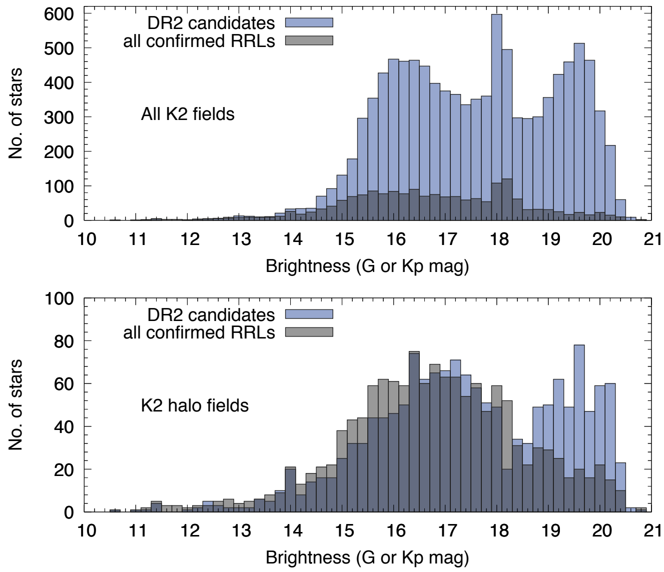

Fig. 7. Brightness distributions of the *Gaia* DR2 candidates (blue) by Holl et al. (2018) in the K2 fields vs. all known confirmed RR Lyrae stars within in the same fields, including stars not in the DR2 variability catalogue. The *upper panel* shows all fields; the spike at 18 mag is the Sgr stream. The *lower panel* shows the halo fields only.

to see how different the selection function of the two missions are. Based on their capabilities alone, *Kepler* is only limited by source confusion, but it is able to observe stars below the faint limit of *Gaia* (see e.g. the RR Lyrae stars in Leo IV; Molnár et al. 2015). However, the sample observed by *Kepler* was limited by the input catalogues used for target selection. The large difference in the upper panel of Fig. 7 comes from the large number of bulge stars that were not observed in the K2 mission. The comparison of the halo fields only (i.e. excluding Campaigns 2, 7, and 11) shows a much better agreement in the lower panel. For stars brighter than 16.5 mag (here we used either G or Kp magnitudes for stars, given the good agreement between the two), the K2 observations and the input catalogues we used provide a sample more complete than that of Holl et al. (2018). Interestingly, the *Gaia* DR2 sample shows another excess below 18.5 mag. Since most of these stars have no K2 light curves, we cannot decide if these stars are contaminants or bona fide RR Lyrae stars that were not detected by other surveys before. Unfortunately, this brightness range will not be accessible to the TESS space telescope either.

4.2.1. RR Lyrae subclass statistics

Based on their K2 light curves, we identified 1371 objects as RR Lyrae stars, 1211 (88%) RRab, 142 (10%) RRc, and 17+1 RRd and anomalous RRd stars (1%). Our classifications do not always agree with those of *Gaia* DR2. We found the *Gaia* DR2 RRAB class to be nearly pure, with only 1% (14/1243) of contamination from the other classes (stars that turned out to be RRc- or RRd-type pulsators instead). The contamination in the RRC and RRD classes was 8% (12/142) and 50% (5/10), respectively. While contamination rises significantly for these classes, they are much less numerous, therefore the overall rate is only about 2% for the RR Lyrae stars in general.

Moreover, of the 24 stars that we could not confirm as RR Lyrae variables, K2 light curves of 5 RRAB candidates revealed Cepheid variations (1 anomalous and 4 Type II Cepheids). This indicates a low level of cross-contamination between the

Table 1. Confusion matrix of the *Gaia* and K2 results.

	Gaia DR2 RRAB	*Gaia* DR2 RRC	*Gaia* DR2 RRD	Sum
RRab	99.6% (1206)	2.8% (4)	0.1% (1)	1211
RRc	6.3% (9)	90.8% (129)	2.8% (4)	142
RRd	27.8% (5)	44.4% (8)	27.8% (5)	10
Neither	95.8% (23)	4.2% (1)	0% (0)	24
Contam.	3.0% (37)	9.1% (13)	50% (5)	

Notes. The *Gaia* classifications are compared to our findings. Contamination of each *Gaia* class is presented in the bottom row.

RR Lyrae and Cepheid classes. We reclassify two of these, V1637 Oph (EPIC 234649037, $P_{\rm K2}$ = 1.327 d), previously classified as an RR Lyrae, and FZ Oph (EPIC 251248334, $P_{\rm K2}$ = 1.500 d), an under-observed variable, as short-period Type II Cepheids, also known as BL Her-type stars, based on their K2 light curves. The results are summarised in the confusion matrix in Table 1.

Figure 8 shows the distribution of the classification scores for the various RR Lyrae subclasses. The left panels group the stars according to their *Gaia* DR2 classification types. We indicate the portion of stars that we reclassified into different classes with darker hues. The right panels show the variability types based on the K2 light curves, with the same colours denoting the number of stars that we reassigned a different RR Lyrae subclass. The difference between the classification score distributions of RRab and RRc/RRd stars is striking: half of the RRab stars have scores above 0.8, while the distribution of the RRc and RRd stars is essentially flat, and for RRc stars the score falls off near 1.0. The flat distribution suggests that the RRC and RRD classes were harder to identify, likely because of the competition with the dominant RRAB class. Figure 8 also confirms that the RRab sample is nearly pure (almost all stars are from the RRAB class), while the RRd sample includes *Gaia* classifications of all three subclasses. For the RRc stars, the sample appears to be nearly pure above classification score 0.6, and about 15% of RRc stars with scores < 0.6 ended up in the RRAB or RRD classes of the *Gaia* DR2 classification.

The distributions of the median brightnesses in the G band of the three RR Lyrae subclasses are shown in Fig. 9. In contrast to the distribution from the original *Kepler* field, we do not see an increase of unconfirmed variables toward the faint end, as they appear rather evenly spread. Interestingly, RRc stars brighter than $G \sim 15$ mag seem to be missing from our particular selection of *Gaia* DR2 classifications (centre left panel), although we reclassified a few of them from the other subclasses (centre right panel). However, the number of RRab stars also decreases rapidly from 14–15 mag towards the bright end. Therefore, the apparent lack of a few RRc stars most certainly stems from small number statistics here.

