The ALMA Survey of Star Formation and Evolution in Massive Protoclusters with Blue Profiles (ASSEMBLE): Core Growth, Cluster Contraction, and Primordial Mass Segregation

Fengwei Xu, Ke Wang, Tie Liu, Mengyao Tang, Neal J. Evans, Aina Palau, Kaho Morii, Jinhua He, Patricio Sanhueza, Hong-Li Liu, et al.

The Astrophysical Journal Supplement Series, 270:9 (31pp), 2024 January https://doi.org/10.3847/1538-4365/acfee5

The ALMA Survey of Star Formation and Evolution in Massive Protoclusters with Blue Profiles (ASSEMBLE): Core Growth, Cluster Contraction, and Primordial Mass Segregation

Fengwei Xu[1,2], Ke Wang[1], Tie Liu[3], Mengyao Tang[4], Neal J. Evans II[5], Aina Palau[6], Kaho Morii[7,8], Jinhua He[9,10,11], Patricio Sanhueza[8,12], Hong-Li Liu[13], Amelia Stutz[14], Qizhou Zhang[15], Xi Chen[3,16,17], Pak Shing Li[3], Gilberto C. Gómez[6], Enrique Vázquez-Semadeni[6], Shanghuo Li[18], Xiaofeng Mai[3], Xing Lu[3], Meizhu Liu[13], Li Chen[13], Chuanshou Li[13], Hongqiong Shi[13], Zhiyuan Ren[19], Di Li[19,20,21], Guido Garay[11], Leonardo Bronfman[11], Lokesh Dewangan[22], Mika Juvela[23], Chang Won Lee[24,25], S. Zhang[1], Nannan Yue[1], Chao Wang[1,2], Yifei Ge[1,2], Wenyu Jiao[1,2,18], Qiuyi Luo[3], J.-W. Zhou[26], Ken'ichi Tatematsu[8,12], James O. Chibueze[27,28,29], Keyun Su[1], Shenglan Sun[1], I. Ristorcelli[30], and L. Viktor Toth[31,32]

[1] Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China; kwang.astro@pku.edu.cn
[2] Department of Astronomy, School of Physics, Peking University, Beijing, 100871, People's Republic of China
[3] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, People's Republic of China
[4] Institute of Astrophysics, School of Physics and Electronic Science, Chuxiong Normal University, Chuxiong 675000, People's Republic of China
[5] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA
[6] Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro 8701, Ex-Hda. San José de la Huerta, 58089 Morelia, Michoacán, Mexico
[7] Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[8] National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[9] Yunnan Observatories, Chinese Academy of Sciences, 396 Yangfangwang, Guandu District, Kunming, 650216, People's Republic of China
[10] Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100101, People's Republic of China
[11] Departamento de Astronomía, Universidad de Chile, Las Condes, 7591245 Santiago, Chile
[12] Astronomical Science Program, The Graduate University for Advanced Studies, SOKENDAI, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[13] Department of Astronomy, Yunnan University, Kunming 650091, People's Republic of China
[14] Departamento de Astronomía, Universidad de Concepción, Casilla 160-C, Concepción, Chile
[15] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[16] Center for Astrophysics, Guangzhou University, Guangzhou 510006, People's Republic of China
[17] Peng Cheng Lab, Shenzhen, 518066, People's Republic of China
[18] Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
[19] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
[20] Research Center for Intelligent Computing Platforms, Zhejiang Laboratory, Hangzhou 311100, People's Republic of China
[21] NAOC-UKZN Computational Astrophysics Centre (NUCAC), University of KwaZulu-Natal, Durban, 4000, South Africa
[22] Physical Research Laboratory, Navrangpura, Ahmedabad-380 009, India
[23] Department of Physics, University of Helsinki, PO Box 64, FI-00014 Helsinki, Finland
[24] Korea Astronomy and Space Science Institute, 776 Daedeok-daero, Yuseong-gu, Daejeon 34055, Republic of Korea
[25] University of Science and Technology, 217 Gajeong-ro, Yuseong-gu, Daejeon 34113, Republic of Korea
[26] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany
[27] Department of Mathematical Sciences, University of South Africa, Cnr Christian de Wet Rd and Pioneer Avenue, Florida Park, 1709, Roodepoort, South Africa
[28] Centre for Space Research, Physics Department, North-West University, Potchefstroom 2520, South Africa
[29] Department of Physics and Astronomy, Faculty of Physical Sciences, University of Nigeria, Carver Building, 1 University Road, Nsukka 410001, Nigeria
[30] Univ. Toulouse, CNRS, IRAP, 9 Av. du colonel Roche, BP 44346, F-31028, Toulouse, France
[31] Eötvös University Budapest, Institute of Physics and Astronomy, Pazmany P.s. 1/A, H-1117 Budapest, Hungary
[32] University of Debrecen, Faculty of Science and Technology, Egyetem tér 1, H-4032 Debrecen, Hungary
Received 2023 July 18; revised 2023 September 14; accepted 2023 September 23; published 2023 December 28

Abstract

The Atacama Large Millimeter/submillimeter Array (ALMA) Survey of Star Formation and Evolution in Massive Protoclusters with Blue Profiles (ASSEMBLE) aims to investigate the process of mass assembly and its connection to high-mass star formation theories in protoclusters in a dynamic view. We observed 11 massive ($M_{\rm clump} \gtrsim 10^3\ M_\odot$), luminous ($L_{\rm bol} \gtrsim 10^4\ L_\odot$), and blue-profile (infall signature) clumps by ALMA with resolution of ~2200–5500 au (median value of 3500 au) at 350 GHz (870 μm). We identified 248 dense cores, including 106 cores showing protostellar signatures and 142 prestellar core candidates. Compared to early stage infrared dark clouds (IRDCs) by ASHES, the core mass and surface density within the ASSEMBLE clumps exhibited a significant increment, suggesting concurrent core accretion during the evolution of the clumps. The maximum mass of prestellar cores was found to be 2 times larger than that in IRDCs, indicating that evolved protoclusters have the potential to harbor massive prestellar cores. The mass relation between clumps and their most massive core (MMCs) is observed in ASSEMBLE but not in IRDCs, which is suggested to be regulated by multiscale mass accretion. The mass correlation between the core clusters and their MMCs has a steeper slope compared to that observed in stellar clusters, which can be due to fragmentation of the MMC and stellar multiplicity. We observe a

decrease in core separation and an increase in central concentration as protoclusters evolve. We confirm primordial mass segregation in the ASSEMBLE protoclusters, possibly resulting from gravitational concentration and/or gas accretion.

Unified Astronomy Thesaurus concepts: Protoclusters (1297); Star formation (1569); Massive stars (732); Interstellar medium (847); Protostars (1302); Star forming regions (1565)

Supporting material: machine-readable tables

1. Introduction

Observations suggest that massive stars form either in bound clusters (Lada & Lada 2003; Longmore et al. 2011, 2014; Motte et al. 2018) or in large-scale hierarchically structured associations (Ward & Kruijssen 2018). However, the process of stellar mass assembly, which includes fragmentation and accretion, remains poorly understood. This is a critical step in determining important parameters such as the number of massive stars and their final stellar mass. Also, it is important to note that the fragmentation of molecular gas and core accretion are both time dependent, as the instantaneous physical conditions in the cloud vary during ongoing star formation and feedback (e.g., radiation and outflow).

Over the past decades, researchers have focused on massive clumps associated with infrared dark clouds (IRDCs), which are believed to harbor the earliest stage of massive star and cluster formation (e.g., Rathborne et al. 2006, 2007; Chambers et al. 2009; Zhang et al. 2009; Wang et al. 2011; Zhang & Wang 2011; Sanhueza et al. 2012; Wang et al. 2014; Zhang et al. 2015; Yuan et al. 2017; Pillai et al. 2019; Huang et al. 2023). Despite their large reservoir of molecular gas at high densities $>10^4$ cm^{-3}, IRDC clumps show few signs of star formation. For example, only 12% in a sample of 140 IRDCs has water masers (Wang et al. 2006). Moreover, IRDCs have consistently lower gas temperatures and line widths, with studies in NH_3 finding temperatures of $\lesssim$15 K (Pillai et al. 2006; Ragan et al. 2011; Wang et al. 2012, 2014; Xie et al. 2021) and line widths of $\lesssim$2 km s^{-1} averaged over a spatial scale of 1 pc (Wang et al. 2008; Ragan et al. 2011, 2012). Both two parameters are lower than those observed in high-mass protostellar objects (HMPOs) with a temperature of $\sim$20 K and line widths of $\sim$2 km s^{-1} (Molinari et al. 1996; Sridharan et al. 2002; Wu et al. 2006; Longmore et al. 2007; Urquhart et al. 2011), and those observed in ultracompact H II (UCH II) regions with $>$25 K and $\gtrsim$3 km s^{-1} (Churchwell et al. 1990; Harju et al. 1993; Molinari et al. 1996; Sridharan et al. 2002). Therefore, there is a clear evolutionary sequence from IRDCs to HMPOs and then to UCH II regions, which sets the basis for a time-dependent study of massive star formation.

Taking advantage of the low contamination from stellar feedback in IRDCs, great efforts have been made to investigate the initial conditions of massive star formation therein. For example, Zhang et al. (2009) first conducted arcsecond resolution studies of the IRDC G28.34+0.06 with the Submillimeter Array (SMA) and found that dense cores giving rise to massive stars are much more massive than the thermal Jeans mass of the clump. This discovery challenges the notion in the "competitive accretion" model that massive stars should arise from cores of thermal Jeans mass (Bonnell et al. 2001). The larger core mass in the fragments demands either additional support from turbulence and magnetic fields (Wang et al. 2012) or a continuous accretion onto the core (Vázquez-Semadeni et al. 2023). On the other hand, observations also find that the mass of these cores does not contain sufficient material to form a massive star (Sanhueza et al. 2017, 2019; Morii et al. 2023), and the cores typically continue to fragment when observed at higher angular resolution (Wang et al. 2011, 2014; Zhang et al. 2015; Olguin et al. 2021, 2022), or at slightly later evolutionary stages (e.g., Palau et al. 2015; Beuther et al. 2018). Therefore, the idea of monolithic collapse (McKee & Tan 2003) for massive star formation does not match the observations. On the simulation side, recent work by Pelkonen et al. (2021) has shed light on the inadequacies of both core-collapse and competitive accretion scenarios. Their findings reveal a lack of a direct correlation between the progenitor core mass and the final stellar mass for individual stars, as well as a lack of an increase in accretion rate with core mass.

However, Padoan et al. (2020) suggested a scenario where massive stars are assembled by large-scale, converging, inertial flows that naturally occur in supersonic turbulence. Very recently, He & Ricotti (2023) performed a high resolution up to $\sim$7 au and found that gas should be continuously supplied from larger scales beyond the mass reservoir of the core. Such a continuous mass accretion is observed directly (Dewangan et al. 2022; Redaelli et al. 2022; Xu et al. 2023a) or indirectly (Contreras et al. 2018). More recent observations of IRDCs with the Atacama Large Millimeter/submillimeter Array (ALMA) routinely reach a mass sensitivity far below the thermal Jeans mass and detect a large population of low-mass cores in the clumps that are compatible with the thermal Jeans mass (Sanhueza et al. 2019; Svoboda et al. 2019; Morii et al. 2023). These cores may form low-mass stars in a cluster. To summarize, these observations point to a picture of massive star formation in which dense cores continue to gain material from the parental molecular clump, while the embedded protostar undergoes accretion (see a review in Section 1.1 in Xu et al. 2023a).

Mass assembly is a dynamic process that occurs over time after all, and it is essential to compare the predictions of theoretical models and numerical simulations with observations of massive clumps at a broad range of evolutionary stages to understand high-mass star and cluster formation. While the state-of-the-art understanding of massive star formation suggests gas transfers along filamentary structures to feed the massive dense cores where protostars grow in mass (Gómez & Vázquez-Semadeni 2014; Motte et al. 2018; Naranjo-Romero et al. 2022; Xu et al. 2023a), observational evidence is required to provide more straightforward constraints on the physical processes during protocluster evolution, which will yield a time-tracked understanding of high-mass star and cluster formation.

