Distributed YSOs in the Perseus Molecular Cloud from the Gaia and LAMOST Surveys

Xiao-Long Wang1,2, Min Fang1,3, Yu Gan1,4,8, Hong-Xin Zhang3,5, Gregory J. Herczeg6,7, Hong-Jun Ma1, En Chen1,3, and Xing-Yu Zhou6,7

1 Purple Mountain Observatory, Chinese Academy of Sciences, No.10 Yuanhua Road, Qixia District, Nanjing 210023, People’s Republic of China
2 University of Chinese Academy of Sciences, No.19A Yuquan Road, Shijingshan District, Beijing 100049, People’s Republic of China
3 School of Astronomy and Space Science, University of Science and Technology of China, Hefei, Anhui 230026, People’s Republic of China
4 Department of Astronomy, Xiamen University, Xiamen, Fujian 361005, People’s Republic of China
5 Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei, Anhui 230026, People’s Republic of China
6 Department of Astronomy, Peking University, Yiheyuan Road 5, Haidian District, Beijing 100871, People’s Republic of China
7 Kavli Institute for Astronomy and Astrophysics, Peking University, Yiheyuan Road 5, Haidian District, Beijing 100871, People’s Republic of China

Identifying the young optically visible population in a star-forming region is essential for fully understanding the star formation event. In this paper, we identify 211 candidate members of the Perseus molecular cloud based on Gaia astrometry. We use LAMOST spectra to confirm that 51 of these candidates are new members, bringing the total census of known members to 856. The newly confirmed members are less extincted than previously known members. Two new stellar aggregates are identified in our updated census. With the updated member list, we obtain a statistically significant distance gradient of 4.84 pc deg−1 from west to east. Distances and extinction corrected color–magnitude diagrams indicate that NGC 1333 is significantly younger than IC 348 and the remaining cloud regions. The disk fraction in NGC 1333 is higher than elsewhere, consistent with its youngest age. The star formation scenario in the Perseus molecular cloud is investigated and the bulk motion of the distributed population is consistent with the cloud being swept away by the Per-Tau Shell.

Unified Astronomy Thesaurus concepts: Pre-main sequence stars (1290); Spectroscopy (1558); Statistical parallax (1573); Hertzsprung Russell diagram (725); Two-color diagrams (1724); Star forming regions (1565); Proper motions (1295)

Supporting material: machine-readable table

1. Introduction

A complete census of young stellar objects (YSOs) in star-forming regions is important for measuring the statistical properties of young stellar populations and for estimating the star formation rates (e.g., Hsieh & Lai 2013; Young et al. 2015; Mercík et al. 2017). Previous studies identified YSOs mainly based on their infrared (IR) excess emission (e.g., Harvey et al. 2007; Evans et al. 2007) or elevated X-ray emission (Stelzer et al. 2012). Various color–color diagrams (CCDs) and color–magnitude diagrams (CMDs) are efficient in selecting sources with IR excess emission. However, these selections are always contaminated by non-YSO sources, such as broad-line active galactic nuclei (AGNs), star-forming galaxies, and asymptotic giant branch (AGB) stars, that have IR properties similar to those of YSOs (Robitaille et al. 2008; Koenig et al. 2012; Manara et al. 2018; Herczeg et al. 2019; Lee et al. 2021). In addition, these CCDs and CMDs are inefficient in selecting evolved YSOs without circumstellar material, which have colors similar to those of main-sequence stars.

Accurate astrometric measurements from the Gaia mission (Gaia Collaboration et al. 2016) have changed this situation. With measurements of its distance and proper motion of individual sources, membership can be assessed directly, without any assumptions for the source properties. Different methods have been developed to identify new members in nearby star-forming regions, based on Gaia astrometric data, including the following examples. Using the data set from Gaia DR2 (Gaia Collaboration et al. 2018), Cánovas et al. (2019) identified 166 new candidates in the ρ Ophiuchi region applying different clustering algorithms, including DBSCAN (Ester et al. 1996), HDBSCAN (Campello et al. 2013, 2015; McInnes et al. 2017), and OPTICS (Ankerst et al. 1999). Also in the ρ Ophiuchi region, Grassier et al. (2021) identified additional ~200 new YSO candidates applying the OCSVM algorithm developed by Ratzenböck et al. (2020), based on the data set from Gaia EDR3 (Gaia Collaboration et al. 2021). Applying a simple thresholding method and utilizing the data set from Gaia DR2, Luhman (2018) refined the sample of known members and identified new candidates in the Taurus star-forming region. Kubiak et al. (2021) increased the census of known ε Cha members by more than 40%. Both simple thresholding methods and clustering algorithms are powerful tools for identifying new members with high reliability in nearby star-forming regions.

The Perseus molecular cloud is one of the nearest (~300 pc from the Sun, Ortiz-León et al. 2018) low mass star-forming regions (Reipurth 2008). It is a well-studied star-forming region (e.g., Sargent 1979; Ladd et al. 1993; Kirk et al. 2006; Ridge et al. 2006a; Enoch et al. 2007; Arce et al. 2010) with a range of star-forming environments, both clustered (e.g., the two young clusters NGC 1333 and IC 348) and distributed, as well
as several small dense clumps (B1, B5, L1448, and L1455) of the type that often produces one or a few stars (Young et al. 2015). The Perseus molecular cloud has been observed in multiwavelength studies (e.g., X-ray: Winston et al. 2010; Alexander & Preibisch 2011; Optical: Zhang et al. 2015; IR: Young et al. 2015; submillimeter: Kirk et al. 2006; Enoch et al. 2007; radio: Ridge et al. 2006a; Arce et al. 2010). Many YSOs have been identified in this region previously (Evans et al. 2009; Hsieh & Lai 2013; Young et al. 2015). In a census of members of IC 348 and NGC 1333 using multi-epoch IRAC astrometry, Luhman et al. (2016) identified many members in the two clusters. However, for the remaining cloud regions, only IR excess emission was used to select large samples of YSO candidates (YSOc; Evans et al. 2009; Hsieh & Lai 2013; Young et al. 2015), so many undiscovered YSOs without disks may be present in the remaining cloud regions. Recently, Pavlidou et al. (2021) identified five new groups in the Perseus star-forming complex based on the astrometric data from Gaia DR2 (Gaia Collaboration et al. 2018), but they mainly focused on the off-cloud regions specifically, instead of the main cloud regions outside the two young clusters.