Finally, Fig. 10 shows the distribution of stars according to the number of G-band observations per RR Lyrae subclass. Here, the unconfirmed variables are clearly grouped at the low end, especially for the RRAB class. For stars with 20 or fewer FoV transits, the contamination is about 3%, and for the few stars with fewer than 12 FoV transits, it rises to 8%. These numbers indicate that the *Gaia* DR2 sample of RR Lyrae classifications is relatively pure even for stars with only a few observations.

These estimates illustrate that the classifications of *Gaia* DR2 are already well suited to identifying single-mode pulsators,

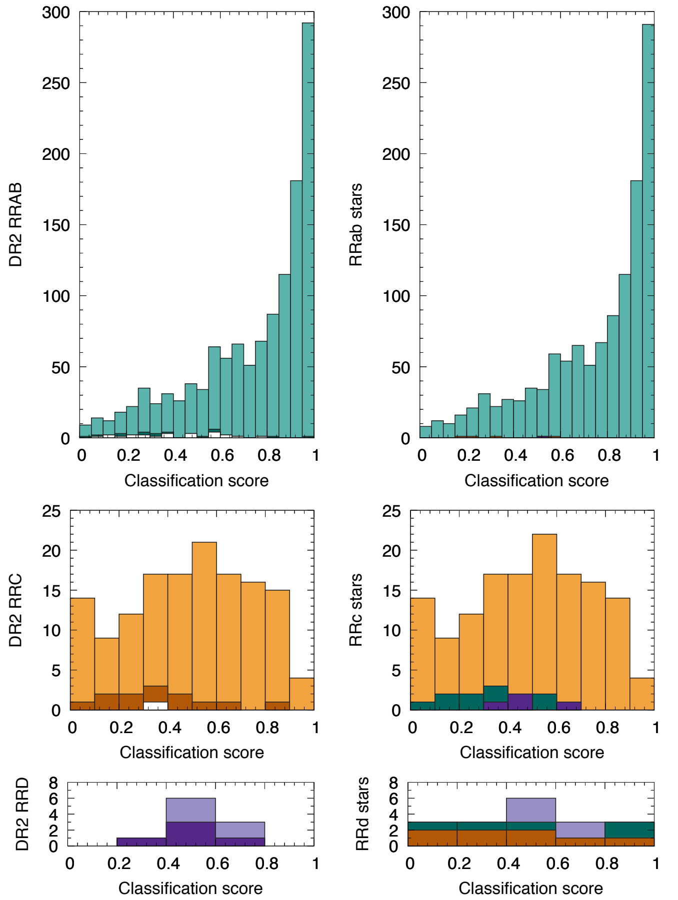

Fig. 8. Distribution of the various classification scores of the RR Lyrae subtypes. *Left panels*: *Gaia* DR2 classification type; *right panels*: as classified based on the K2 data. The distribution is skewed towards score values of 1.0 for RRab stars but is flat for RRc and RRd stars. Light green, orange, and pale violet mark bona fide RRab, RRc, RRd stars, respectively. Dark green, brown, and dark lilac denote stars that we classified into a different subclass (which are distributed in different panels on the right-hand side). White bars refer to the counts of stars for which the RR Lyrae variability was not confirmed or was rejected.

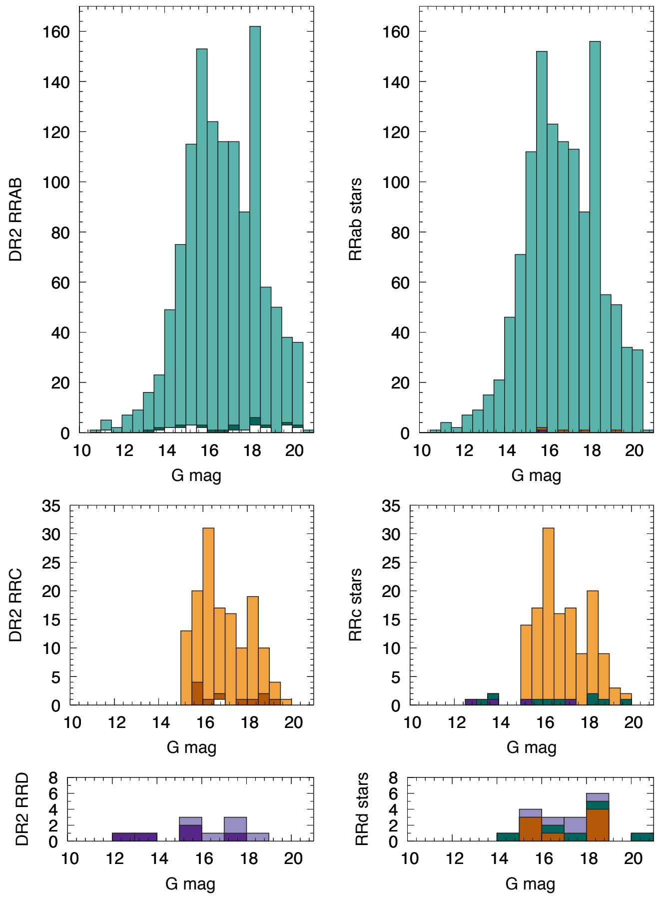

Fig. 9. Distribution of the median G magnitudes per RR Lyrae subtype. The colour-coding is the same as in Fig. 8.

even at low numbers of FoV transits, but multiperiodic objects like RRd stars will need more extended observations.

4.2.2. Ambiguous identifications

As mentioned in Sect. 3, the Fourier parameters in Fig. 3 mostly separate the stars into two groups corresponding to RRab- and RRc-type stars, but we found a few stars that we could not classify unambiguously. We flagged four stars as potentially anomalous Cepheids: three fundamental-mode candidates and one first-overtone candidate. We must emphasize, however, that the Fourier parameters of anomalous Cepheids overlap with those of RRab stars for certain period ranges, and therefore we could not rule out that these four objects are RRab stars based on their light-curve shapes alone. These four stars are identified in K2 as EPIC 206010651, 206175324, 212459957, and 234523936 (the last is the overtone candidate), corresponding to *Gaia* DR2 `source_id` 2600030303142307968, 2614960399737036672, 3606980678405498368, and 4134356134978875904, respectively.

Some outliers of the distribution of Φ^c_{21} in Fig. 3 can be attributed to a peculiar group of modulated RRab stars that exhibit a strong Blazhko effect, with a full rotation in the Φ^c_{i1} parameters during the minimum-amplitude cycle (Guggenberger et al. 2012; Bódi et al. 2018). This phenomenon could shift the average value of the parameters when the duration of the light curve is comparable to the modulation period. This is the case, for example, for EPIC 245954410, with $\log P/d = -0.29$ and $\Phi^c_{21} = 6.16$ rad, and for EPIC 212545143, an extremely modulated RRab star with $\log P/d = -0.27$ and $\Phi^c_{21} = 0.99$ rad. We note that two other stars at low Φ^c_{21} but with shorter periods are not flagged as outliers. Although most RRc stars appear in a tight group, others spread out from 0 to 2π, and these two stars represent that subgroup of RRc stars.