Therefore, we conduct the ALMA Survey of Star Formation and Evolution in Massive Protoclusters with Blue Profiles (ASSEMBLE), designed to study mass assembly systematically, including fragmentation and accretion, and its connection to high-mass star formation theories. The survey

aims at providing a "dynamic" view from two main perspectives: (1) answering a series of kinematics questions such as when infall starts and stops, how gas transfers inwards, and where infalling gas goes; (2) and unveiling the evolution of key physical parameters in the protoclusters since the sample in the survey provides more evolved protoclusters compared to early stage IRDCs. The first idea is reflected in our sample selection that all the 11 massive clumps are chosen from pilot single-dish surveys with evident blue profiles indicating global infall motions and rapid mass assembly. The sample also benefits from synergy with ALMA Three-millimeter Observations of Massive Star-forming regions (ATOMS; Liu et al. 2020), supporting gas kinematics analyses (Xu et al. 2023a). The second idea is to compare the ASSEMBLE results with those in early stage IRDCs reported by Sanhueza et al. (2019), Morii et al. (2023) as well as Svoboda et al. (2019). Sanhueza et al. (2019), Morii et al. (2023) are both included in series work "The ALMA Survey of 70 μm Dark High-mass Clumps in Early Stages" (ASHES hereafter), which focus on a pilot sample of 12 (ASHES Pilots; Sanhueza et al. 2019) and a total sample of 39 (ASHES Totals; Morii et al. 2023) of carefully chosen IRDCs, respectively. The mean temperature of these IRDCs is ∼15 K, with a range of 9–23 K, and the luminosity-to-mass ratio ranges from 0.1 to 1 $L_{\odot}/M_{\odot}$, supporting the idea that these clumps host the early stages of massive star formation (Morii et al. 2023).

In this paper, we present comprehensive analyses of dust continuum emission from a carefully selected sample comprising 11 massive protocluster clumps that exhibit evidence of gas infall. Our study focuses on investigating the physical properties and evolution of cores within these clumps, including their mass, spatial distribution, and comparison with earlier stages. The paper is structured as follows. Section 2 describes the criteria used for the selection of our sample. Section 3 provides a summary of the observation setups and details the data reduction process. In Section 4, we present the fundamental results derived from the ASSEMBLE data. Section 5 offers in-depth discussions on the implications and significance of the observed results. To gain further insights into protocluster evolution, Section 6 presents comparative analyses with the ASHES data and contributes to the development of a comprehensive understanding of the protocluster evolution. Finally, in Section 7, we summarize the key findings and provide future prospects.

2. Sample Selection

2.1. Massive Clumps with Infall Motion

The ASSEMBLE sample, consisting of 11 carefully selected sources, owes its creation to advanced observational tools such as the IRAS, Spitzer, and Herschel satellites, as well as various ground-based surveys focusing on dust continuum and molecular lines. Bronfman et al. (1996) conducted a comprehensive and homogeneous CS(2–1) line survey of 1427 bright IRAS point sources in the Galactic plane candidates that were suspected to harbor UCH II regions. Subsequently, Faúndez et al. (2004) conducted a follow-up survey of 146 sources suspected of hosting high-mass star formation regions (bright CS(2–1) emission of $T_b > 2$ K, indicative of reasonably dense gas), using 1.2 mm continuum emission. The same set of 146 high-mass star-forming clumps was then surveyed by Liu et al. (2016a), using HCN(4–3) and CS(7–6) lines with the 10 m Atacama Submillimeter Telescope Experiment telescope. With the most reliable tracer of infall motions HCN(4–3) lines (Chira et al. 2014, Xu et al. 2023b), they identified 30 infall candidates based on the "blue profiles."

Out of the 30 infall candidates, 18 are further confirmed by HCN(3–2) and CO(4–3) lines observed with the Atacama Pathfinder Experiment (APEX) 12 m telescope (Yue et al. 2021). Furthermore, the 18 sources were found to have virial parameters below 2, indicating that they are likely undergoing global collapse. All the 18 sources are covered by both the APEX Telescope Large Area Survey of the Galaxy (ATLASGAL; Urquhart et al. 2018) and the Herschel Infrared Galactic Plane Survey (Molinari et al. 2010; Elia et al. 2017, 2021), allowing well-constrained estimates of clump mass and luminosity from the infrared spectral energy distribution (SED) fitting (Urquhart et al. 2018). Given ASSEMBLE's goal of investigating massive and luminous star-forming clumps, the study adopts additional selection criteria that the clump should be massive and luminous.

To summarize, the sample of 11 ASSEMBLE targets meets the following key criteria: (1) the CS(2–1) emission has a brightness temperature of $T_b > 2$ K; (2) the HCN (3–2), HCN(4–3), and CO(4–3) lines exhibit *blue profiles*; (3) the clump masses range from 8×10^2 to 2×10^4 $M_{\odot}$, with a median value of $\sim 4 \times 10^3$ $M_{\odot}$; and (4) the bolometric luminosities range from 1×10^4 to 6×10^5 $L_{\odot}$, with a median value of $\sim 1 \times 10^5$ $L_{\odot}$.

2.2. Physical Properties of Selected Sample

Table 1 presents the basic properties of the ASSEMBLE sample, including the clump kinematic properties, distances, and physical characteristics. The velocity in the local standard of rest (V_{lsr}) was determined from the $C^{17}O$(3–2) lines in the APEX observations (Yue et al. 2021), which is listed in column (5). The line asymmetric parameter ($\delta V \equiv (V_{CO} - V_{C^{17}O})/\Delta V_{C^{17}O}$) in column (6) defines the line as having a blue profile. The kinematic distance as well as its upper and lower uncertainties is estimated using the latest rotation curve model of the Milky Way (Reid et al. 2019) and is listed in column (7). The clump radius is derived from 2D Gaussian fitting and is listed in column (8). The radius is derived from the 2D Gaussian fitting, the same method as adopted in Sanhueza et al. (2019) to better compare with. The dust temperature (T_{dust}), clump mass (M_{cl}), bolometric luminosity (L_{bol}), and luminosity-to-mass ratio (L/M) are obtained from the far-IR (70–870 μm by Herschel and ATLASGAL survey) SED fitting (Table 5 in Urquhart et al. 2018) and are listed in columns (9)–(12), respectively. The clump surface density, $\Sigma_{cl} = M_{cl}/\pi R_{cl}^2$, is listed in column (13). It is noteworthy that all of the ASSEMBLE clumps have a surface density of $\Sigma_{cl} \gtrsim 1$ g cm^{-2}, significantly surpassing the threshold (0.05 g cm^{-2}) for high-mass star formation proposed by Urquhart et al. (2014), He et al. (2015), which further justifies our sample selection.

The background Spitzer three-color composite map (blue, 3.6 μm; green, 4.5 μm; red, 8 μm) in Figure 1 displays the infrared environment. All the 11 targets exhibit bright infrared sources indicating active massive star formation, although the L_{bol}/M_{cl} derived from Table 1 columns (10)–(11) varies from 12 to 80. The differences in the L_{bol}/M_{cl} suggest potential variations in the evolutionary stages among the samples. For instance, I16272-4837 (also known as SDC335; Peretto & Fuller 2009) with the value of 12 is in an early stage of high-mass star formation embedded in a typical IRDC (Xu et al. 2023a). Another example is I15520-5234, where extended

Table 1
Physical Properties of the ASSEMBLE Clumps

ASSEMBLE Clump		Position		V_{lsr}	δV	Dist.	R_{cl}[c]	T_{dust}	log(M_{cl})	log(L_{bol})	L/M	Σ_{cl}	Color ID
IRAS[a]	AGAL[b]	α(J2000)	δ(J2000)	(km s^{-1})		(kpc)	(pc [″])	(K)	($M_\odot$)	($L_\odot$)	$L_\odot/M_\odot$	(g cm^{-2})	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
I14382-6017	G316.139-00.506	14:42:02.76	−60:30:35.1	−60.17	−0.60	$3.62^{+0.68}_{-0.57}$	0.49(28)	28.0	3.65	5.20	35	1.21	
I14498-5856	G318.049+00.086	14:53:42.81	−59:08:56.5	−49.82	−0.66	$2.90^{+0.52}_{-0.47}$	0.23(16)	26.7	3.01	4.43	26	1.29	
I15520-5234	G328.809+00.632	15:55:48.84	−52:43:06.2	−41.67	−0.46	$2.50^{+0.38}_{-0.40}$	0.16(13)	32.2	3.23	5.13	79	4.17	
I15596-5301	G329.406-00.459	16:03:32.29	−53:09:28.1	−74.20	−0.61	$4.21^{+0.41}_{-0.40}$	0.36(17)	28.5	3.93	5.47	34	4.40	
I16060-5146	G330.954-00.182	16:09:52.85	−51:54:54.7	−90.44	−0.56	$5.08^{+0.54}_{-0.44}$	0.30(12)	32.2	3.95	5.78	67	6.50	
I16071-5142	G331.133-00.244	16:11:00.01	−51:50:21.6	−86.80	−0.47	$4.84^{+0.55}_{-0.35}$	0.41(17)	23.9	3.68	4.78	12	1.87	
I16076-5134	G331.279-00.189	16:11:27.12	−51:41:56.9	−88.37	−0.75	$4.99^{+0.45}_{-0.45}$	0.52(21)	30.1	3.57	5.27	49	0.92	
I16272-4837	G335.586-00.291	16:30:59.08	−48:43:53.3	−46.57	−0.40	$2.96^{+0.41}_{-0.30}$	0.22(15)	23.1	3.24	4.29	11	2.34	
I16351-4722	G337.406-00.402	16:38:50.98	−47:27:57.8	−40.21	−0.76	$2.80^{+0.35}_{-0.40}$	0.21(15)	30.4	3.22	4.88	46	2.57	
I17204-3636	G351.041-00.336	17:23:50.32	−36:38:58.1	−17.74	−0.83	$2.76^{+0.60}_{-0.69}$	0.26(19)	25.8	2.88	4.17	19	0.77	
I17220-3609	G351.581-00.352	17:25:24.99	−36:12:45.1	−94.94	−0.77	$7.54^{+0.31}_{-0.18}$	0.79(21)	25.4	4.35	5.66	20	2.38	

Notes. The IRAS and ATLASGAL names of the ASSEMBLE clumps are listed in columns (1)–(2). Field centers of the ALMA mosaic observations are listed in columns (3)–(4), which are not always consistent with the continuum peak but cover most of the emission of the clump. Kinematic properties in columns (5)–(6) are derived from the CO(4–3) and $C^{17}O$(3–2) lines in APEX observations (Yue et al. 2021). The distance and its uncertainties are estimated using the latest rotation curve model of the Milky Way (Reid et al. 2019) and listed in column (7). Clump radius (R_{cl}) is derived from 2D Gaussian fitting, shown in column (8). Clump properties including dust temperature (T_{dust}), mass (M_{cl}), bolometric luminosity (L_{cl}), and luminosity-to-mass ratio (L/M) are retrieved from Urquhart et al. (2018), which are listed in columns (9)–(12). The clump surface density, in column (13), is calculated by $\Sigma_{cl} = M_{cl}/\pi R_{cl}^2$. Identical color, in column (14), is used to distinguish in statistical plots hereafter.

[a] The "IRAS" is replaced by "I" for short.

[b] The ATLASGAL name from Urquhart et al. (2018).

[c] Effective clump radius by Source Extractor (Urquhart et al. 2014).