In this paper, we use Gaia astrometry and photometry to identify candidate members in the Perseus molecular cloud, especially in the remaining cloud regions outside the two clusters, and use the LAMOST spectroscopic survey to confirm the memberships of these candidates. We describe the data sets used in this work in Section 2, and the source selection procedures in Section 3. The results are presented in Section 4 and discussed in Section 5. We give our summary and conclusions in Section 6.

2. Data sets

2.1. Gaia Data

We use the astrometric and photometric data from the Gaia survey (Gaia EDR3) (Gaia Collaboration et al. 2021; Riello et al. 2021; Fabricius et al. 2021) to identify new candidates in the Perseus molecular cloud. We consider only the main cloud region that is covered by the $^{12}$CO $J = 1 \rightarrow 0$ map from the COMPLETE (Coordinated Molecular Probe Line Extinction and Thermal Emission) Survey of Star-forming Regions (Ridge et al. 2006b), corresponding to $15^\circ < l < 161^\circ$ and $-22^\circ < b < -16^\circ$. Gaia satellite provides us with very accurate astrometric measurements of almost 2 billion sources, including ~60,000 within our survey area.

Based on the previously proposed distances to the Perseus molecular cloud, we include only sources with parallaxes between 2.2 and 5.0 mas, corresponding to distances of about 200–450 pc. Only sources whose observations are consistent with the five-parameter model (Lindgren et al. 2018) are retained for further analysis, that is we keep only sources with $ruw < 1.4$ (Gaia Collaboration et al. 2021; Lindgren 2018). Additional quality cuts are applied to extract sources with high-quality astrometry and photometry. Only sources with parallaxes over errors larger than 5 and with proper motion errors less than 0.5 mas are considered to produce an astrometrically precise and reliable data set. Sources with magnitude errors in $G$ or $RP$ bands larger than 0.1 mag are removed to produce a high-quality catalog. With these cuts, we extract about 3000 Gaia sources with high quality toward the direction of the Perseus molecular cloud.

2.2. Optical Spectroscopic Data from the LAMOST Survey

We use the spectroscopic data from data release 7 of the LAMOST survey (LAMOST DR7; Luo et al. 2022) to confirm the memberships of the candidates in the cloud. LAMOST, the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (also called the Guoshoujing Telescope), is a quasi-meridional Schimd telescope located at Xinglong Observatory Station in Hebei, China. The telescope has an effective aperture of 3.6–4.9 m and a field of view of about 5° in diameter. The telescope is equipped with 16 spectrographs and 4000 fibers, each spectrograph has a resolution of $\sim 2.5 \, \AA$, and the wavelength coverage is 3690–9100 Å (Cui et al. 2012; Zhao et al. 2012; Liu et al. 2015).

LAMOST DR7 contains more than 10 million spectra, more than 9.5 million of which are stellar spectra. In the direction of the Perseus molecular cloud, LAMOST obtained more than 9000 spectra of about 5500 distinct sources.

2.3. Additional Ancillary Data

The Perseus molecular cloud has been well covered by many large-sky surveys, including Pan-STARRS1 (Hodapp et al. 2004), Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), and Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010). The Pan-STARRS1 (PS1) survey images the whole sky in five broadband filters, $grizy$, with wavelength coverage from 0.4–1 $\mu$m (Stubbs et al. 2010). We retrieve the $griz$ photometry from the PS1 DR1 catalog (Chambers et al. 2016). Sources that are brighter than 14 mag or that have magnitude errors larger than 0.05 mag, in the $r$ or $i$ filters, are removed to avoid saturation problems and to produce a high-quality catalog. The $HJK_S$ magnitudes are retrieved from the 2MASS All-Sky Point Source Catalog (Skrutskie et al. 2003), which reaches limiting magnitudes of 15.8, 15.1, and 14.3 mag at 10 $\sigma$ for the $J$, $H$, and $K_S$ bands, respectively (Skrutskie et al. 2006). The WISE survey is a mid-infrared full-sky survey undertaken in four bands: W1, W2, W3, and W4 bands with wavelengths centered at 3.35, 4.60, 11.56, and 22.09 $\mu$m, respectively (Wright et al. 2010). We take the WISE photometry from the AllWISE source catalog (Wright et al. 2019) to determine the presence or absence of circumstellar disks around YSOs.