4.3. RR Lyrae completeness and purity ranges

The observations of the two missions of *Kepler* provide different samples that are not straightforward to combine. Here we discuss the various samples and statistics obtained. We can focus on the original *Kepler* field where the combination of the targeted, continuous light curves and the FFI photometry provides a sample that is not biased by prior selection of targets, but hindered by the sparse nature of the FFI light curves. Alternatively, we can gather all targeted observations from the two missions, resulting in a larger sample that is based on the same data acquisition method. Finally, we can calculate purity and completeness values for all three samples (*Kepler*, *Kepler* FFI, and K2) combined. Through these combinations we provide a range of estimates, which are summarised in Table 2.

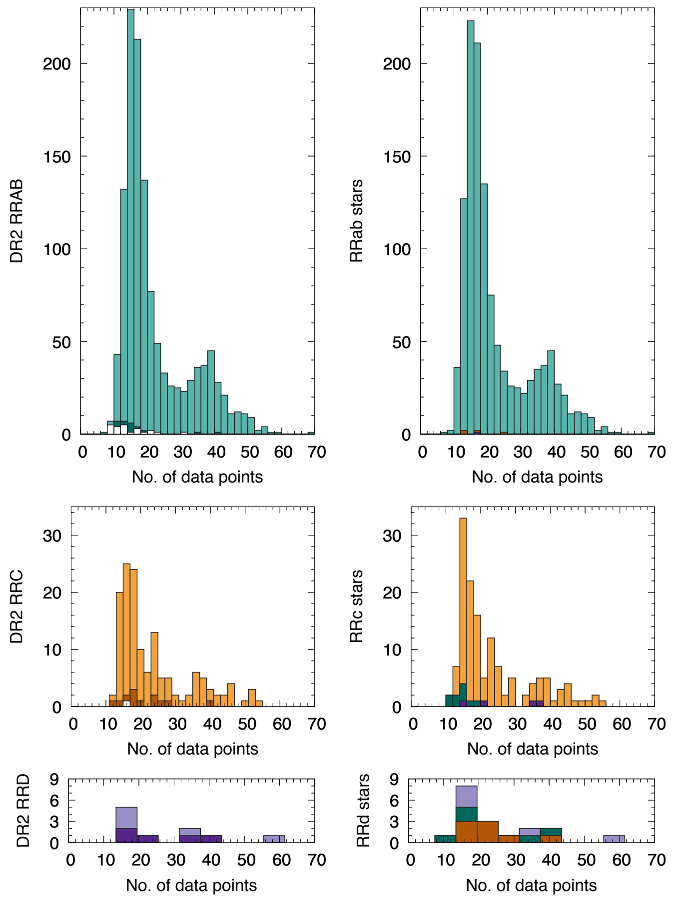

Fig. 10. Distribution of the number of FoV transits in the G band per RR Lyrae subtype. The colour-coding is the same as in Fig. 8.

Overall, we were able to identify 0.9% of the RR Lyrae variables classified by Rimoldini et al. (2018) over the whole sky in the observations of the two missions of *Kepler*, outside the bulge. Inside the *Kepler* and K2 FoVs, light curves were collected for 15% of the DR2 targets (1706 stars out of 11 702 potential, observable targets from the two missions). As stated above, however, the actual coverage is much higher for the original *Kepler* FoV (near-complete) and the halo fields in the K2 mission (52%).

For the *Kepler* and/or K2 targeted observations, we derived a purity rate between 92 and 98% (with lower limits at 51% and 75%), and a completeness in the range of 70–78%. The completeness is estimated by simply comparing the number of cross-matched and confirmed *Gaia* sources to those also in the fields but not classified as RR Lyrae variables in *Gaia* DR2. The lowest completeness value of 70% here is reached if we only use the targeted observations but count the FFI stars among the missed ones (even though they were not known before DR2). The DR2 classifications have a completeness of 75% within the K2 fields, while the combination of all data sets leads to an overall completeness rate of 78%.

We obtain a very high completeness ratio of 96% for the original *Kepler* field when the FFI stars are included. The lack of further sources here can be attributed to the fact that the field is at low Galactic latitudes and thus was largely avoided by deep surveys that could have identified faint RR Lyraes there. Therefore, the high completeness in this field reflects our limited knowledge that has now been expanded by the *Gaia* DR2 identifications, so we consider this value as an upper limit.

Table 2. Completeness and purity values for the RR Lyrae and Cepheid stars in *Gaia* DR2 based on the *Kepler* and K2 light curves.

Type	Data	*Gaia* matches	Compl.	Purity
RRL	*Kepler*	48	78%	92%
RRL	*Kepler* + FFI	311	<96%	>75%
RRL	K2	1395	75%	(51–)98%
RRL	*Kepler* + K2	1443	70–76%	(51–)98%
RRL	All	1706	78%	(55–)94%
CEP	All	41		~66%

Notes. Absolute lower limits of purity rates (from the worst-case scenario in the halo K2 fields) are indicated in brackets.

Even more diverse limits can be derived for the purity. The estimate of 92% originates from the targeted *Kepler* observations, but that sample alone is much smaller than the rest. Both the K2 and combined *Kepler*+K2 observations show 98% purity. This is in broad agreement with the findings of Holl et al. (2018). We conclude that the single-mode group RRAB is nearly pure, whereas the RRC group suffers from some contamination, but most of the confusing sources are RRd stars, and not eclipsing binaries. *Gaia* DR2 all-sky classifications are not well suited yet to identifying double-mode (RRd) stars, but as these are intrinsically rare, they have little effect on the overall population statistics.

The inclusion of the FFI stars into the statistics of the original field, however, warns us that the *Kepler* and K2 samples might include some bias against contaminating sources in *Gaia* DR2 that have not been proposed for observation in the *Kepler* and K2 missions. The combined statistics from the original field gives us a lower limit of 75% for purity, although this value is likely affected by faint RR Lyrae sources whose FFI photometry was of insufficient quality.