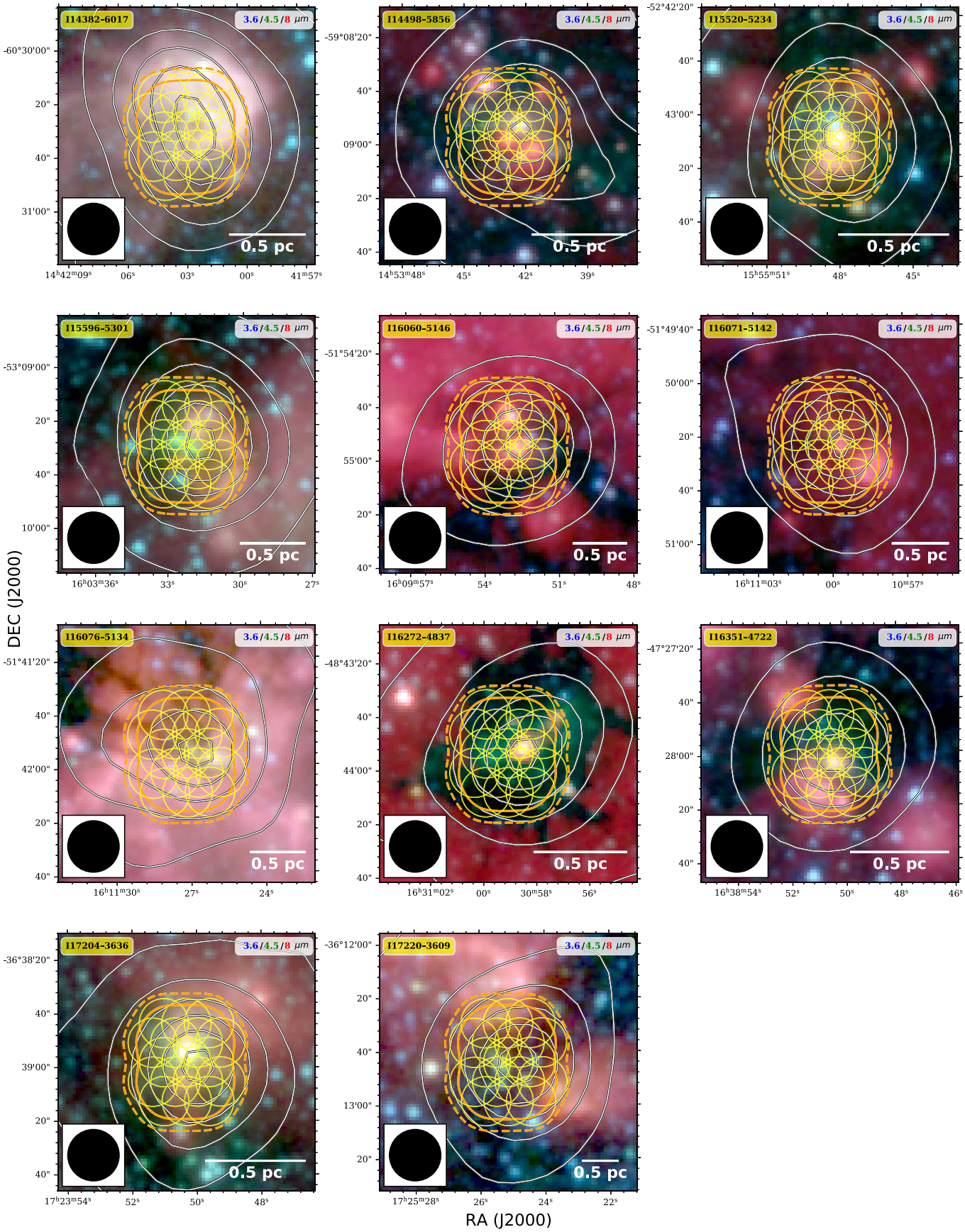

Figure 1. Showcase of the ASSEMBLE sample. Background shows the Spitzer infrared three-color map (blue, 3.6 μm; green, 4.5 μm; red, 8 μm). White contours are ATLASGAL 870 μm continuum emission, with levels starting from 5σ increasing in steps of $f(n) = 3 \times n^p + 2$, where $n = 1, 2, 3, \ldots N$. The beam of the ATLASGAL continuum map is 19.″2 as shown on the bottom left. The ALMA mosaicked pointings are shown with yellow circles, and the primary beam responses of 0.5 and 0.2 are outlined by orange solid and dashed lines respectively. The scale bar of 0.5 pc is labeled in the bottom right corner.

radio ($\nu = 8.64$ GHz) continuum emission indicates evolved UCH II regions (see Figure 4 of Ellingsen et al. 2005). Accordingly, I15520-5234 has the highest value of $L_{\rm bol}/M_{\rm cl} \simeq 80$.

In Appendix A, we present additional information on the radio emission derived from the MeerKAT Galactic Plane Survey 1.28 GHz data (Padmanabh et al. 2023; S. Goedhart et al. 2023, in preparation). All of the protoclusters included in our study exhibit embedded radio emission, which can originate from UCH II regions, or extended radio emission, which may arise from radio jets or extended H II regions. For instance, in the case of I14382-6017, we observe cometary radio emission that exhibits a spatial correlation with the 8 μm emission, outlining the extended shell of the H II region. However, at an early stage, I16272-4837 displays two radio point sources associated with two UCH II regions (Avison et al. 2015). Additionally, I17720-3609 exhibits northward extended radio emission that is linked to a blueshifted outflow (Baug et al. 2020, 2021).

3. Observations and Data Reduction

3.1. ALMA Band 7 Observing Strategy

The 17-pointing mosaicked observations were carried out with ALMA using the 12 m array in Band 7 during Cycle 5 (Project ID: 2017.1.00545.S; PI: Tie Liu) from 2018 May 18 to May 23. The observations have been divided into six executions; 48 antennas were used to obtain a total of 1128 baselines with lengths ranging from 15 to 313.7 m across all the executions.

The mosaicked observing fields of ALMA are designed to cover the densest part of the massive clumps traced by the ATLASGAL 870 μm continuum emission. The fields are outlined by the yellow dashed loops in Figure 1. The field center of each mosaicking field is shown in columns (2)–(3) of Table 1. Each mosaicked field has a uniform size of $\sim$46″ and a sky coverage of $\sim$0.58 arcmin^2. The on-source time per pointing is 2.7–3.7 minutes, which is listed in column (2) of Table 2.

The four spectral windows (SPWs) numbered 31, 29, 25, and 27 are centered at frequencies of 343.2, 345.1, 354.4, and 356.7 GHz, respectively. SPWs 25 and 27 possess a bandwidth of 468.75 MHz and spectral resolution of 0.24 km s^{-1}, which are specifically designed to observe the HCN(4–3) and HCO$^+$(4–3) strong lines, respectively. These lines serve as reliable tracers for infall and outflow (e.g., Chira et al. 2014). On the other hand, SPWs 29 and 31 have a bandwidth of 1875 MHz and a spectral resolution of 0.98 km s^{-1}, which are intended to cover a wide range of spectral lines. The CO(3–2) line serves as outflow tracer according to Sanhueza et al. (2010), Wang et al. (2011), and Baug et al. (2020, 2021). High-density tracers, such as H^{13}CN(4–3) and CS(7–6), can determine the core velocity. Additionally, some sulfur-bearing molecules, such as H_2CS and SO_2, can serve as tracers of rotational envelopes, while shock tracers include SO $^3\Sigma$ (8_8–7_7) and $CH_3OH(13_{1,12}$–$13_{0,13})$. The hot-core molecular lines, such as H_2CS, CH_3OCHO, and CH_3COCH_3, have a sufficient number of transitions to facilitate rotation-temperature and chemical abundance studies. A summary of the target spectral lines can be found in Table A1 of Xu et al. (2023a).

3.2. ALMA Data Calibration and Imaging

The pipeline provided by the ALMA observatory was utilized to perform data calibration in CASA (McMullin et al. 2007) version 5.1.15. The phase, flux, and bandpass calibrators are listed in columns (7)–(8) of Table 2. The imaging was conducted through the `TCLEAN` task in CASA 5.3. To aggregate the continuum emission, the line-free channels were meticulously selected by visual inspection, with the bandwidth and its percentage of total bandwidth listed in column (3). A total of three rounds of phase self-calibration and one round of amplitude self-calibration were run to enhance the dynamic range of the image. For self-calibration, antenna DA47 was designated as the reference antenna. During imaging, the deconvolution was set as "hogbom" while the weighting parameter was set as "briggs" with a robust value of 0.5 to balance sensitivity and angular resolution. The primary beam correction is conducted with `pblimit=0.2`. Following self-calibration, the sensitivity and dynamic range of the final continuum image were significantly improved, as indicated in column (4), ranging from 0.5 to 1.7 mJy beam^{-1} with a mean value of $\sim$1 mJy beam^{-1}. The beam size (i.e., angular resolution) with 0.″8–1.″2 and maximum recoverable scale (MRS) with 7.″2–9.″2 are presented in columns (5)–(6).

4. Results

4.1. Dust Continuum Emission

Figure 2 presents the ALMA 870 μm dust continuum images without primary beam correction for a uniform rms noise. As a comparison, the dust continuum emission at the same wavelength from the single-dish survey ATLASGAL (Schuller et al. 2009) with a beam size of 19.″2 is overlaid as black contours. In all of the 11 targets, the small-scale structures resolved by ALMA show a good spatial correlation with the large-scale structures seen by ATLASGAL. In other words, the dense structures surviving in the interferometric *filtering-out* effect are mostly distributed in the densest part of the clump. However, the small-scale structures present various morphologies: some present elongated filaments (e.g., I14382-6017 and I16071-5142); some have centrally concentrated cores (e.g., I16272-4837); some have spiral-like dust arms (e.g., I15596-5301 and I16060-5146).

4.2. Core Extraction and Catalog

Here, we present the extraction of core-like structures (or cores) and the measurement of fundamental physical parameters including integrated flux, peak intensity, size, and position. The choice of core extraction algorithm should be carefully made based on actual physical scenarios and scientific expectations. In this work, we use the `getsf` extraction algorithm that spatially decomposes the observed images to separate relatively round sources from elongated filaments as well as their background emission (Men'shchikov 2021). As suggested by Xu et al. (2023a), the `getsf` algorithm is a better choice than `astrodendro` in the case study of SDC335 (one of the ASSEMBLE sample), because it can (1) deal with uneven background and rms noise; (2) can separate the blended sources and/or filaments; (3) extract extended emission features.

We perform the `getsf` algorithm on the continuum emission maps without primary beam correction (`unpbcor`).

Table 2
ALMA Observational Parameters

ASSEMBLE Clump	On-source Time (min/pointing)	Line-free Bandwidth[a] (MHz [%])	rms Noise (mJy beam^{-1})	Beam Sizes ($'' \times ''$)	MRS[b] ($''$)	Phase Calibrators	Flux and Bandpass Calibrators
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I14382-6017	3.2	3569 (76.1)	1.0	1.175 × 0.793	8.433	J1524-5903	J1427-4206
I14498-5856	3.2	1407 (30.0)	1.0	1.065 × 0.785	8.179	J1524-5903	J1427-4206
I15520-5234	3.7	1249 (26.6)	1.5	0.852 × 0.699	9.197	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I15596-5301	3.7	1347 (28.7)	0.5	0.824 × 0.691	9.071	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I16060-5146	3.7	58 (1.2)	2.5	0.820 × 0.681	8.953	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I16071-5142	3.7	70 (1.5)	1.3	0.810 × 0.666	8.974	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I16076-5134	3.7	354 (7.5)	0.6	0.734 × 0.628	8.755	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I16272-4837	3.7	72 (1.5)	1.4	0.799 × 0.662	8.450	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I16351-4722	3.7	131 (2.6)	1.2	0.771 × 0.651	8.221	J1650-5044	J1924-2914, J1427-4206, J1517-2422
I17204-3636	2.7	2432 (51.9)	0.5	0.790 × 0.660	7.246	J1733-3722	J1924-2914
I17220-3609	2.7	320 (6.8)	1.7	0.813 × 0.653	7.237	J1733-3722	J1924-2914

Notes. The ASSEMBLE target names are listed in column (1).
[a] The bandwidth used for continuum subtraction as well as the percentage to total bandwidth.
[b] Maximum recoverable scale.