3. Source Selection

3.1. Census from Previous Studies

Luhman et al. (2016) compiled a thorough census for the two young clusters, and provided samples of 478 and 203 sources with well-confirmed memberships for IC 348 and NGC 1333, respectively. Beginning with this tabulation, we search the literature for additional members in the Perseus molecular cloud, especially in the regions outside the two clusters, with spectroscopically confirmed memberships. We add to our sample the eight and two members identified by Esplin & Luhman (2017) and Luhman & Hapich (2020), respectively. From the Kounkel et al. (2019) search for close companions around young stars, we include in our census the 114 members that have radial velocities between 0 and 25 km s$^{-1}$. This velocity range is consistent with the radial velocities of members in IC 348 (Cottaar et al. 2015) and NGC 1333.
(Foster et al. 2015). Kounkel et al. (2019) reported a radial velocity of 143 km s\(^{-1}\) for the source 2MASS J03453345 + 3145553, but Cottaar et al. (2015) reported a radial velocity of around 20 km s\(^{-1}\) from multiple observations, so we retain this object in our member list. Combining all these tabulations together, we arrive at a sample of 805 known members with spectroscopically confirmed memberships in the Perseus molecular cloud.

Cross matching this list of known members with the Gaia catalog described in Section 2.1, using 2\(^{\prime}\) tolerance, we obtain astrometric and photometric measurements for 406 known members (~50\% of the sample). It is not surprising that about half of the known members are not matched with the Gaia catalog, since we consider only sources with high-quality Gaia measurements as described in Section 2.1. We would obtain 598 matches, i.e., ~75\% of the sample, if no quality cuts were applied to the Gaia catalog. In addition, many of the known members are either too faint (~150 members in our sample are fainter than \(J = 16\) mag, all without matches in the Gaia catalog) or too embedded (~60 members in the census are deeply embedded class 0/I protostars) to be observed by the Gaia satellite.

Visually inspecting the properties of the 406 matched members, we find that the bulk of these members share common distances, motions, and ages, with limited scatter, as expected from their memberships. Figure 1 displays the distributions of parallaxes and proper motions, as well as the \(G - RP\) versus \(G\) CMD for these matched known members.

In the parallax distribution in Figure 1, nearly all members have parallaxes within the 3\(\sigma\) interval from the weighted mean parallax. The weighted mean and the weighted standard deviation of the parallax are \(\overline{\pi} = 3.29\) mas and \(\sigma_\pi = 0.19\) mas, respectively, calculated by adopting the inverse of the parallax errors as the weight for each parallax.

In the proper motion distribution in Figure 1, 394 of the matched members have proper motions within the 3\(\sigma\) confidence ellipse (the red ellipse in the plot). The confidence ellipse is calculated as follows. Visually inspecting the distribution of the proper motions, we identify six sources with extreme proper motions (outside the viewing range of the plot). In fact, these outliers are companions around young stars (Kounkel et al. 2019). The weighted means and the weighted standard deviations of the proper motion are \(\overline{\mu_\alpha^*} = 5.3\) mas yr\(^{-1}\), \(\overline{\mu_\delta} = -7.2\) mas yr\(^{-1}\) and \(\sigma_{\mu_\alpha^*} = 1.43\) mas yr\(^{-1}\), \(\sigma_{\mu_\delta} = 1.54\) mas yr\(^{-1}\), respectively, calculated by adopting the inverse of the proper motion errors as the weight for each proper motion, excluding the six extreme cases. We also calculate the Pearson correlation coefficient of \(-0.7\) between \(\mu_\alpha^*\) and \(\mu_\delta\) excluding the six extreme cases, considering measurement uncertainties of both axes, by using the Monte Carlo method proposed by Curran (2014). Using the weighted means, the weighted standard deviations, and the correlation coefficient, we construct the 3\(\sigma\) confidence ellipse.

From the CMD in Figure 1, we find that all but one members are above or around the 10 Myr isochrone from the PARSEC stellar model (Bressan et al. 2012) with solar metallicity, scaled to a distance of 300 pc. The only source well below the 10 Myr isochrone is 2MASS J03291082 + 3116427. Considering its location on the CMD and the presence of a disk (Luhman et al. 2016), this object is likely harboring an edge-on disk.

### 3.2. Candidates and New Members

In the previous section, we find that nearly all the known members have parallaxes within the 3\(\sigma\) of the median parallax for the Perseus members and have proper motions within the 3\(\sigma\) confidence ellipse, as well as that all members are above or around the 10 Myr isochrone. Based on these statistical properties of the known members, we identify new candidates from the Gaia catalog. Specifically, we select a source as a new candidate if it has parallax between 2.71 and 3.87 mas and its proper motion is within the 3\(\sigma\) confidence ellipse defined in the previous section. We also require that the candidate is above the 10 Myr isochrone. As the PARSEC stellar model does not extend down to masses below the hydrogen burning limit, we artificially extend the 10 Myr isochrone vertically to include low mass members. This extension line fits the low mass members fairly well. With these constraints, we select 211 additional candidate members from the Gaia catalog.

Kounkel et al. (2022) studied the Per OB2 association using Gaia EDR3, applying the clustering algorithm HDBSCAN (Campello et al. 2013, 2015; McInnes et al. 2017). In the same region studied in our work, they discovered 129 new candidate members. In this work, we recover most of these candidates, and identified 109 new YSO candidates. These sources are distributed in the regions with low YSO surface density (~50 deg\(^{-2}\)), and are missed by the HDBSCAN algorithm used in Kounkel et al. (2022).