Of course, the *Kepler*+FFI sample is disjunct from the K2 FoVs, but it is the only one without prior selection bias. The excess of faint unobserved stars in the selection function (Fig. 7) suggests that the DR2 RR Lyrae sample might be less pure than our statistics suggests. In absence of these light curves, only a simple lower limit can be estimated for K2, considering that in halo fields, 52% of the *Gaia* sources were observed with 98% purity. Even if all the remaining 48% were false positives, this sets an absolute lower limit of 51% for the purity in the Galactic halo. The combination of halo and original-field stars raises that limit slightly to 55%.

4.4. Cepheids in the Kepler and K2 data

Cepheids are much less numerous than RR Lyrae stars, and many of them populate the Galactic bulge, the Magellanic Clouds, and the disc of the Milky Way. The *Kepler* and K2 missions largely avoided these areas to prevent source confusion and blending, therefore the overlap between these missions and the Cepheids in *Gaia* DR2 is much more limited than for the RR Lyrae stars. Because of the low number of targets, we decided to also include the Bulge stars in our statistics. Since Cepheids are intrinsically luminous, their variations can be recognised even when blended with other stars in the *Kepler* observations.

In the Lyra-Cygnus field of the original *Kepler* mission, *Gaia* DR2 classifications correctly include V1154 Cyg as CEP and HP Lyr as T2CEP (which is a potential RV Tau star; Graczyk et al. 2002), while it missed DF Cyg and misclassified V677 Lyr

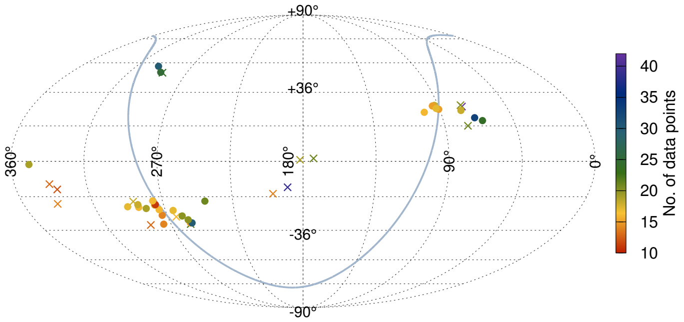

Fig. 11. Distribution in the sky of various Cepheid-type stars from the *Gaia* DR2 classifications cross-matched with *Kepler* and K2 measurements. Crosses mark stars that we rejected or were not able to confirm: half of them are at high Galactic latitudes. The notation is the same as in Fig. 2.

as T2CEP (known to be a longer-period semi-regular variable; Gorlova et al. 2015).

In the K2 fields, 73 stars classified as Cepheids in *Gaia* DR2 were observable (Fig. 11). Of these, we found data for 38 from the three subclasses (3 ACEP, 13 CEP, and 22 T2CEP stars), of which we were able to confirm 20 as a members of the Cepheid family, and 5 more as likely Cepheid stars. Given the small sample size, we did not separate the stars further into subclasses. With the inclusion of the stars from the RRAB class that we classified as Cepheids, the number of Cepheid variables common between *Kepler*/K2 and *Gaia* is 32 stars (2 from *Kepler* and 25+5 from K2).

The stars we were not able to confirm as Cepheids include four eclipsing binaries and/or rotational variables, five long-period variables, and four stars with unidentified variations. We classified one star as an RRab variable, further confirming a low level of cross-identification between Cepheids and RR Lyrae stars. This star was included in a K2 proposal as an RR Lyrae target, and thus it has been accounted for in our completeness estimate.

We were able to inspect the light curves of 54% (41/76) of the stars that fell into the FoVs of the two missions. The stars identified in the two *Kepler* missions account for only 0.9% (observable) and 0.5% (observed) of the 8550 Cepheid variables identified by Rimoldini et al. (2018). However, most of the DR2 detections are concentrated in the Magellanic Clouds, therefore our results are more relevant for the Cepheid population in the Milky Way. The findings indicate a purity on the order of 66% for this size-limited sample. Considering the very low number of sources available to us and the limited temporal coverage of the K2 data, this result is more or less in agreement with the 15% contamination in the sky-uniform test results of Holl et al. (2018).

5. Conclusions

RR Lyrae and Cepheid stars are used for many purposes, such as mapping the substructures and stellar populations of the Milky Way and the Magellanic Clouds. The *Gaia* Data Release 2 features a large collection of RR Lyrae and Cepheid candidates (among others) over the entire sky that can be exploited for such studies. However, many of the identifications are based on a low number of observations. The *Kepler* space telescope observed a selection of these targets in great detail during the *Kepler* and K2 missions. The fact that *Kepler* is able to provide continuous light curves for the entire brightness range of the *Gaia* DR2 targets provides a great opportunity to validate these classifications. We investigated the targets in common between the missions, consisting of 0.9% of both the RR Lyrae and Cepheid candidates from the all-sky classification of DR2 (Rimoldini et al. 2018). Within the areas of the *Kepler* and K2 FoVs (between Campaigns 0–8 and 11–13), 15% and 54% of the DR2 RR Lyrae and Cepheid candidates, respectively, had LC data recorded by *Kepler*. For the original FoV we also extracted sparse photometry from the full field images to obtain a sample more complete than the targeted observations alone.

We found that the photometry in *Gaia* DR2 is already suitable to properly identify single-mode pulsators. The RRAB class was found to be nearly pure, with very few contaminants. The RRC class and the various Cepheid classes are somewhat more contaminated. The few RRD (double-mode) stars we found included single-mode pulsators as well, and bona fide RRd stars have *Gaia* DR2 classifications from all RR Lyrae subclasses, indicating that reliable identification of multimode pulsators will require more observations. Overall, we found the purity to be in the 92–98% range, based on the targeted observations of the *Kepler* space telescope, with lower limits of 75% (FFI stars) and 51% (worst-case scenario for the halo K2 fields). For the classification of Cepheids in *Gaia* DR2, we provide a purity estimate in the order of 66%.

Based on the visual examination of the contaminating sources (stars that were not found to be RR Lyrae or Cepheid stars), we conclude that contaminants are more likely to be pulsators or rotational variables than eclipsing binaries.

We estimate the completeness of the *Gaia* DR2 RR Lyrae classifications in the *Kepler* and K2 fields to be around 70–78% for the targeted observations with an upper limit of 96% for the original field, if we included the FFI stars as well.

All of the estimates presented here are summarised in Table 2, and they are in agreement with the limited validation tests presented in Holl et al. (2018), indicating that the observations of the *Kepler* space telescope can indeed be used to validate surveys that collect sparse photometry.