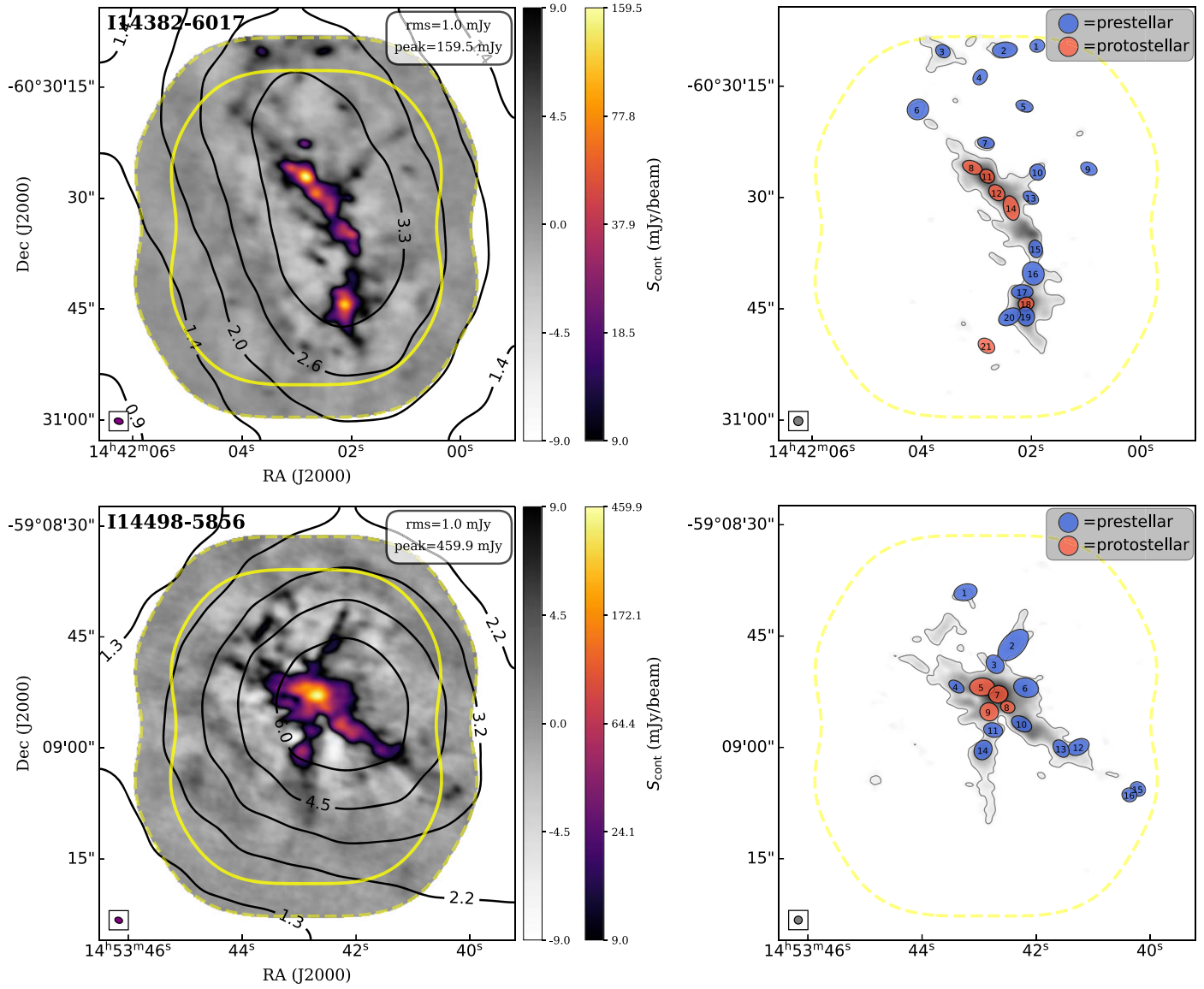

Figure 2. The ALMA 870 μm dust continuum emission without primary beam correction as well as extracted cores for two ASSEMBLE clumps (I14382-6017 and I14498-5856). The ALMA mosaicked primary beam responses of 0.5 and 0.2 are outlined by yellow solid and dashed lines respectively. Only the primary beam response of 0.2 is shown on the right panel. The beam size of each continuum image is shown in the bottom left corner. Left: the background color map shows the ALMA 870 μm emission with two colorbars, the first one (gray scale) showing -9 to $+9$ times the rms noise on a linear scale, then a second one (color-scheme) showing the range $+9$ times the rms noise to the peak value of the image in an arcsinh stretch. The rms noise and peak intensity are given on the top right. The black contours are from the ATLASGAL 870 μm continuum emission, with power-law levels that start at 5σ and end at $I_{\rm peak}$, increasing in steps following the power law $f(n) = 3 \times n^p + 2$ where $n = 1, 2, 3, \ldots N$, and p is determined from $D = 3 \times N^p + 2$ ($D = I_{\rm peak}/\sigma$, the dynamic range; $N = 8$, the number of contour levels). The values of each contour level are labeled with a unit of Jy beam^{-1}. Right: the background gray-scale map shows the arcsinh-stretch part in the left panel, outlined by the 5σ contour. The ALMA continuum emission map is smoothed to a circular beam with a size equal to the major axis of the original beam. The cores extracted by `getsf` algorithm are presented by red and/or blue ellipses, as well as black IDs, with numbers in order from north to south. The red and blue ones represent protostellar and prestellar cores defined in Section 4.3.

The `unpbcor` map is first smoothed into one with a circular beam whose size is equal to the major axis of the original beam. The `getsf` is set to extract sources whose sizes should be larger than the beam size but smaller than the MRS. As suggested by Men'shchikov (2021), significantly detected sources are defined as follows: (1) signal-to-noise ratio larger than unity; (2) peak intensity at least 5 times larger than the local intensity noise; (3) total flux density at least 5 times larger than the local flux noise; (4) ellipticity not larger than 2 to ensure a core-like structure; (5) footprint-to-major-axis ratio larger than 1.15 to rule out cores with abrupt boundary emission. After core extraction as well as fundamental measurement by `getsf`, two flux-related parameters (integrated flux and peak intensity) are corrected by the primary beam response, depending on the core location in the continuum emission maps with primary beam correction (`pbcor`). Fundamental measurements of the core parameters are listed in Table 3.

To evaluate how much flux is recovered by the ALMA observations, we integrate the ATLASGAL 870 μm flux over the field of view of the ASSEMBLE clumps. If all the sources and filaments extracted by `getsf` are included, then the recovered flux by ALMA ranges from 10% to 25%. Although the flux recovery can be further improved by including short-baseline observations (e.g., the Atacama Compact Array), some SMA/ALMA observations show a typical flux recovery between 10%

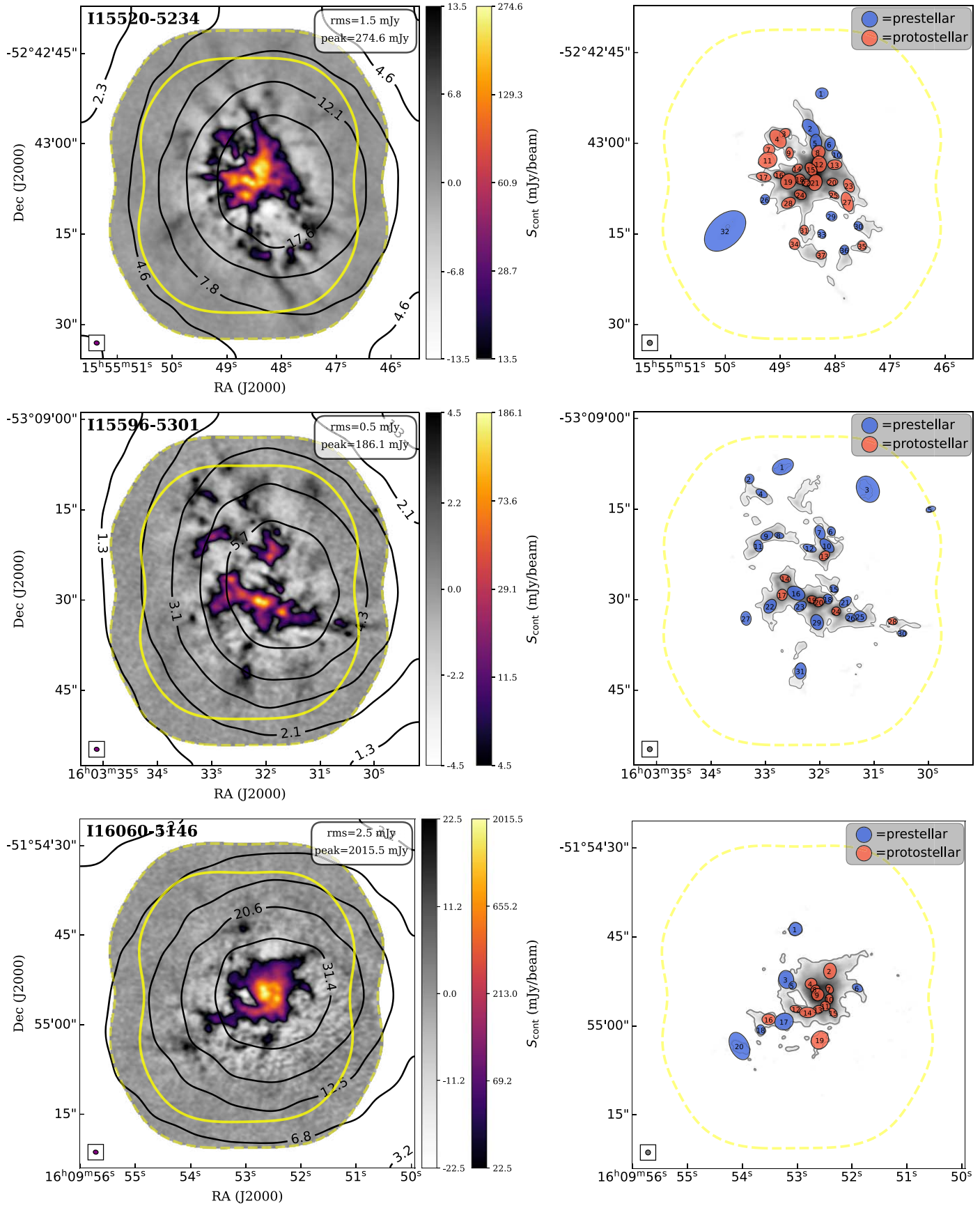

Figure 2. (Continued.)

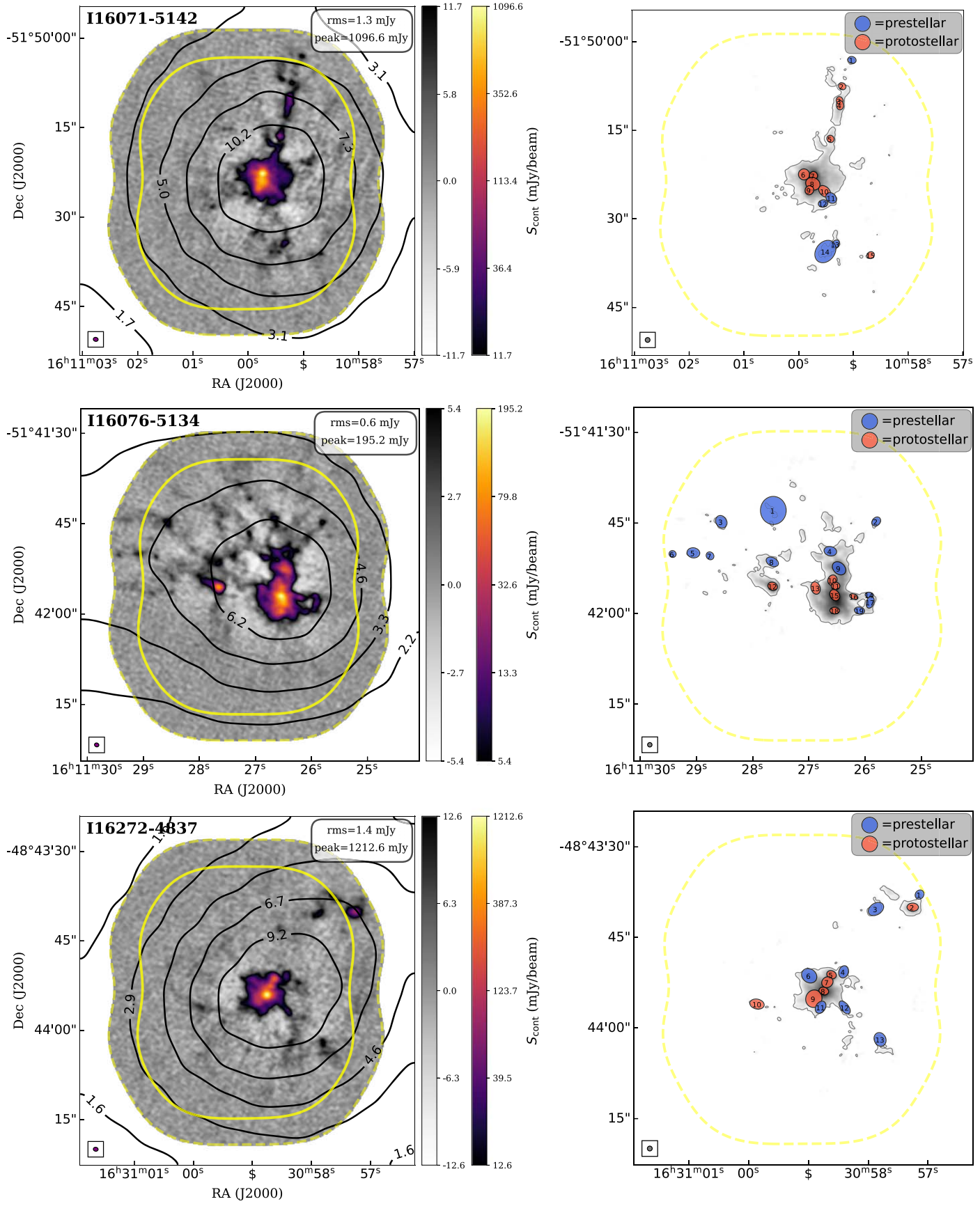

Figure 2. (Continued.)