By cross matching with the LAMOST archive, we obtain optical spectra for 55 candidates. Of these candidates, 38 have broad H\(\alpha\) emission lines (characteristic of ongoing accretion activities, Hartmann et al. 1994; Muzerolle et al. 2001) or Li I \(\lambda 6707\) Å absorption lines (indicator of youth) in their optical spectra (as shown in Figure 2), indicating that they are bona fide members of the Perseus molecular cloud. Detailed analyses of these emission lines and corresponding accretion activities will be presented in a future paper (X.-L. Wang et al., in preparation). For the remaining 17 candidates, we do not detect significant emission lines or Li I \(\lambda 6707\) Å absorption lines in their optical spectra. These 17 candidates include eight A-type, four F-type, two G-type, one K-type, and two M-type stars (see Section 4.1 for details of the spectral classification) and most of them show clear H\(\alpha\) absorption lines in their optical spectra. Inspecting their locations on the Hertzsprung–Russell (H–R) diagram, we confirm 13 of them as young members. The remaining four objects are classified as main-sequence stars or AGB stars and are rejected from our list of candidates (see Section 5.1 for a discussion).

In summary, we identify 211 candidate members sharing common distances, motions, and ages in the Perseus molecular cloud and confirm 51 of them as bona fide members, based on evidence of youth, including the presence of emission lines and Li I \(\lambda 6707\) Å absorption lines in their optical spectra, and their locations on the H–R diagram. We confirm more than 90\% (51/55) of the candidates have LAMOST spectra as Perseus members. These new members and candidates are listed in Table 1. The spatial distributions of all members (805 known members and 51 new members) and candidates (156 candidates) are displayed in Figure 3. Nearly all the members and candidates newly identified in this work are spread over the whole cloud outside the two young clusters.

\(^{10}\) Twenty-seven sources are excluded due to large \(\alpha\) values, large photometric errors (Section 2.1) or being older than 10 Myr.
4. Stellar Properties

4.1. Spectral Types

In this section, we estimate spectral types for the newly confirmed members, listed in Table 1. Due to the complexity of classifying YSOs with late spectral types (emission lines and molecular bands), the LAMOST pipeline (Wu et al. 2011; Luo et al. 2015) may give incorrect spectral types during the chi-square fitting procedures. Thus, for objects with automated spectral types from the LAMOST pipeline later than K0, we measure spectral type from the LAMOST spectra by applying the classification scheme from Fang et al. (2017), which focuses on specific spectral features (i.e., VO, TiO, and CaH bands). For sources with earlier types, we adopt the spectral classifications from the LAMOST archive directly. To test the validity of these classification schemes, we apply these classification schemes to a subsample of the known members that have been observed by LAMOST as well. The spectral types obtained here and that from the literature are consistent with each other (Figure 4).

The distributions of spectral types for previously known and newly confirmed members are displayed in Figure 5. Most of the new members are late K and M types. Although we identify many A-, F-, and G-type members, the overall distribution of the spectral types remains unchanged. Due to the sensitivity of the LAMOST survey, we did not select any new members later than about M6.

4.2. Extinction Corrections

The extinction of each star is determined individually. Intrinsic $g - r$, $r - i$, $i - z$, $z - J$, $J - H$, and $H - K_s$ colors are estimated with corresponding spectral types for the stars with determined spectral types. For sources later than F0, the relationship from Fang et al. (2017) is adopted, and for earlier type stars that from Covey et al. (2007) is used. These intrinsic colors are converted from the Sloan Digital Sky Survey photometric system to the Pan-STARRS1 photometric system (Tonry et al. 2012), and the extinction values are determined,
Figure 2. Example spectra of a representative sample of the newly confirmed members. The prominent emission lines and the Li\textsc{i}$\lambda$6707 Å absorption lines are marked in each panel, if present. The Gaia EDR3 source designations and the corresponding spectral types are labeled as well.

Table 1
Properties of New Members and Candidates

| Gaia EDR3       | $\alpha$ | $\delta$ | $v$ | $\mu_\alpha$ | $\mu_\delta$ | SPT  | A_V  | Membership |
|-----------------|----------|----------|-----|---------------|---------------|------|------|------------|
| (deg)           | (deg)    | (mas)    |     | (mas/yr)      | (mas/yr)      |      | (mag)|            |
| 12044379950667098 | 55.085941 | 30.986461 | 3.39 | 6.37          | -7.86         | K9.7| 1.2 | Member     |
| 120460024334304128 | 54.553027 | 31.074960 | 3.35 | 6.20          | -8.28         | M3.7| 0.5 | Member     |
| 120463116710737920 | 54.875402 | 31.110591 | 3.40 | 6.69          | -8.81         | M4.2| 0.5 | Member     |
| 120903501181107968 | 52.162684 | 30.251784 | 3.45 | 7.97          | -10.13        | ... | ... | Candidate  |
| 120934910278041600 | 52.333973 | 30.563745 | 3.53 | 7.59          | -10.34        | ... | 0.42| Candidate  |
| 120952429448891520 | 52.122005 | 30.726234 | 3.05 | 4.18          | -5.18         | ... | 3.45| Candidate  |
| 217475121620315264 | ...      | ...      | ... | ...           | ...           | ... | ... | ...        |
| 217481269670674048 | 55.806566 | 32.607210 | 3.11 | 3.08          | -3.37         | F0  | 0.2 | Member     |
| 217482850218656512 | 55.893907 | 32.724620 | 3.32 | 4.11          | -7.29         | ... | 1.28| Candidate  |
| 217510376662979152 | 55.243254 | 32.506634 | 3.43 | 6.63          | -9.60         | K5  | 0.4 | Member     |
| 217510892061734208 | 55.156662 | 32.497486 | 3.40 | 7.17          | -9.08         | ... | ... | Candidate  |
| 217863938374329472 | 56.342371 | 32.921457 | 3.25 | 7.45          | -10.73        | ... | 0.76| Candidate  |

Note. The table is available in its entirety in a machine-readable format.

(This table is available in its entirety in machine-readable form.)
combined with observed colors and the extinction law from Wang & Chen (2019).