Acknowledgements. This research received funding from the Hungarian Academy of Sciences through the Lendület grants LP2014-17 and LP2018-7, and the János Bolyai Research Scholarship (L.M. and E.P.); the Hungarian National Research, Development, and Innovation Office through the NKFIH grants K-115709, PD-116175, and PD-121203; and the ÚNKP-17-3 program of the Ministry of Human Capacities of Hungary (Á.L.J.). L.M. has been supported by the Premium Postdoctoral Research Program of the Hungarian Academy of Sciences. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, some of which participate in the *Gaia* Multilateral Agreement, which include, for Switzerland, the Swiss State Secretariat for Education, Research and Innovation through the ESA Prodex program, the "Mesures d'accompagnement", the "Activités Nationales Complémentaires", and the Swiss National Science Foundation. This paper includes data collected by the *Kepler* mission. Funding for the *Kepler* and K2 missions are provided by the NASA Science Mission Directorate. The *Kepler*/K2 data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts. We thank Carolyn Doherty for proofreading the manuscript, and Benjamin Montet for the fruitful discussions about the `f3` code.

References

Barros, S. C. C., Demangeon, O., & Deleuil, M. 2016, A&A, 594, A100
Beaton, R. L., Freedman, W. L., Madore, B. F., et al. 2016, ApJ, 832, 210
Benkő, J. M., Kolenberg, K., Szabó, R., et al. 2010, MNRAS, 409, 1585

Benkő, J. M., Plachy, E., Szabó, R., Molnár, L., & Kolláth, Z. 2014, ApJS, 213, 31
Bódi, A., Szatmáry, K., & Kiss, L. L. 2016, A&A, 596, A24
Bódi, A., Molnár, L., Plachy, E., & Szabó, R. 2018, in Proceedings of the Polish Academy of Sciences, The RR Lyrae 2017 Conference, 6, 282
Borucki, W. J. 2016, Rep. Prog. Phys., 79, 036901
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Clementini, G., Ripepi, V., Molinaro, R., et al. 2018, A&A, in press, DOI: 10.1051/0004-6361/201833374 (*Gaia* 2 SI)
Derekas, A., Plachy, E., Molnár, L., et al. 2017, MNRAS, 464, 1553
Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, ApJ, 696, 870
Drake, A. J., Catelan, M., Djorgovski, S. G., et al. 2013, ApJ, 763, 32
Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Gorlova, N., Van Winckel, H., Ikonnikova, N. P., et al. 2015, MNRAS, 451, 2462
Graczyk, D., Mikolajewski, M., Leedjarv, L., et al. 2002, Acta Astron., 52, 293
Guggenberger, E., Kolenberg, K., Nemec, J. M., et al. 2012, MNRAS, 424, 649
Hernitschek, N., Schlafly, E. F., Sesar, B., et al. 2016, ApJ, 817, 73
Holl, B., Audard, M., Nienartowicz, K., et al. 2018, A&A, 618, A30
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Huber, D., Bryson, S. T., Haas, M. R., et al. 2016, ApJS, 224, 2
Juhász, A. L., & Molnár, L. 2018, in Proceedings of the Polish Academy of Sciences, The RR Lyrae 2017 Conference, 6, 298
Kolenberg, K., Bryson, S., Szabó, R., et al. 2011, MNRAS, 411, 878
Leavitt, H. S. 1908, Ann. Harvard Coll. Obs., 60, 87
Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100
Lund, M. N., Handberg, R., Davies, G. R., Chaplin, W. J., & Jones, C. D. 2015, ApJ, 806, 30
Minniti, D., Dékány, I., Majaess, D., et al. 2017, AJ, 153, 179
Molnár, L. 2018, Proc. Polish Acad. Sci., 6, 106
Molnár, L., Pál, A., Plachy, E., et al. 2015, ApJ, 812, 2
Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017, ApJ, 851, 116
Moskalik, P., Smolec, R., Kolenberg, K., et al. 2015, MNRAS, 447, 2348
Mullally, F., Barclay, T., & Barentsen, G. 2016, Astrophysics Source Code Library, [record ascl:1601.009]
Nemec, J. M., Cohen, J. G., Ripepi, V., et al. 2013, ApJ, 773, 181
Plachy, E., Molnar, L., Szabo, R., Kolenberg, K., & Banyai, E. 2016, Comm. Konkoly Obs., 105, 19
Plachy, E., Klagyivik, P., Molnár, L., Sódor, Á., & Szabó, R. 2017, Eur. Phys. J. Web Conf., 160, 04009
Ricker, G. R., Vanderspek, R., Winn, J., et al. 2016, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, Proc. SPIE, 9904, 99042B
Rimoldini, L., Holl, B., Audard, M., et al. 2018, A&A, submitted [arXiv:1811.03919]
Sesar, B., Ivezić, Ž., Stuart, J. S., et al. 2013, AJ, 146, 21
Simon, N. R., & Lee, A. S. 1981, ApJ, 248, 291
Simon, N. R., & Teays, T. J. 1982, ApJ, 261, 586
Sódor, A. 2012, Konkoly Obs. Occ. Techn. Notes, 15, 1
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2009, Acta Astron., 59, 1
Soszyński, I., Smolec, R., Dziembowski, W. A., et al. 2016, MNRAS, 463, 1332
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2017, Acta Astron., 67, 103
Still, M., & Barclay, T. 2012, Astrophysics Source Code Library, [record ascl:1208.004]
Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, PASP, 124, 985
Szabó, R., Kolenberg, K., Molnár, L., et al. 2017, Eur. Phys. J. Web Conf., 160, 04004
Udalski, A., Szymański, M. K., & Szymański, G. 2015, Acta Astron., 65, 1
Van Cleve, J. E., & Caldwell, D. A. 2016, Kepler Instrument Handbook, Technical Report., NASA Ames Research Center
Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948
Vega, L. D., Stassun, K. G., Montez, R., Jr., Boyd, P. T., & Somers, G. 2017, ApJ, 839, 48
Vinícius, Z., Barentsen, G., Gully-Santiago, M., et al. 2017, KeplerGO/PyKE, DOI: 10.5281/zenodo.835583

Appendix A: *Gaia* DR2 and KIC/EPIC cross-match

In the following tables, we provide the cross-match of the sources in common between *Gaia* DR2 classifications and the *Kepler* and K2 observations. Tables A.1–A.3 list stars from the original *Kepler* FoV that we classified as RRab, RRc, or likely RRab/RRc variables. Tables A.4–A.6 include the RRab, RRc, and RRd stars identified in the K2 observations, respectively. Longer tables are available at the CDS in full length. Table A.7 lists the cross-matched Cepheid-type stars. Finally, stars that we rejected as RR Lyrae or Cepheid stars based on their K2 data are presented in Tables A.8 and A.9.