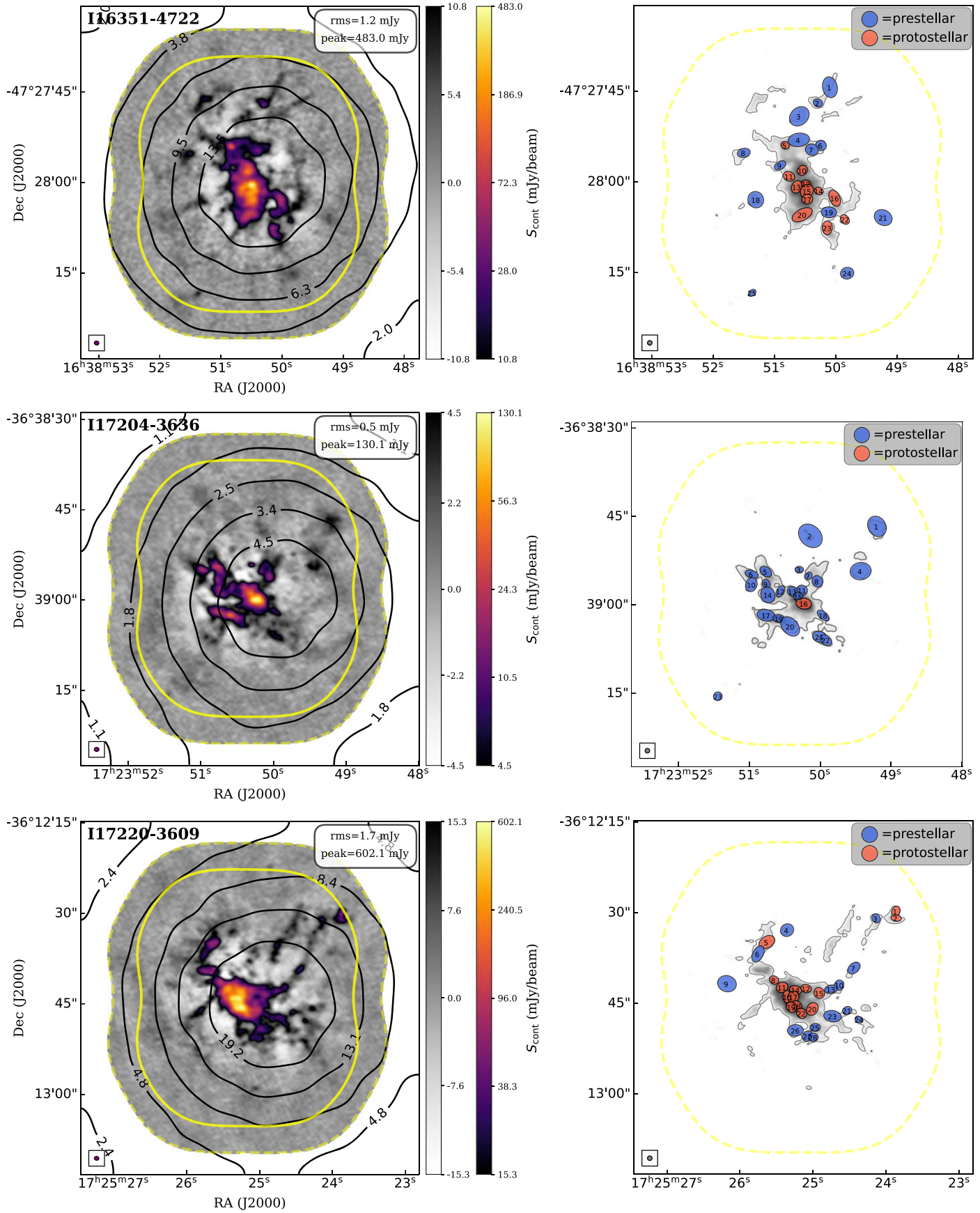

Figure 2. (Continued.)

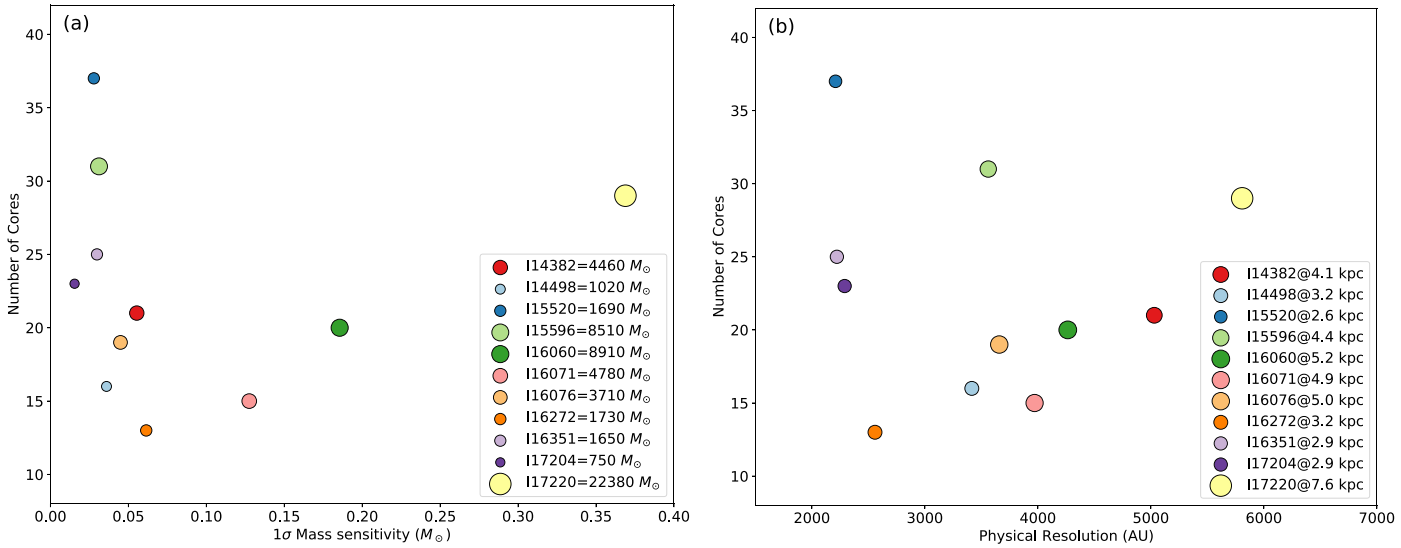

Figure 3. (a) Number of cores detected against the 1σ mass sensitivity. The mass of clump is coded as the size of circle. (b) Number of cores detected against the physical resolution. The distance of clump is coded as the size of circle.

Table 3
Fundamental Measurements of Core Parameters from `getsf`

ASSEMBLE Clump	Core Name	Position		Peak Intensity	Integrated Flux	$\theta_{maj} \times \theta_{min}$	PA	θ_{deconv}	Core Classification[a]
		α(J2000)	δ(J2000)	(mJy beam^{-1})	(mJy)	($''$ × $''$)	(deg)	($''$)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
I14382	1	14:42:01.91	−60:30:09.7	25.1(2.2)	29.8(1.7)	1.33 × 1.13	100.7	1.23	0
I14382	2	14:42:02.50	−60:30:10.3	38.9(5.6)	72.3(5.7)	2.24 × 1.40	93.7	1.3	0
I14382	3	14:42:03.63	−60:30:10.4	51.2(7.2)	51.2(5.6)	1.35 × 1.11	58.7	1.23	0
I14382	4	14:42:02.95	−60:30:13.9	12.2(3.8)	13.2(2.9)	1.47 × 1.18	142.7	1.23	0
I14382	5	14:42:02.15	−60:30:17.8	10.3(1.6)	10.8(1.2)	1.58 × 1.01	72.7	1.23	0

Notes. The ASSEMBLE clump and extracted core ID are listed in columns (1) and (2). The core IDs are in order from the north to the south. The equatorial coordinate centers of the cores are listed in columns (3)–(4). The peak intensity and integrated flux are listed in columns (5)–(6). The fitted FWHM of the major and minor axes convolved with the beam and the position angle (counterclockwise from the north) are listed in columns (7)–(8). The deconvolved FWHM of the core size is shown in column (9). The core classification in column (10) is based on Section 4.3.

[a] Core classification: 0 = prestellar candidate, 1 = only molecular outflow is detected, 2 = only warm-core line is detected, 3 = both outflow and warm-core line are detected, 4 = both outflow and hot-core line are detected, and 5 = only hot-core line is detected.

(This table is available in its entirety in machine-readable form.)

to 30% (e.g., Wang et al. 2014; Sanhueza et al. 2017; Liu et al. 2018; Sanhueza et al. 2019). In our case, the maximum recoverable scale is $\sim 9''$. Therefore, most of the mass in the massive clump is not confined in dense structures (cores).

As shown in Figure 3(a), we found no overall correlation between the number of detected cores and the mass sensitivities with a Pearson correlation coefficient of 0.08. Likewise, panel (b) reveals no overall correlation between the number of detected cores and the physical resolution with a Pearson correlation coefficient of −0.04. Therefore, the number of detected cores is basically independent of the mass sensitivity and spatial resolution provided by the observations.

4.3. Core Classification and Evolutionary Stages

All the ASSEMBLE clumps have infrared bright signatures. As an example of a relatively early stage, I16272-4837 has extended 4.5 μm emission, which is a common feature of outflows (Cyganowski et al. 2008). A more evolved example of I14382-6017 is totally immersed in a cometary H II region traced by the polycyclic aromatic hydrocarbon emission in the 8 μm emission. Therefore, at least some cores in each clump are in an active star formation stage.

The classification of the evolutionary stages of the 248 cores is based on the identification of star formation indicators, including molecular outflows, H_2CS multiple transition lines, and CH_3OCHO multiple transition lines.