For sources lacking determined spectral types, we can only roughly determine their extinctions with the method described in Fang et al. (2013), in which the extinctions are obtained by employing the \(H - K_S\) versus \(J - H\) CCD. Observed \(J - H\) and \(H - K_S\) colors are compared to intrinsic colors of main-sequence stars (Bessell & Brett 1988) and T Tauri stars (Meyer et al. 1997), and extinction values are estimated for individual stars, using the extinction law from Wang & Chen (2019) (see Figure 6). The typical error of the extinction values in the \(V\) band is \(\sim 0.6\) mag and may be higher for stars with accretion disks.

These extinction values are listed in Table 1 and displayed in Figure 7. The median values of extinction for previously known members, new members, and candidates are \(A_V = 3.6, 2.2,\) and \(2.2\) mag, respectively. A two sample K-S test also indicates that newly confirmed members and candidates have
smaller extinction values than previously known sources. This result is consistent with the fact that most of our newly confirmed members and candidates are spread evenly across the cloud, located at less extincted regions, while most previous searches focused on the two clusters.

5. Discussion

5.1. Membership Confirmations

We assess memberships for about 70% of the candidates having LAMOST spectra as bona fide members based on the existence of Hα emission lines or Li I λ 6707 Å absorption lines in their optical spectra in Section 3.2. Similar to Cieza et al. (2012), we assess memberships for the remaining candidates having LAMOST spectra, but lacking emission lines or Li I λ 6707 Å absorption lines in their optical spectra, according to their locations on the H-R diagram in this section.

To construct the H-R diagram, we convert the spectral types to effective temperatures using the empirical relations determined by Fang et al. (2017) for members later than F0 types and by Pecaut & Mamajek (2013) for earlier spectral types. The observed J magnitudes are dereddened using the extinction values determined in Section 4.2 and distance corrected assuming distances of 300 pc for all sources (see the red filled histogram in panel (a) in Figure 1). The stellar luminosities are then calculated as the following:

$$\log \left( \frac{L_*}{L_\odot} \right) = - \frac{BC_T + M_J - M_{bol,\odot}}{2.5},$$

where $M_J$ is the extinction corrected absolute J magnitude, $BC_T$ is the bolometric correction in the J band, and $M_{bol,\odot}$ is the bolometric magnitude of the Sun. We take $M_{bol,\odot} = 4.755$ mag (Mamajek 2012) in this work. The $BC_T$ values from Fang et al. (2017) are adopted for spectral types later than F0 and that from Pecaut & Mamajek (2013) for earlier spectral types.

As shown in the H-R diagram (Figure 8), sources having emission lines or Li I λ 6707 Å absorption lines in their optical spectra are all above or around the 10 Myr isochrone, as expected from their memberships. Among the remaining 17 objects lacking emission lines or Li I λ 6707 Å absorption lines in their optical spectra, 13 are located above or around the 10 Myr isochrone and are consistent with being members in the Perseus molecular cloud, three are well below the 10 Myr isochrone (marked with open diamonds), and are likely to be field main-sequence stars, and one object is well above the stellar birth line (marked with a square) and is a possible AGB star.
the remaining cloud regions

Figure 9. Left: histograms of the stellar surface densities for members in different cloud regions, i.e., IC 348 (black open histogram), NGC 1333 (hatched histogram), the remaining cloud regions (gray open histogram), and Autochthe (blue bars). Right: the same as in the left panel, but for densities including candidate members. The red bars indicate the Barnard 1 group found in Section 5.2 and marked in Figure 3, stacked on top of that for Autochthe.

5.2. Stellar Surface Densities and Substructures in the Perseus Cloud

YSOs tend to form in clusters and are expected to have higher surface densities than field stars. As mentioned above, there are both clustered and distributed star formation in the Perseus molecular cloud. We have divided the cloud into three regions (i.e., NGC 1333, IC 348, and the remaining cloud regions, following Rebull et al. 2007). We calculate the stellar surface densities \( \Sigma \) for individual YSOs as (Casertano & Hut 1985),

\[
\Sigma = \frac{n - 1}{\pi r_n^2},
\]

where \( n \) is the \( n \)th nearest star and \( r_n \) is the distance to the \( n \)th nearest star. We adopt \( n = 6 \) as a surface density reference, the same as that used in Gutermuth et al. (2009).

The resulting stellar surface density distributions are shown in Figure 9 for different cloud regions. The median densities are 3298, 4368, and \( 42 \text{ deg}^{-2} \) in IC 348, NGC 1333, and the remaining cloud regions, respectively, if only confirmed members (previously known members and newly confirmed members) are included in the density calculations. The stellar surface densities are much higher in the two clusters than in the remaining cloud regions. The star formation environment in the remaining cloud regions is different from that in the clustered regions. The remaining cloud regions represent an environment for distributed star formation. Including the candidate members in the calculation does not change the stellar surface densities in the cluster regions (3236 and 4378 \( \text{deg}^{-2} \) for IC 348 and NGC 1333, respectively), but nearly doubles the stellar density in the remaining cloud regions (69 \( \text{deg}^{-2} \)).

Visual inspection of Figure 3 shows two additional groupings of YSO(c)s with enhanced stellar densities outside the main clusters IC 348 and NGC 1333. The stellar aggregate toward the western part of the cloud consists of six known members, 10 new members, and 12 new candidates. This group is associated with the dark cloud L1448 (Lynds 1962) and was originally identified by Pavlidou et al. (2021), who named the group Autochthe. The stellar aggregate to the east of NGC 1333 consists of 14 known members, two new members, and six new candidates. This group is associated with the dark cloud Barnard 1 (Barnard 1919) and we designate it the “Barnard 1 group”.