Table A.1. Cross-match of RRab-type stars in the original *Kepler* observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 class	KIC	*Kp* (mag)	Data type
2051319231466329600	290.0443778	38.2875548	15.852	RRAB	3111002	15.989	FFI
2051756764073824512	291.5422064	37.2395927	18.121	RRAB	1721534	16.913	FFI
2051849058633285632	292.8211913	37.6989397	16.471	RRAB	2309247	16.387	FFI
2051927669415635072	292.6072021	38.2211264	16.179	RRAB	3121676	15.959	FFI
2051930903531763072	292.5824442	38.2520216	16.165	RRAB	3121566	16.400	FFI
2052112730962078080	294.4600661	38.2577522	17.047	RRAB	3129996	17.002	FFI
2052124653792193280	294.5614553	38.4515596	17.952	RRAB	3354775	17.637	FFI
2052162282013412992	295.0290071	38.9723355	15.410	RRAB	3864443	15.593	target
2052226633499326336	294.5504877	39.1471277	14.583	RRAB	4069023	14.650	FFI
2052259898026804480	293.0292280	38.4096197	17.767	RRAB	3348493	17.827	FFI
...							

Notes. The *Kp* brightness here refers to the values found in KIC, data type is either target (LC data available) or FFI (no LC data, just FFI photometry). The following information was extracted from the *Gaia* DR2 archive: RA and Dec coordinates (`ra` and `dec` fields of the `gaiadr2.gaia_source` table), the median *G*-band magnitudes (rounded values from those in the `median_mag_g_fov` field of the `gaiadr2.vari_time_series_statistics` table), and the DR2 class (content of the `best_class_name` field of the `gaiadr2.vari_classifier_result` table). The full table is available at the CDS.

Table A.2. Cross-match of RRc-type stars in the original *Kepler* observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 class	KIC	*Kp* (mag)	Data type
2053453246098423424	291.2093828	40.2449799	14.652	RRC	5097329	14.896	FFI
2073611761018174208	298.7963502	40.6390445	17.183	RRC	5476906	17.125	FFI
2076328310640589312	296.0172668	39.9738297	17.391	RRC	4852066	17.393	FFI
2076789487048365952	296.9555846	41.1889926	17.642	RRC	5895400	17.796	FFI
2079302622726841728	298.4694906	45.0157871	16.093	RRC	8839123	16.229	FFI
2080332143568818304	296.6246981	46.5426500	17.569	RRC	9783052	17.730	FFI
2080566717498680448	296.7116973	47.2295668	16.997	RRC	10221234	17.050	FFI
2086448009499589376	298.3179484	47.7771822	15.909	RRC	10553801	15.700	FFI
2086583008909904896	296.9578235	47.5964425	17.165	RRC	10418799	17.125	FFI
2100304359969626624	285.1271038	39.4015144	17.685	RRC	4345865	17.716	FFI
2100508422455676288	287.5221537	39.5553736	18.080	RRC	4451334	18.263	FFI
2102965800881905792	288.3245722	43.5415773	15.384	RRC	7812805	15.546	FFI
2102980060173158144	287.9872854	43.3258509	15.749	RRC	7672313	15.864	FFI
2103421170493346816	284.5119233	40.3957360	16.541	RRC	5166889	16.784	FFI
2104875824376341888	283.7965235	42.0339167	16.012	RRC	6584320	16.053	FFI
2105082292044513536	283.5884772	43.0995012	15.724	RRC	7422845	15.867	FFI
2105292058247822976	281.8111161	44.1887140	17.201	RRC	8211945	17.258	FFI
2105292195685604608	283.0499329	43.4426085	17.006	RRC	7733600	17.038	FFI
2105850953752583680	284.8012931	43.9498940	16.619	RRC	8081725	16.927	FFI
2106317524639255680	286.2291335	44.7836872	16.065	RRC	8612183	16.198	FFI
2106527496999999488	285.9604747	46.0288732	13.296	RRC	9453114	13.419	target
2106890954312188288	283.3226482	45.0401760	15.321	RRC	8801073	15.56	FFI
2106998156695528448	283.3651904	45.5335706	14.838	RRC	9137819	14.991	target
2116749617944811264	280.8543785	42.5792557	18.126	RRC	7006857	17.999	FFI
2126843276427186432	290.0433736	43.7155517	14.864	RRC	7954849	14.862	FFI
2132728824731016192	288.8330915	50.2067867	16.846	RRC	11909124	16.706	FFI
2133142859577849856	289.2404017	50.8121588	16.904	RRC	12204812	17.119	FFI

Table A.3. Stars with uncertain but likely classifications from the original *Kepler* observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 class	KIC	*Kp* (mag)	Data type	*Kepler* class
2052863426827875456	290.6708966	38.8452285	17.444	RRAB	3744225	17.710	FFI	RRab?
2073860422436544000	297.4519701	40.9893283	14.835	RRAB	5727406	14.972	FFI	RRab?
2099377506028502400	288.2585772	38.4944088	19.527	RRAB	3331090	19.427	FFI	RRab?
2128491611857019776	292.9388450	47.8195878	18.659	RRAB	10602685	18.171	FFI	RRab?
2126654744543334784	291.5287676	45.1965107	16.213	RRAB	8884795	16.261	FFI	RRc?
2077964177781795200	295.0209805	42.9274675	17.784	RRC	7373841	17.267	FFI	RRc?
2078178681323184768	294.4612324	43.3732497	19.252	RRC	7690615	19.224	FFI	RRc?
2080190551386659456	296.2246715	45.7374787	17.731	RRC	9292598	17.933	FFI	RRc?
2080284112954652544	297.5274861	46.4933085	17.164	RRC	9725750	17.141	FFI	RRc?
2086553047221169408	298.1033115	48.5651069	16.978	RRC	11045097	16.933	FFI	RRc?
2104482469792786688	283.3977429	41.7620072	18.390	RRC	6343434	18.275	FFI	RRc?
2105141974910639232	283.2231541	43.9378973	14.954	RRC	8078670	15.143	FFI	RRc?
2130857971336396032	287.5319452	47.2799976	17.003	RRC	10198186	17.157	FFI	RRc?
2132959408636529792	287.6619814	50.4778068	19.039	RRC	12006372	18.943	FFI	RRc?
2052170249169053568	294.1764164	38.5791661	19.825	RRAB	3454674	18.810	FFI	RRc/EB?
2085218132725526528	300.3009309	45.8995148	19.121	RRAB	9368129	18.842	FFI	RRc/EB?
2100223412723866880	284.7592318	39.2885991	18.378	RRC	4136159	18.187	FFI	RRc/EB?
2133079873879600384	289.9087494	50.5891289	17.925	RRC	12059064	17.620	FFI	RRc/EB?