For molecular outflows, Baug et al. (2020) used CO(3–2), HCN(4–3), and SiO(2–1) emission lines to confirm the presence of 32 bipolar and 41 unipolar outflows in the 11 ASSEMBLE clumps, and then, a total of 42 continuum cores are associated with outflows. In this study, we updated the outflow catalogs by a channel-by-channel analysis of the outflow lobes to determine their association with the extracted cores, and subsequently assigned the outflows accordingly. A total of 39 (∼16%) cores are assigned bipolar or unipolar outflows. If a core is assigned outflows, then it is classified as protostellar (Nony et al. 2023). Some cores even show multipolar outflows (e.g., I16272-4837 ALMA8; Olguin et al. 2021), indicating either the precession of accreting and outflowing protostars or the presence of multiple outflows from multiple systems. However, we should acknowledge that the method will miss those weak outflows associated

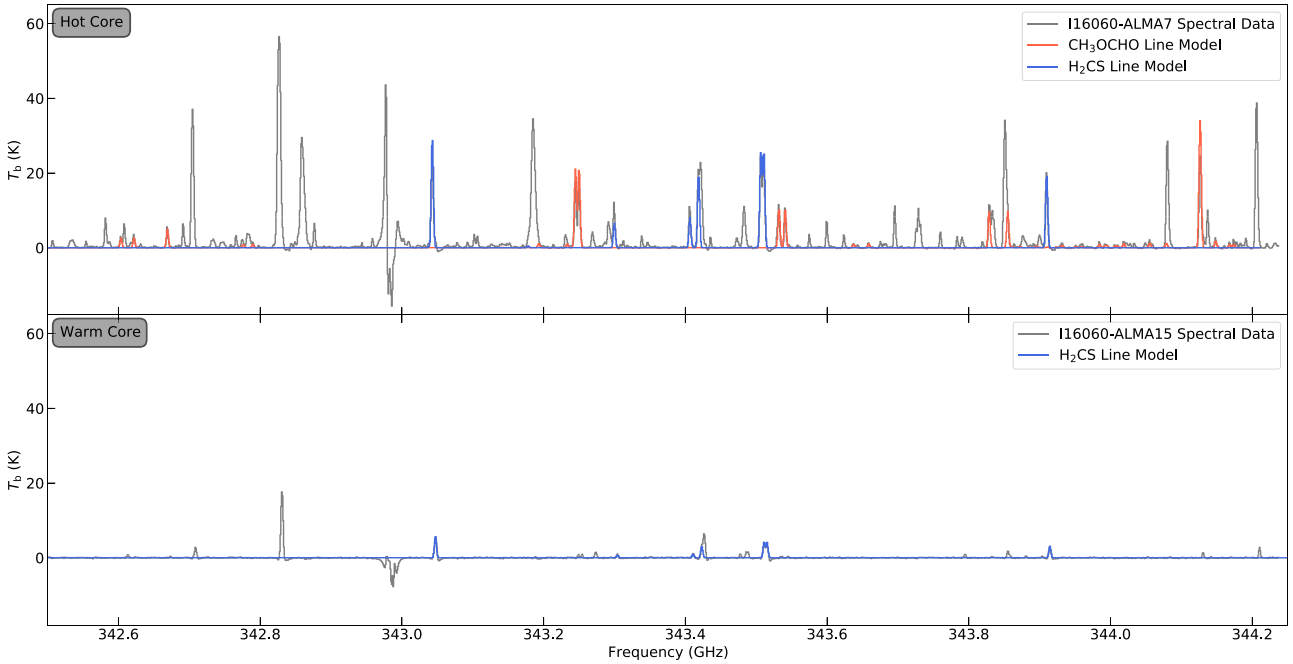

Figure 4. Examples of *hot-core* and *warm-core* spectra. The gray lines are the real spectral data extracted from I16060-ALMA7 and I16060-ALMA15 dense cores. The best-fit line models of CH_3OCHO and H_2CS are shown in red and blue, respectively. The core temperatures are assumed to be 112 and 89 K for the hot and warm cores, respectively.

with the lowest-mass objects, especially for the more distant regions, naturally yielding a lower limit in the number of protostellar objects.

Owing to its comparatively abundant nature, the emission of H_2CS is observed extensively in the core population. However, the relatively low abundance of CH_3OCHO species restricts its detection to hot molecular cores with line-rich features. In this paper, we first classify those cores with robust ($>3\sigma$) detections of both CH_3OCHO and H_2CS multiple transition lines as *hot cores*, especially those that have robust rotation temperature estimation by both CH_3OCHO and H_2CS molecules. Since the "hot cores" are believed to be the result of warm-up processes by central protostar(s) to 100–300 K (Gieser et al. 2019), there should be a stage of dense cores with temperature of <100 K and without line-rich features, which are called "warm cores" (Sanhueza et al. 2019). Then, we define cores with only robust detection of H_2CS but without detection of CH_3OCHO lines as warm cores. We present examples of both *hot cores* and *warm cores* in Figure 4, where I16060-ALMA7 is a typical hot core with line-rich feature, and evident detection of multiple transitions of CH_3OCHO, as well as H_2CS. However, I16060-ALMA15 has a paucity of hot molecular lines, including CH_3OCHO, but with only H_2CS.

Among the 248 ASSEMBLE cores, H_2CS line emissions have been identified in 92 cores, of which 35 display *line-rich* features and are further categorized as hot cores, while the other 57 cores are classified as *warm cores* based on the detection of enough H_2CS lines. Among these warm cores, 22 have insufficient H_2CS transitions available for the calculation of temperature. 142 cores without the star-forming indicators mentioned above (outflows, H_2CS, or CH_3OCHO lines) are then classified as a prestellar core candidate, implying a stage preceding the protostellar phase. Based on the classification above, we mark the core in the column (10) of Table 3: 0 = prestellar candidate, 1 = only molecular outflow is detected, 2 = only H_2CS line is detected, 3 = both outflow and H_2CS line are detected, and 4 = both CH_3OCHO line and outflow are detected, and 5 = only CH_3OCHO line is detected.

The caveats of the core classification results are as follows: (1) external heating by hot cores in the vicinity can also excite H_2CS lines in some prestellar cores, so some warm cores can have no stars form inside; (2) prestellar core candidates may include both prestellar and protostellar cores that are gravitational bound, and cores that are not bound and unable to form a star. To keep consistent, we do not distinguish the two and refer to them as prestellar core candidates in the following part of the paper. We note that spectral analyses of these cores can further constrain their dynamic states.

It is noteworthy that outflows have been observed in all of the ASSEMBLE clumps, providing evidence of star-forming activities with a 100% occurrence rate in our clump sample. However, two massive star-forming clumps, namely I14382-6017 and I17204-3636, do not exhibit any detection of hot cores. This absence of hot cores has been confirmed by crossmatching with the ALMA Band 3 data set, ensuring their nonexistence (Qin et al. 2022). In the case of the protocluster I14382-6017, the extended spherical morphology of H40α line emission is spatially consistent with the MeerKAT Galactic Plane Survey 1.28 GHz data (Padmanabh et al. 2023; S. Goedhart et al. 2023, in preparation). As identified by Zhang et al. (2023), it represents an UCH II region with an electron density of $0.15–0.16 \times 10^4$ cm^{-3}. The protocluster I14382-6017 is situated on the outskirts of the UCH II region, suggesting the possibility of a second generation of cores (refer to Figure 14). As a result, the absence of hot cores in this particular region can be attributed to the relatively young age of the newly formed protocluster. The absence of hot cores in I17204-3636 can be a different issue, as the H40α and the 1.28 GHz emission are spatially correlated with dense cores (see Figure 14). But we note that I17204-3636 has the lowest mass of 760 $M_\odot$, with the maximum core mass of 2.9 $M_\odot$ (refer to Section 4.4). Furthermore, the temperature of the only warm

Table 4
Calculated Properties for the Core Sample

ASSEMBLE Clump	Core Name	T_{dust}		M_{core}	R_{core}	$n(H_2)$	Σ	N_{peak} (H_2)
		(K)	Method[a]	($M_\odot$)	(au)	($\times 10^6$ cm^{-3})	(g cm^{-2})	($\times 10^{23}$ cm^{-2})
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
I14382	1	28(5)	G	1.3(0.9)	2200	3.71(2.43)	0.44(0.16)	0.94(0.34)
I14382	2	28(5)	G	3.2(2.1)	4700	0.93(0.62)	0.69(0.28)	1.46(0.60)
I14382	3	28(5)	G	2.3(1.6)	2200	6.38(4.37)	0.90(0.37)	1.92(0.78)
I14382	4	28(5)	G	0.6(0.4)	2200	1.64(1.24)	0.21(0.12)	0.46(0.25)
I14382	5	28(5)	G	0.5(0.3)	2200	1.35(0.92)	0.18(0.08)	0.39(0.16)

Notes. The ASSEMBLE clump and extracted core ID are listed in columns (1) and (2). Dust temperature and its estimation methods are listed in columns (3) and (4). The mass, radius, volume density, surface density, and peak column density are listed in columns (5)–(9).
[a] Temperature estimation method: G = global clump-averaged temperature in column (9) of Table 1; H = H_2CS rotation temperature; C = CH_3OCHO rotation temperature.

(This table is available in its entirety in machine-readable form.)

core I17204-ALMA16 is 88(±7) K (Section 4.4), which is not so high as $\gtrsim$100 K to be a hot core. Therefore, in the case of I17204-3636, the cores may not be massive and hot enough to excite hot molecular lines or initiate hot-core chemistry.

4.4. Core Physical Properties

The temperature estimation utilizes three hybrid methods (clump-averaged temperature, H_2CS line, and CH_3OCHO line) based on the core properties. H_2CS lines are chosen due to their strong spatial correlation with dust as demonstrated in Xu et al. (2023a), and their widespread distribution. The ASSEMBLE spectral window encompasses multiple hyperfine components of the $J = 10 - 9$ transitions, with upper energy levels from 90 to 420 K (see Table C 1 in Xu et al. 2023). However, H_2CS lines could be optically thick toward massive hot cores, therefore only tracing the core envelope. To trace the dust temperature of hot cores, CH_3OCHO molecule with upper energy up to ∼589 K is employed instead. The temperatures obtained from CH_3OCHO (mean value of 110 K) are consistently higher than those derived from H_2CS (mean value of 95 K), indicating that CH_3OCHO is a suitable tracer of the inner and denser gas. In the cases where neither H_2CS nor CH_3OCHO lines are detected, it is assumed that the core either lacks sufficient column density or is too cold to excite the lines. This suggests that the core has not developed its own temperature gradient and thus is assumed to share the same temperature as the clump from the SED fitting. The temperature as well as the method to obtain it are listed in columns (3)–(4) of Table 4.

Assuming that all the emission comes from dust in a single T_{dust} and that the dust emission is optically thin, the core masses are then calculated using

$$M_{core} = \mathcal{R}\frac{F_\nu^{int} D^2}{\kappa_\nu B_\nu(T_{dust})}, \quad (1)$$

where F_ν^{int} is the measured integrated dust emission flux of the core, $\mathcal{R}$ is the gas-to-dust mass ratio (assumed to be 100), D is the distance, κ_ν is the dust opacity per gram of dust, and $B_\nu(T_{dust})$ is the Planck function at a given dust temperature T_{dust}. In our case, κ_ν is assumed to be 1.89 cm^2 g^{-1} at $\nu \sim 350$ GHz (Xu et al. 2023a), which is interpolated from the given table in Ossenkopf & Henning (1994), assuming grains with thin ice mantles and the Mathis–Rumpl–Nordsieck (Mathis et al. 1977) size distribution and a gas density of 10^6 cm^{-3}. Substituting the temperature in Equation (1), the core masses are then calculated and listed in the column (5).

Cores are characterized by 2D Gaussian-like ellipses with the FWHM of the major and minor axes (θ_{maj} and θ_{min}), and position angle (PA) listed in columns (7)–(8) of Table 3. Following Rosolowsky et al. (2010), Contreras et al. (2013), the angular radius can be calculated as the geometric mean of the deconvolved major and minor axes:

$$\theta_{core} = \eta[(\sigma_{maj}^2 - \sigma_{bm}^2)(\sigma_{min}^2 - \sigma_{bm}^2)]^{1/4}, \quad (2)$$

where σ_{maj} and σ_{min} are calculated from $\theta_{maj}/\sqrt{8\ln 2}$ and $\theta_{min}/\sqrt{8\ln 2}$ respectively. The σ_{bm} is the averaged dispersion size of the beam (i.e., $\sqrt{\theta_{bmaj}\theta_{bmin}/(8\ln 2)}$ where θ_{bmj} and θ_{bmin} are the FWHM of the major and minor axis of the beam). η is a factor that relates the dispersion size of the emission distribution to the angular radius of the object determined. We have elected to use a value of $\eta = 2.4$, which is the median value derived for a range of models consisting of a spherical, emissivity distribution (Rosolowsky et al. 2010). Therefore, the core physical radius can be directly calculated by $R_{core} = \theta_{core} \times D$, as shown in the column (6) of Table 4.

The number density, n, is then calculated by assuming a spherical core,

$$n = \frac{M_{core}}{(4/3)\pi\mu_{H_2} m_H R_{core}^3}, \quad (3)$$

where μ_{H_2} is the molecular weight per hydrogen molecule, and m_H is the mass of a hydrogen atom. Throughout the paper, we adopt the molecular weight per hydrogen molecule $\mu_{H_2} = 2.81$ (Evans et al. 2022), and derive the number density of hydrogen molecule $n(H_2)$.

The core-averaged surface density can be calculated by $\Sigma = M_{core}/(\pi R_{core}^2)$. The peak column density is estimated from

$$N_{peak}(H_2) = \mathcal{R}\frac{F_\nu^{peak}}{\Omega\mu_{H_2} m_H \kappa_\nu B_\nu(T_{dust})}, \quad (4)$$

where F_ν^{peak} is the measured peak flux of core within the beam solid angle Ω.[33] The calculated volume, surface, and peak column densities are shown in columns (7)–(9) of Table 4.