We also perform the density-based clustering algorithm DBSCAN (Ester et al. 1996) on our data set for further examination. When only confirmed members are considered, the algorithm returns 3 clusters, i.e., IC 348, NGC 1333, and Autochthe. The algorithm returned 4 clusters, i.e., IC 348, NGC 1333, Autochthe, and the Barnard 1 group, when the candidates identified in this work are included in the input data set. These results are consistent with our visual inspection.

5.3. Distance Gradient

A single distance can not describe the whole cloud accurately, due to the complex structures of the Perseus molecular cloud. Ortiz-León et al. (2018) obtained distances of 321 pc and 293 pc for IC 348 and NGC 1333 respectively, based on VLBA and Gaia measurements.\(^{11}\)

In this work, we identify about 200 new members and candidates in the Perseus molecular cloud outside the two clusters, and these members spread nearly evenly across the cloud. With these distributed YSOs, as well as the clustered

\(^{11}\) Multiple measurements of the distance to the Perseus molecular cloud have been performed in the past. Cernis (1990) proposed a distance of 200–250 pc to the Perseus molecular cloud according to extinction studies. Very long baseline interferometry studies of H2O masers gave distances of 235 pc for NGC 1333 (Hirota et al. 2008) and 232 pc toward L1448 (Hirota et al. 2011). A distance of 240 pc was estimated by comparing the density of foreground stars with the Robin et al. (2003) Galactic model (Lombardi et al. 2010). A distance of 250 pc toward the Perseus molecular cloud has been widely used by various studies (e.g., Young et al. 2015; Zhang et al. 2015; Enoch et al. 2006).
YSOs, we explore potential distance gradient toward the cloud. We measure distances of 310 and 291 pc toward IC 348 and NGC 1333, respectively, consistent with that measured by Ortiz-León et al. (2018) within 1σ. In Figure 10, we plot the changes in distances along the R.A. direction for all members. We also calculate the median distances of all members in corresponding R.A. bins. As shown in Figure 10, the cloud is getting closer toward us gradually, from east to west. The Spearman’s rank correlation coefficient between R.A. and distance is $|\rho| = 0.88$ with $p = 0.4\%$, indicating a statistically significant correlation between R.A. and distance and that the distance gradient is real. The distance gradient is 4.84 pc deg$^{-1}$ from NGC 1333 toward IC 348.

Sargent (1979) found a smoothly varying LSR velocity across the cloud from their CO observations, and Enoch et al. (2006) pointed out that there may also be a distance gradient across the cloud given the velocity gradient. However, we are unable to convert the LSR velocity gradient to a distance gradient directly, since the Perseus molecular cloud is located toward the Galactic anti-center. With our updated member list, we quantify this distance gradient. But it is worth noting that we can not rule out the possibility that the clusters reside in different clouds along the line of sight rather than in a single cloud (Bally et al. 2008).

5.4. Evolutionary Stages in Different Cloud Regions

In this section, we investigate the evolutionary stages in different cloud stages in different cloud regions. Bell et al. (2013) measured an age of $\sim$6 Myr for IC 348, whereas an age of $<1$ Myr was measured for NGC 1333 (Wilking et al. 2004). Young et al. (2015) obtained a similar conclusion based on the fractions of YSOs at different evolutionary stages. Dividing the Perseus molecular cloud into three regions, IC 348, NGC 1333 clusters, and the remaining cloud regions (Rebull et al. 2007), we construct the dereddened and distance corrected $r - i$ versus $r$ CMDs for each region (Figure 11). For the cluster regions, we use the median distance for each cluster to do the distance correction, and individual distances are used for individual YSOs in the remaining cloud regions. We fit isochrones from the PARSEC stellar model (Bressan et al. 2012) and estimate ages of 5.4, 2.9, and 5.7 Myr for IC 348, NGC 1333, and the remaining cloud regions, respectively. The ages of the remaining cloud regions may be underestimated, since we select only candidates above the 10 Myr isochrone on the $G - RP$ versus $G$ CMD (panel (c) in Figure 1).

Recently, Kounkel et al. (2022) obtained much younger ages of 2.5 Myr for both IC 348 and NGC 1333, by isochrone fitting of the Gaia photometry to the MIST isochrones (Choi et al. 2016). This discrepancy is mainly due to the difference between the PARSEC stellar model (Bressan et al. 2012) and the MIST isochrones (Choi et al. 2016). We obtain ages of 3.1 and 2.0 Myr for IC 348 and NGC 1333, respectively, performing isochrone fitting to the MIST isochrones (Choi et al. 2016). We obtain a much older age for IC 348 than for NGC 1333, which is consistent with a higher fraction of diskless objects in IC 348 than in NGC 1333 (Young et al. 2015; Luhman et al. 2016).

We obtain an age for IC 348 similar to that in previous studies (Bell et al. 2013; Luhman et al. 2016). Luhman et al. (2016) proposed an older age for NGC 1333 than those found in previous studies (e.g., Wilking et al. 2004), and they attributed this discrepancy to the distance they assumed for NGC 1333. We obtain a much younger age than Luhman et al. (2016), with our updated distance measurements. The remaining cloud regions are significantly older than NGC 1333 and nearly coeval with IC 348. Young et al. (2015) found that the number ratio of PMS stars to protostars in the remaining cloud regions is similar to that in NGC 1333, but significantly smaller than that in IC 348. This discrepancy can be partly explained by the fact that we estimated ages based on optical CMDs, which may bias toward members at later evolutionary stages and thus older ages.