Table A.4. Cross-match of RRab-type stars in the K2 observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 Class	EPIC	K2 C	*Kp* (mag)	P (d)
3374190873985045888	94.4518789	19.7203845	15.109	RRAB	202064530	0	15.159	0.49697
3376253420360404992	97.3380133	22.1837461	16.122	RRAB	202064526	0	16.044	0.56076
3379937093550342528	101.2078923	24.3026555	14.439	RRAB	202064516	0	14.531	0.52379
3381448299265993728	101.3520857	24.5421433	14.789	RRAB	202064491	0	14.565	0.57253
3425627711558357248	93.8482864	23.8313169	15.895	RRAB	202064531	0	15.409	0.50795
3602396707054332160	179.8946495	−1.8598515	19.006	RRAB	201339783	1	19.573	0.53086
3794039453472320256	174.1404128	−1.3380417	18.571	RRAB	201375063	1	18.591	0.62803
3795681505368933760	176.5094136	0.3493481	19.828	RRAB	201488452	1	19.561	0.59032
3796741163995579136	172.7052697	−0.9883901	19.391	RRAB	201398311	1	19.624	0.46457
3797093488752748032	171.4873933	−0.1616250	14.282	RRAB	201454019	1	14.326	0.70842
...								

Notes. K2 C refers to the observing campaign. The *Kp* brightness refers to the flux-calibrated values. For the K2 observations we also list the pulsation periods. The full table is available at the CDS.

Table A.5. Cross-match of RRc-type stars in the K2 observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 Class	EPIC	K2 C	*Kp* (mag)	P (d)
3786001409293244800	173.4343361	−5.1026035	15.862	RRC	201158092	1	16.061	0.31014
3896812458883270400	175.4991113	3.9916983	15.412	RRC	201720727	1	15.477	0.30975
66183697982876416	60.1350351	24.4304789	16.015	RRC	211091644	4	16.041	0.29042
607162729019029888	135.5588208	14.1677961	16.358	RRC	211573254	5	16.495	0.31959
609116694325015680	130.7046858	13.3769410	16.293	RRC	211516905	5	16.275	0.34025
612376063402765312	135.0281818	18.6079110	15.357	RRC	211891936	5	15.386	0.34697
657965796923993216	129.2065123	15.9333259	16.456	RRC	211701322	5	16.311	0.33911
658334137615243648	131.1006087	17.2093019	19.544	RRC	211792469	5	19.968	0.30588
659746632100473984	129.2783707	18.9322997	15.905	RRC	211913888	5	15.942	0.2977
665174886647140736	130.4816877	22.0112793	15.745	RRC	212099502	5	15.767	0.32141
...								

Notes. The full table is available at the CDS.

Table A.6. Cross-match of RRd-type stars in the K2 observations.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 Class	EPIC	K2 C	*Kp* (mag)	P_0 (d)	P_1 (d)
3796490612783265152	176.8340702	1.8239436	15.839	RRC	201585823	1	15.774	0.48260	0.35942
6248239227324924416	239.1581229	−18.8471338	14.745	RRAB	205209951	2	14.757	0.47077	0.34877
51156844364167552	58.0512607	20.3530932	15.504	RRC	210831816	4	15.541	0.48878	0.36380
612194609624700928	135.5928672	18.5610970	18.786	RRC	211888680	5	19.185	0.48300	0.35938
3610631916003219328	204.2092377	−11.7283006	15.674	RRD	212547473	6	15.619	0.54507	0.40643
3620942277055055488	197.4515385	−13.8228783	16.517	RRC	212449019	6	16.617	0.48777	0.36339
4071397068375975296	282.0255672	−29.2124249	18.162	RRC	229228184	7	18.793	0.45927	0.34111
4071405658308934144	282.1318806	−29.0042516	18.076	RRAB	229228194	7	19.183	0.52268	0.38971
4071509081124919040	281.9373883	−28.3133952	18.531	RRC	229228175	7	18.858	0.47000	0.34945
4072051140361909632	282.7750015	−27.2508593	15.832	RRC	214147122	7	15.902	0.54103	0.40366
6761560112114793216	282.6946631	−29.1564157	16.901	RRAB	213514736	7	17.208	0.50359	0.37522
2576293393286532224	16.9326178	5.9752228	18.050	RRC	229228811	8	18.651	0.50021	0.37291
2580012972403894528	19.7918896	9.8452558	17.320	RRAB	220636134	8	17.390	0.50138	0.37374
3682596906250805632	192.1245496	−2.3467725	18.053	RRD	228952519	10	18.500	0.55142	0.40453
3698207256945765760	185.5512423	0.0526262	20.317	RRAB	248369176	10	20.952	0.56839	0.42407
3699831549153899648	185.4119454	0.8164297	17.024	RRD	201519136	10	17.019	0.46348	0.34432
2413839863087928064	349.3069278	−10.1488848	16.844	RRD	245974758	12	17.281	0.47527	0.35331
2438582821787698176	351.4208931	-8.0499976	17.044	RRD	246058914	12	17.105	0.45295	0.33611