The major sources of uncertainty in the mass calculation come from the gas-to-dust ratio and the dust opacity. We adopt

[33] Beam solid angle: $\Omega = \frac{\pi\theta_{maj}\theta_{min}}{4\ln(2)}$.

Table 5
Statistics of the ALMA Cores in Each Clump

ASSEMBLE Clump	1σ Mass Sensitivity	Number of Cores	Core Mass		Mean Value of					Number of Prestellar/ Protostellar Cores
			Min.	Max.	Mass	Radius	$n(H_2)$ ($\times 10^6$ cm^{-3})	Σ (g cm^{-2})	N_{peak} (H_2) ($\times 10^{23}$ cm^{-2})	
	($M_\odot$)		($M_\odot$)	($M_\odot$)	($M_\odot$)	(au)				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
I14382	0.044	21	0.39	4.65	1.73	2760	4.1	0.5	1.1	15/6
I14498	0.030	16	0.37	9.70	2.25	3210	3.9	0.9	1.9	12/4
I15520	0.026	37	0.12	3.75	0.99	2250	8.4	0.8	1.7	11/26
I15596	0.029	31	0.46	8.14	2.14	3210	7.1	0.7	1.5	24/7
I16060	0.182	20	1.67	52.61	14.34	4800	22.2	3.2	6.9	7/13
I16071	0.128	15	1.57	49.03	8.82	3420	17.8	2.5	5.4	5/10
I16076	0.046	19	0.56	12.07	2.41	3910	3.6	0.7	1.4	12/7
I16272	0.054	13	0.17	19.28	4.58	2730	40.6	2.7	5.7	7/6
I16351	0.029	25	0.18	2.76	1.28	2880	6.9	0.9	2.0	13/12
I17204	0.014	23	0.22	2.89	1.05	2900	7.4	0.7	1.5	22/1
I17220	0.372	28	5.37	52.27	19.57	6190	7.9	2.3	4.8	14/14

Note. The ASSEMBLE clump is listed in column (1). The 1σ mass sensitivity and the number of extracted cores are listed in columns (2) and (3). The minimum, maximum, and mean values of the mass are listed in columns (4)–(6). The mean values of the core radius, volume density, surface density, and peak column density are listed in columns (7)–(10). The numbers of prestellar and protostellar cores are listed in column (11).

the uncertainties derived by Sanhueza et al. (2017) of 28% for the gas-to-dust ratio and of 23% for the dust opacity, contributing to the ∼36% uncertainty of the specific dust opacity. The uncertainties of the core flux (∼14%), temperature (∼20%), and distance (assumed to be 10%) are included. Monte Carlo methods are adopted for an uncertainty estimation, and 1σ confidence intervals are given for core mass, volume density, surface density, and peak column density in columns (5) and (7)–(9) of Table 4.

We also summarize the statistics of the core physical parameters in Table 5. The number of cores in each clump is listed in column (3). The minimum, maximum, and mean core mass are listed in columns (4)–(6). The mean values of the core radius, volume density, surface density, and column density are listed in columns (7)–(10). The numbers of prestellar and protostellar cores are listed in column (11).

5. Discussion

5.1. Coevolution of Clump and Most Massive Core

The ability of a clump to form massive stars is directly linked to the amount of material within the natal clump (Beuther et al. 2013). Furthermore, Traficante et al. (2023) show that physical properties such as mass and surface density of the fragments and their parent clumps are tightly correlated. Therefore, it is essential and straightforward to study the relation between clump and its most massive core (MMC), which is most likely to form massive stars inside the clump. The left panel of Figure 5 shows the core masses (M_{core}) versus the mass of the clump (M_{clump}) of the ASSEMBLE clumps, with the maximum value, that is, the mass of MMC (M_{max}) labeled. As demonstrated in the right panel, a positive sublinear correlation is observed between $M_{ASSEMBLE,max}$ and $M_{ASSEMBLE,clump}$, with a power-law index of 0.75(0.08). The Pearson and Spearman correlation coefficients are calculated to be 0.67 and 0.73, respectively. Significantly, both correlation coefficients exhibit p-values below 0.05, indicating a high level of statistical significance for the observed correlation. This positive correlation indicates a coevolution between the clump and MMC, i.e., a more massive clump contains a more massive core, which is consistent with what has been found in Anderson et al. (2021).

Furthermore, the coevolution of the massive clump and its MMC can be connected to gas kinematics in a dynamic picture. In massive star-forming regions, filamentary gas accretion flows frequently connect clump and core scales in both observations (Peretto et al. 2013, 2014; Liu et al. 2016b; Lu et al. 2018; Yuan et al. 2018; Dewangan et al. 2020; Sanhueza et al. 2021; Li et al. 2022; Xu et al. 2023a; Yang et al. 2023) and simulations (Schneider et al. 2010; Naranjo-Romero et al. 2022), which can play a crucial role in regulating mass reservoirs at different scales. Notably, Xu et al. (2023a) found four spiral-like gas streams conveying gas from the natal clump directly to the MMC, with a continuous and steady gas accretion rate across 3 mag. Therefore, we suggest that such a "conveyor belt" (Longmore et al. 2014) should be the main reason for coevolution. If all the massive clumps are undergoing a quick mass assembly, the sublinearity of the mass scaling relation also suggests that the clump-to-cores efficiency should vary among different clumps (Xu et al. 2023b). To more directly understand the dynamic picture of the coevolution of clump and core, detailed gas kinematics analyses should be systematically performed in a sample with a wide range of evolutionary stages.

5.2. High-mass Prestellar Cores in Protoclusters?

High-mass prestellar cores, defined as cores with masses greater than 30 $M_\odot$ (following the definition of Sanhueza et al. 2019), are crucial in discriminating between different models of high-mass star formation. Specifically, they provide a key discriminator for the turbulent core accretion model (McKee & Tan 2003; Tan et al. 2013, 2014) versus the competitive accretion model (Bonnell et al. 2001, 2004) or the global hierarchical collapse model (Heitsch et al. 2008; Vázquez-Semadeni et al. 2009; Ballesteros-Paredes et al. 2011; Vázquez-Semadeni et al. 2017; Ballesteros-Paredes et al. 2018). Despite numerous observational searches for high-mass

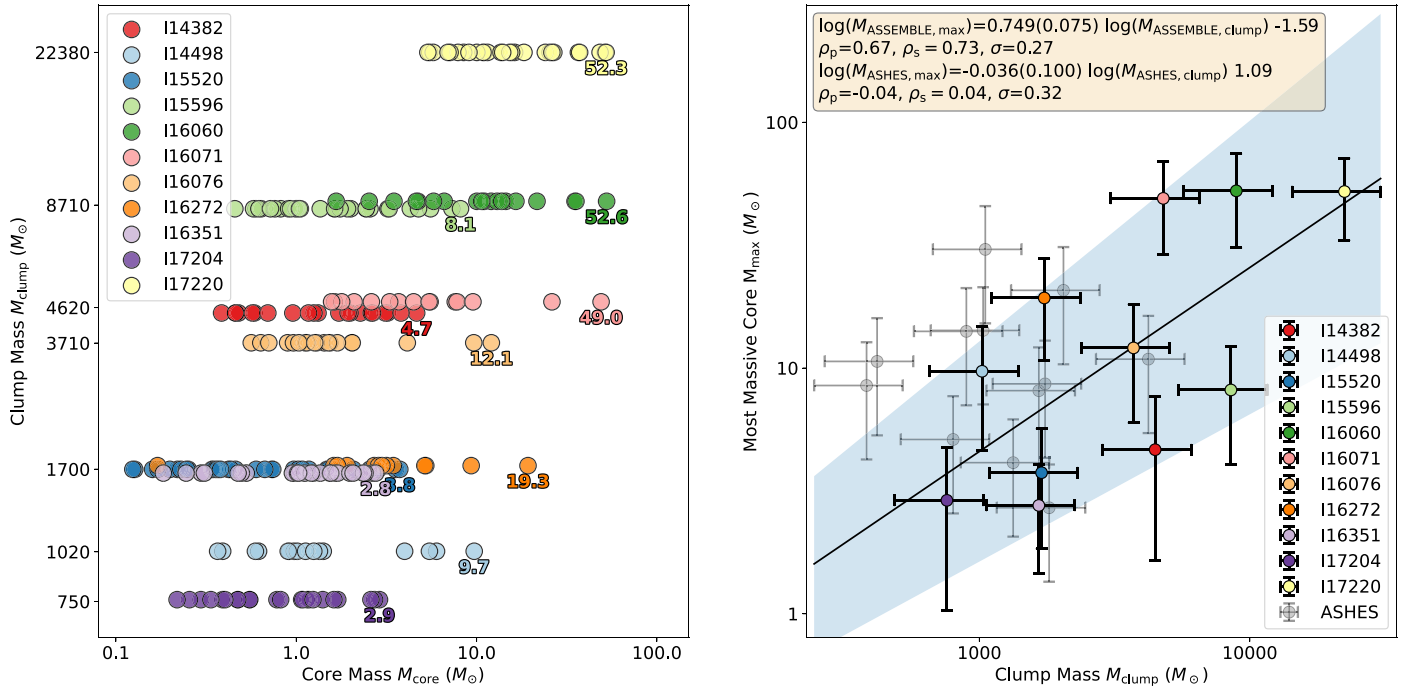

Figure 5. Left: Core masses M_{core} vs. the clump masses M_{clump}. The masses of the most massive cores M_{max} in each clump are labeled with the corresponding colors. Right: The scaling relation between M_{max} and M_{clump}, with the result of linear regression shown on the top left. The shaded area shows the 2σ uncertainty of the fitting result of the ASSEMBLE sample.

prestellar cores (e.g., Zhang et al. 2009; Zhang & Wang 2011; Wang et al. 2012; Cyganowski et al. 2014; Wang et al. 2014; Kong et al. 2017; Sanhueza et al. 2017; Louvet 2018; Li et al. 2019; Molet et al. 2019; Sanhueza et al. 2019; Svoboda et al. 2019; Morii et al. 2021), only a few promising candidates have been identified to date: including G11P6-SMA1(Wang et al. 2014) and G28-C2c1a (Barnes et al. 2023). The rarity of high-mass prestellar cores suggests either that the initial fragmentation of massive clumps does not produce such massive starless cores or that these objects have short lifetimes.

It is worth noting that most of the efforts in the search for massive starless cores have been focused on IRDCs. However, several numerical simulations suggest that thermal feedback from OB protostars and strong magnetic field protostellar clusters can play a crucial role in reducing the level of further fragmentation and producing more massive dense cores (Krumholz et al. 2007; Offner et al. 2009; Krumholz et al. 2011; Myers et al. 2013), and hints for such a reduction of fragmentation for strong magnetic fields have actually been suggested observationally (Palau et al. 2021). Observations also suggest that a $5\,M_\odot$ zero-age main sequence star can produce radiation feedback to support high-mass fragments (Longmore et al. 2011). In particular, massive starless core candidates such as G9.62+0.19MM9 (Liu et al. 2017) and W43-MM1#6 (Nony et al. 2018) have been found in evolved protostellar clusters. Moreover, Contreras et al. (2018) reported a relatively massive but highly subvirial collapsing prestellar core with mass $17.6\,M_\odot$, that is heavily accreting from its natal cloud at a rate of $1.96 \times 10^{-3}\,M_\odot\,\mathrm{yr}^{-1}$. If the accretion rate persists during the lifetime of the massive starless clump ($\lesssim$1–3 $\times\,10^4$ yr), then the mass of the prestellar core can be doubled at the beginning of the protostellar stage. Therefore, it would be even more promising to search for high-mass prestellar cores in protostellar clusters than in prestellar clusters.

Within the ASSEMBLE protoclusters, the most massive prestellar core I17220-ALMA9 has a mass of $18.3\,M_\odot$ within 0.065 pc, which is about 2 times larger than the ones found in the ASHES IRDCs. The second massive prestellar core I16060-ALMA17 has a mass of $16.5\,M_\odot$ within 0.045 pc. However, we should note that (1) the ASSEMBLE data only have the ALMA 12 m array configuration, so the core flux can be underestimated with extended flux filtered out; (2) we adopt the clump-averaged temperature as the temperature of the prestellar core, which can be overestimated, resulting in an underestimated core mass. Complementary short-baseline configuration and a better estimation of temperature should give a better estimate of the prestellar core mass. At any rate, the available evidence strongly suggests that (1) prestellar cores are becoming more massive, which can be due to the continued mass accumulation along with the natal clump (see Section 6.1); (2) high-mass prestellar cores can survive in protostellar clusters. However, to demonstrate the causality between the survival of high-mass prestellar cores and the protocluster environment, both a systematic search for high-mass prestellar cores in massive protoclusters and determination of environmental effects are needed.