We estimate the fraction of sources harboring disks as another proxy for evolutionary stages of different cloud regions. Luhman et al. (2016) estimated disk fractions of around 40% and 60% for IC 348 and NGC 1333, respectively, with disk fraction defined as $N(II)/N(II + III)$. Based on the source counts reported in Young et al. (2015) (their Table 4), we estimate disk fractions as high as 82%, 92%, and 89% for IC 348, NGC 1333, and the remaining cloud regions, respectively. The discrepancy between the two is mainly due to Young et al. (2015) using only infrared-selected YSOs, whereas Luhman et al. (2016) included many optically visible members, to calculate the disk fractions.

Since many of the new members and candidates identified in this work lack enough Spitzer data to determine the presence or absence of a disk, we use the $K_S - W2$ versus $H - K_S$ CCD (Figure 12) to determine the presence or absence of a disk. As shown in the plot, the sources between the two dashed lines are reddened dwarfs. We consider objects with $K_S - W2 > 0.98 \times (H - K_S) + 0.22$ to be those harboring disks. We estimate disk fractions of 45%, 65%, and 30%–40% for IC 348, NGC 1333, and the remaining cloud regions, respectively, with our updated member list. We estimate similar disk fractions as that estimated by Luhman et al. (2016) for IC 348 and NGC 1333. Including or excluding the new members and candidates in the calculation has little impact on the disk.
fractions for IC 348 and NGC 1333, since we select few new members and candidates in these regions. But for the remaining cloud regions, we measure a much lower disk fraction compared to that excluding the new members and candidates in the calculation (30%–40% compared to 55%). We find that the remaining cloud region is more coeval with IC 348 than with NGC 1333. This result suggests the importance of including Gaia astrometric data in identifying YSOs at more evolved stages, with star-like spectral energy distributions.

5.5. A Possible Scenario for Star Formation in the Perseus Molecular Cloud

Recently, Zucker et al. (2022) proposed that star formation near the Sun is driven by expansion of the Local Bubble (Cox & Reynolds 1987) with the stellar traceback method. They found that nearly every well-known molecular cloud in the vicinity of the Sun lies on the surface of the Local Bubble, except the Perseus molecular cloud. They attributed this exception to the displacement of the recently discovered

![Figure 11. Pan-STARRS1 $r$ vs. $r$ CMDs for known (blue circles) and new (red squares) members of the Perseus molecular cloud, divided into three regions, IC 348 (left), NGC 1333 (middle), and the remaining cloud regions (right). The points have been dereddened and distance corrected. In each panel, the black lines are 1, 3, 10, 30, and 100 Myr isochrones, from top to bottom, from the PARSEC stellar model (Bressan et al. 2012), with solar metallicity, and the red solid line is the best-fitting isochrone. The resulting ages are labeled.](image1)

![Figure 12. Left panel: infrared color–color plot for members and candidates in IC 348. The blue circles, red squares, and orange diamonds mark known members, new members, and candidate members, respectively. The black solid line is the locus of dwarfs from Pecaut & Mamajek (2013), and the dashed lines correspond to the extinction law from Wang & Chen (2019), enclosing the color space of dwarfs due to reddening. The dashed–dotted line is 0.15 mag (about 3 times the typical uncertainties in $K_S – W2$ colors) redder than the right border of the dwarf locus, serving as the boundary between disk and diskless YSOs. The middle and right panels are the same as the left panel, but for NGC 1333 and the remaining cloud regions, respectively.](image2)
Per-Tau Shell (Bialy et al. 2021), containing Perseus on its far side and Taurus on its near side. The stellar traceback data provided by Zucker et al. (2022) indicate that IC 348 and NGC 1333 have been moving away from the center of the Per-Tau Shell since about 6–8 Myr ago, and that IC 348 has already escaped the shell whereas NGC 1333 is passing through the surface of the shell (as illustrated in the left panel of Figure 13).

Since we double the census of members in the remaining cloud regions, we can investigate the distributed populations with respect to the feedback from the Per-Tau Shell or the Local Bubble. As these distributed populations spread over large areas, it is improper to apply the stellar traceback method to these data, considering the remaining cloud regions as a whole. We investigate motions and locations of individual members relative to the center of the Per-Tau Shell, nearly all of these distributed members have velocities pointing outward the shell with separation angles less than 90° (as shown in the right panel of Figure 13), about 60% of them have escaped the shell (see the scatter plot in the left panel of Figure 13). These facts may point to the scenario that expansion of the Local Bubble drives the formation of the surface clouds (Zucker et al. 2022) as well as the Perseus molecular cloud, and subsequent expansion of the Per-Tau Shell blows away the Perseus molecular cloud from the Local Bubble’s surface and results in elongated shape of the cloud. This elongation may partly explain the distance gradient confirmed in Section 5.3.

We have searched the literature as well as the Gaia archive for possible massive stars in the vicinity of the cloud that may be responsible for creating the Per-Tau Shell, but no massive stars were found within the region we studied. Bialy et al. (2021) estimated that two to eight supernovas may be responsible for creating such a large shell structure. Since our study is focused on the main cloud region only, a detailed analysis of searching for the driving source of the shell is beyond the scope of the current study.