Table A.7. Cepheid cross-match.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 class	KIC/EPIC	Campaign	ID
2078709577944648192	297.0644335	43.1269176	8.916	CEP	7548061	*Kepler*	V1154 Cyg
2101097215232231808	290.4127674	39.9355875	10.396	T2CEP	4831185	*Kepler*	HP Lyr
3378049163365268608	100.7812965	20.9391061	9.676	CEP	202064438	0	AD Gem
3423693395726613504	90.6524429	22.2341468	9.549	CEP	202064436	0	RZ Gem
3425495186047291136	92.5806547	24.0208511	10.485	CEP	202064435	0	V371 Gem
3425576270732293632	93.9995283	23.7474871	11.668	T2CEP	202064439	0	BW Gem
6249649277967449216	242.8120587	–16.8611619	12.444	T2CEP	205546706	2	KT Sco
47299585774090112	65.0074841	17.2793921	17.040	T2CEP	210622262	4	–
4085507620804499328	286.6122728	–19.6098184	12.820	T2CEP	217987553	7	V1077 Sgr
4085983537561699584	282.0408160	–20.1265928	13.173	T2CEP	217693968	7	V377 Sgr
4087335043492541696	286.5130651	–18.4282476	12.323	T2CEP	218642654	7	V410 Sgr
6869460685678439040	293.6444501	–19.3611183	12.956	ACEP	218128117	7	ASAS J193435-1921.7
4052361842043219328	274.9408203	–27.1592265	12.806	T2CEP	222668291	9	V1185 Sgr
4066429066901946368	273.2604032	–23.1172942	6.835	CEP	225102663	9	12 Sgr
4093976334264606976	273.8594387	–20.6295526	10.131	CEP	226412831	9	V1954 Sgr
4096140001386430080	275.8297937	–18.5747736	9.909	CEP	227267697	9	AY Sgr
4096341040228858240	275.2730788	–18.4554741	9.832	CEP	227315843	9	V5567 Sgr
4096979650282842112	276.1854194	–16.7971738	8.315	CEP	227916945	9	XX Sgr
4118144527610250880	264.9723348	–20.9930069	14.470	T2CEP	226238697	9	BLG-T2CEP-27
4111834567779557376	256.5229101	–26.5805651	6.835	CEP	232257232	11	BF Oph
3409635486731094400	69.3115536	18.5430127	6.267	CEP	247086981	13	SZ Tau
3415206707852656384	76.3094130	21.7635904	12.300	T2CEP	247445057	13	VZ Tau
3423579012158717184	92.1462986	22.6172200	8.377	CEP	202062191	0	SS Gem (T2CEP?)
4111218875639075712	260.2294555	-23.4355371	13.657	T2CEP	235265305	11	–
4111880369315900032	255.2860252	-26.5951004	11.628	T2CEP	232254012	11	ET Oph
4112437031430794880	257.1786569	-25.1637301	12.962	T2CEP	231047453	11	IO Oph
2638680812622984960	348.8605810	-1.3746326	17.749	ACEP	246385425	12	–

Notes. The three sections represent the *Kepler* field (*top panel*), clear identifications in the K2 fields (*middle panel*), and uncertain identifications (*bottom panel*).

Table A.8. Cross-match of stars that turned out not to be RR Lyrae stars.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 Class	EPIC	K2 C	Comment
3804350643452878848	166.9247344	0.4356445	18.594	RRAB	201494217	1	–
2601409430025491200	338.7320094	–13.0709211	17.999	RRAB	212235345	3	ROT
66847287606832768	56.7996921	25.1162548	11.421	RRAB	211132787	4	EC
4071401500782333696	281.9615772	–29.1119425	18.155	RRAB	213528187	7	ROT/DSCT
4072113675078719360	282.0081017	–27.5777297	15.741	RRAB	214027794	7	Low amplitude
4072275715608921344	282.9797618	–26.8576286	18.365	RRAB	229228258	7	ROT
4073145880282355200	282.3548453	–25.8753240	20.193	RRAB	214689700	7	Low amplitude
4074665130492370560	282.6322558	–25.3923767	16.858	RRC	214895832	7	HADS
4075390224042082688	284.9051268	–23.3645490	14.838	RRAB	215881928	7	BL Her (V839 Sgr)
4078049770897097088	281.5874740	–24.1344566	18.138	RRAB	215479005	7	ROT
4078775551632704896	283.7273346	–21.9521609	15.731	RRAB	216660299	7	–
4082823506754098432	289.0457792	–20.9322003	14.921	RRAB	217235287	7	BL Her (V527 Sgr)
3698191112165374976	184.8087296	–0.1380778	15.335	RRAB	201455676	10	Instrumental
4058445066313103744	262.9527077	–30.1317612	18.940	RRAB	242184466	11	–
4059999702649953152	264.3055426	–29.8639067	19.727	RRAB	240291558	11	–
4060392954243527296	263.8477795	–28.3171369	17.458	RRAB	240709679	11	ROT
4117562851501910912	263.5796074	–22.4795227	20.256	RRAB	225461305	11	Long period
4127629876912023168	255.7664846	–21.4563876	15.481	RRAB	230545230	11	HADS
4134649262221615104	258.3963306	–18.0976697	14.366	RRAB	234649037	11	BL Her (V1637 Oph)
4134727499321491072	258.7677959	–18.1450438	19.791	RRAB	234640705	11	–
6030292619420564352	255.8996713	–27.6952328	13.949	RRAB	251248334	11	BL Her (FZ Oph)
2435741447518507392	354.9756087	–9.0838411	15.281	RRAB	246015642	12	ACEP
2636644688886741248	345.5868479	–3.3788560	19.691	RRAB	246284344	12	Flare
3411660546628972928	74.1444421	20.6563863	14.468	RRAB	247311936	13	Long period

Notes. The last column provides alternative identifications and/or likely variability class, if possible.

Table A.9. Cross-match of stars that turned out not to be Cepheid variables.

Gaia DR2 source_id	DR2 RA (deg)	DR2 Dec (deg)	*G* (mag)	DR2 Class	EPIC	K2 C	Comment
2101228774371560832	288.8005648	39.7139988	11.473	T2CEP	4644922	*Kepler*	SRV (V677 Lyr)
3799150778087557888	173.5765613	1.2033677	12.936	T2CEP	201544345	1	LPV?
2594314457585274240	337.8754104	–18.0702930	13.192	T2CEP	205903217	3	LPV (BC Aqr)
2609074232957002880	338.0786514	–9.5948047	16.970	ACEP	206210264	3	RRab
2612858782044502272	334.2271302	–11.8676581	13.553	T2CEP	206109248	3	–
3609358780321640576	198.9566943	–13.7125905	15.246	T2CEP	212454161	6	EB
4087799999464896512	288.8008055	–17.0581624	13.837	T2CEP	219308521	7	–
6762146937758096640	284.0731029	–27.5238093	10.667	CEP	214047277	7	SRV? (V4061 Sgr)
3578288819399570816	189.6473674	–10.9422293	11.478	T2CEP	228708336	10	LPV?
3699052820043280640	181.8713363	0.6166174	13.984	T2CEP	201506181	10	–
4108800671623776128	256.5415354	–27.1369675	10.173	CEP	232135078	11	LPV?
4116400736547210496	264.3714041	–23.8031748	14.632	T2CEP	224691021	11	–
3392988846324877952	75.3718178	15.0239664	13.588	T2CEP	246736776	13	EB
3419364167475320448	74.4462919	23.5072079	16.317	T2CEP	247671949	13	ROT?
3419422583326134144	75.0288648	24.1428188	14.012	T2CEP	247761523	13	EB

Notes. The last column provides alternative identifications and/or likely variability class, if possible.