Visualizing the five groups identified by Pavlidou et al. (2021) (see Figure 14), we find that the three groups, Electryon, Heleus, and Mestor are far away from the Per-Tau Shell, the remaining two groups Alcaeus and Autochthe, as well as NGC 1333 and IC 348 are near the surface of the shell. Considering their ages (3–5 Myr, Pavlidou et al. 2021) and their distances to the shell, the formation of the three off-shell groups is irrelevant to the Per-Tau Shell’s expansion. Alcaeus shares an age similar to that of IC 348 and Autochthe is coeval with NGC 1333. These groups may form during the same star formation event.
6. Summary

We performed a search for new members and candidates in the Perseus molecular cloud, using spectroscopic data from the LAMOST survey and astrometric data from the Gaia survey. The spectral types and extinction corrections are determined for the newly confirmed members. The spatial distributions of the updated member list and the age differences in different cloud regions are investigated. The main contributions are summarized as follows.

1. We performed a thorough census for new members and candidates in the Perseus molecular cloud and identified 211 candidate members sharing common distances, motions, and ages. We identified 51 candidates as new members of the cloud based on evidence of youth including the presence of emission lines or Li I λ 6707 Å absorption lines in their optical spectra and their locations on the H-R diagram, which brings the total number of known members to 856. Due to the sensitivity of the LAMOST survey, we did not identify any new members later than M6.

2. The stellar aggregate named Autochthone in Pavlidou et al. (2021) is confirmed as a real coeval stellar aggregate with a dozen of confirmed members, with our revised member list. We also identify another small aggregate associated with the dark cloud Barnard 1.

3. The new members are less extincted, and reside in regions with low surface densities. Cloud members in the remaining cloud regions represent the type of distributed star formation, as well as star formation of the type that often produces one or a few stars.

4. A statistically significant distance gradient of 4.84 pc deg^{-1} is measured from west to east. But it is worth noting that the possibility that the two clusters reside in different clouds along the line of sight rather than a single cloud still exists.

5. We estimate similar ages for IC 348 and the remaining cloud regions and NGC 1333 is much younger than the two. The disk fraction in NGC 1333 is higher than elsewhere, consistent with its youngest age.

6. The bulk motion of the distributed populations is pointed outward of the Per-Tau Shell, and the five groups identified by Pavlidou et al. (2021) are moving away from the shell as well. These are consistent with the star formation scenario proposed by Zucker et al. (2022).

We acknowledge the support of the National Natural Science Foundation of China (NSFC, grant No. 11303073) and National Key R&D Program of China (No. 2017YFA0402700). H.Z. gives thanks for the support from the National Key R&D Program of China (2017YFA0402702), the NSFC grant (Nos. 11973039 and 12122303) and the CAS Pioneer Hundred Talents Program. G.J.H. acknowledges the support of the National Natural Science Foundation of China (NSFC, grant No. 12173003). We thank A-Li Luo, Jian-Jun Chen, Wen Hou, and other staff from NAOC for their help in downloading the data and for their comments on the data reduction. 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, in particular the institutions participating in the Gaia Multilateral Agreement. Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). We also acknowledge the various Python packages that were used in the data analysis of this work, including Matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Shapely (Gillies et al. 2007), Scikit-Learn (Pedregosa et al. 2011).

References

Alexander, F., & Preibisch, T. 2011, The X-ray Universe 2011, ed. J.-U. Ness & M. Ehle, 184
Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. 1999, SIGMOD Rec, 28, 49
Arce, H. G., Borkin, M. A., Goodman, A. A., Pineda, J. E., & Halle, M. W. 2010, ApJ, 715, 1170
Astropy Collaboration, Price-Whelan, A. M., Sipocz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bally, J., Walawender, J., Johnstone, D., Kirk, H., & Goodman, A. 2008, in Handbook of Star Forming Regions, Vol. I: The Northern Sky, ed. B. Reipurth (San Francisco, CA: ASP), 308
Barnard, E. E. 1919, ApJ, 49, 1
Bell, C. P. M., Naylor, T., Mayne, N. J., Jeffries, R. D., & Littlefair, S. P. 2013, MNRAS, 434, 806
Bessell, M. S., & Brett, J. M. 1988, PASP, 100, 1134
Bialy, S., Zucker, C., Goodman, A., et al. 2021, ApJL, 919, L5
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013, in Advances in Knowledge Discovery and Data Mining, ed. J. Pei et al. (Berlin: Springer), 160
Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. 2015, ACM TKDD, 10, 1
Cánovas, H., Cañtero, C., Cieza, L. et al. 2019, A&A, 626, A80
Casertano, S., & Hut, P. 1985, ApJ, 298, 80
Cernis, K. 1990, AP&SS, 166, 315
Chambers, K. C., Hut, P. 1985, ApJ, 298, 80
Cieza, L. A., Schreiber, M. R., Romero, G. A., et al. 2012, ApJ, 750, 157
Cottaar, M., Covey, K. R., Foster, J. B., et al. 2015, ApJ, 807, 27
Covey, K. R., Ivezic, Z., Schlegel, D., et al. 2007, AJ, 134, 2398
Cox, D. P., & Reynolds, R. J. 1987, ARA&A, 25, 303
Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, RAA, 12, 1197
Curran, P. A. 2014, arXiv:1411.3816
Eenhoorn, M. L., Glenn, J., Evans, N. J. L., et al. 2007, ApJ, 666, 982
Eenhoorn, M. L., Young, K. E., Glenn, J., et al. 2006, ApJ, 643, 293
Espolin, T. L., & Luhman, K. L. 2017, AJ, 154, 134
Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in Proc. of the Second International Conference on Knowledge Discovery and Data Mining, KDD’96, (Portland, OR: AAAI Press), 226, https://studylab.net/doc/13739397/a-density-based-algorithm-for-discovering-clusters
Evans, N. J. L., Dunham, M. M., Jørgensen, J. K., et al. 2009, ApJS, 181, 321