5.3. Core Separation

To study the spatial distribution of cores, we first build the minimum spanning tree (MST) for each ASSEMBLE core cluster; and the details can be found in Appendix B.

Following the convention of Wang et al. (2016), Sanhueza et al. (2019), we take the "edge" of MST as the separation between the cores. A total of $N-1$ separation lengths are defined in each clump where N is the core number. The upper panel of Figure 6 shows the distributions of core separation of the ASSEMBLE sample in blue, the ASHES total sample (ASHES Totals; Morii et al. 2023) in green color, and the ASHES pilot sample (ASHES Pilots Sanhueza et al. 2019) in gray color, respectively. When normalized into a probability

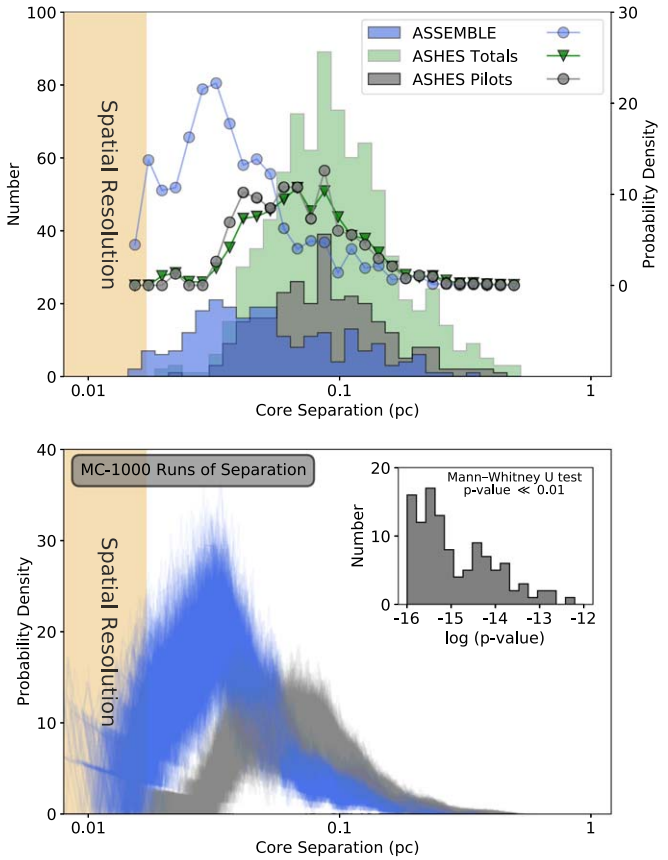

Figure 6. Upper: the number distribution (indicated by the stacked histogram) and probability density distribution (indicated by line-connected points) of core separation are presented in a logarithmic scale, where the ASSEMBLE, the ASHES Total (Morii et al. 2023), and the ASHES Pilot (Sanhueza et al. 2019) samples are presented in blue, green, and gray colors, respectively. The mean spatial resolution of both the ASSEMBLE and the ASHES surveys are ∼0.02 pc, shown with orange shadow. Lower: the 1000 Monte Carlo runs of the probability density distribution of core separation for the ASSEMBLE (blue lines) and the ASHES (gray lines), respectively, considering the Gaussian-like uncertainty of clump distance. The Mann–Whitney U test is performed on each of the sets of core separation distributions, and the distribution of the *p*-value is shown in the top right. The *p*-values are much lower than 0.01, showing that two samples share a significantly different distribution of core separation.

density as shown with line-connected scatter plot, the Kolmogorov–Smirnov test between the distributions of the ASHES Totals and Pilots gives a *p*-value of 0.57 $\gg$0.1, indicating that the ASHES Pilots share the same distribution with that of the ASHES Totals. Therefore, the ASHES Pilots are good enough to represent the ASHES Totals in the case of studying core separation. Since the sample size of the ASHES Pilots is comparable to that of the ASSEMBLE, we only compare core separations from the ASSEMBLE with those from the ASHES Pilots in the following analyses.

The bias of the mass sensitivity and spatial distribution should be excluded. For example, if the ASSEMBLE mass sensitivity is higher than the ASHES one, we are about to detect more low-mass cores, reducing the separation. Thanks to comparable sensitivities of the two samples, we have detected the core population with the same truncation limited by the mass sensitivity. In addition, the ASSEMBLE and ASHES surveys share similar spatial resolutions, as indicated by the orange shadows; and therefore, we can directly compare their core separations.

The Mann–Whitney U test[34] between two groups of core separations gives a *p*-value $\ll$ 0.01, significantly excluding the null hypothesis that two distributions are the same. To further test the effects of the uncertainty of the clump distance, 1000 Monte Carlo runs are adopted to simulate the 1σ distribution dispersion, as shown in the blue and gray extent in the lower panel of Figure 6. The distribution of the *p*-value derived from the Mann–Whitney U test is shown with the subpanel on the upper right corner in the lower panel. Even perturbed by 1σ uncertainty from distance (∼10%–20%), the majority of *p*-values are significantly lower than 0.01, suggesting that two distributions are truly different. In other words, the core separations in the ASSEMBLE protoclusters are systematically smaller than those in the ASHES protoclusters, suggesting that the cluster becomes tighter with closer separations during the clump evolution indicated by L/M. The evolution of core separations is consistent with what has been found in Traficante et al. (2023).

It should be noted that the ASSEMBLE core separation exhibits a significant peak at ∼0.035 pc. The value is twice the spatial resolution (mean value of ∼0.018 pc), suggesting it is not a result of resolution effects. Furthermore, both Tang et al. (2022), Palau et al. (2018) have also observed two peaks in the separation histogram in W51 North and OMC-1S. One of these peaks falls within the range of 0.032–0.035 pc, which aligns with the results we have obtained in our study. Such a consistency between three independent observations (with different spatial resolutions) might suggest a typical level of hierarchical fragmentation at this scale.

5.4. The $\mathcal{Q}$ Parameter

To quantify the spatial distribution of cores, we follow the approach of Cartwright & Whitworth (2004) and define the $\mathcal{Q}$ parameter as

$$\mathcal{Q} = \frac{\bar{m}}{\bar{s}}, \tag{5}$$

where $\bar{m}$ is the normalized mean edge length of the MST given by

$$\bar{m} = \sum_{i=1}^{N_c-1} \frac{L_i}{(N_c A)^{1/2}}, \tag{6}$$

where N_c is the number of cores, L_i is the length of each edge, and A is the area of protocluster as $A = \pi R_{\rm cluster}^2$, with $R_{\rm cluster}$ calculated as the distance from the mean position of cores to the farthest core. $\bar{s}$ is the normalized correlation length,

$$\bar{s} = \frac{L_{\rm av}}{R_{\rm cluster}}, \tag{7}$$

where $L_{\rm av}$ is the the mean separation length between all cores, and $R_{\rm cluster}$ is the cluster radius.

The $\mathcal{Q}$ value serves as a measure of the degree of subclustering and the large-scale radial density gradient in a given region. As indicated by Figure 5 in Cartwright & Whitworth (2004), a value of $\mathcal{Q} \gtrsim 0.8$ indicates a centrally condensed spatial distribution characterized by a radial density

[34] The Mann–Whitney U test is a null hypothesis test, used to detect differences between two independent data sets. The test is specifically for nonparametric distributions, which do not assume a specific distribution for a set of data (Mann & Whitney 1947). Because of this, the Mann–Whitney U test can be applied to any distribution, whether it is Gaussian or not.

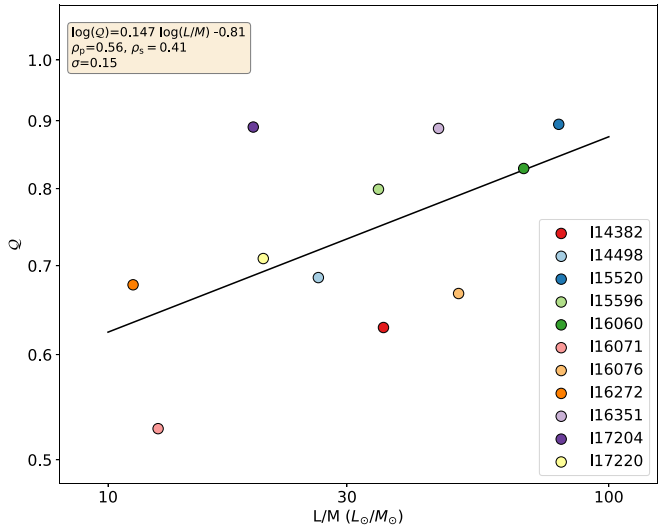

Figure 7. The $\mathcal{Q}$ values vs. luminosity-to-mass ratio L/M in the ASSEMBLE protoclusters. The linear regression results including the fitting model, Pearson correlation coefficient ρ_p, Spearman rank correlation coefficient ρ_s, and scatter σ are shown on the upper left.

profile of the form $n(r) \propto r^{-\alpha}$. On the other hand, when $\mathcal{Q} \lesssim 0.8$, the $\mathcal{Q}$ parameter decreases from approximately 0.80 to 0.45 with an increasing degree of subclustering, ranging from a fractal dimension of $D = 3.0$ (representing a uniform number-density distribution without subclustering), to $D = 1.5$ (indicating strong subclustering).

From the MST results, the derived $\mathcal{Q}$ parameters for the ASSEMBLE clumps range from 0.53 to 0.89, with a median value[35] of 0.71(0.13). We note that there are four protoclusters I15520, I16060, I16351, and I17204 that have $\mathcal{Q}$ greater than 0.8, indicative of a centrally condensed spatial distribution. As shown in Figure 7, the $\mathcal{Q}$ parameter shows a weak correlation with luminosity-to-mass ratio, with Pearson correlation coefficient $R_p = 0.56$. The positive correlation suggests that a protocluster is becoming more centrally condensed as it evolves. In Section 6.5, a correlation among a sample of both ASSEMBLE and ASHES could be more instructive, since a wider dynamic range of L/M is available.

5.5. Mass Segregation

As defined in Allison et al. (2009), Parker & Goodwin (2015), mass segregation refers to a more concentrated distribution of more massive objects with respect to lower-mass objects than that expected by random chance. For dynamically old bound systems (i.e., relaxed and virialized clusters), the process of two-body relaxation has redistributed energy between stars, and they approach energy equipartition whereby all stars have the same mean kinetic energy. Therefore, more massive stars will have a lower velocity dispersion, and they will sink into the deeper gravitational potential, i.e., the center of the cluster (Spitzer 1969).

Despite the observed mass segregation in old stellar clusters, it does not have to be from a canonical two-body relaxation dynamical process. If we observe mass segregation in a region that is so young that two-body encounters cannot have mass segregated the stars, then the mass segregation must be set by some aspect of the star formation process, and is often called "primordial mass segregation" (Parker & Goodwin 2015), which has been found in some simulations of star formation (e.g., Moeckel & Bonnell 2009; Myers et al. 2014). Observationally, Sanhueza et al. (2019) have only found weak mass segregation in 4 out of 12 IRDCs and no mass segregation in the others. The overall conclusion is that there is no significant evidence of primordial mass segregation in IRDCs (Sanhueza et al. 2019; Morii et al. 2023). In contrast, at a similar physical resolution of 2400 au and the same band (1.3 mm) by ALMA, Dib & Henning (2019) have found massive star-forming region W43 exhibits evident mass segregation with maximum mass segregation ratio $\Lambda_{\rm MSR}^{\rm max} = 3.49$ (see definition in Equation (8)).

5.5.1. Λ Plots: Characterization of Mass Segregation

To quantify the mass segregation in the protoclusters, we adopt the mass segregation ratio (MSR), $\Lambda_{\rm MSR}$, which is defined by Allison et al. (2009) and shown to perform best compared to three other methods by Parker & Goodwin (2015). The value of $\Lambda_{\rm MSR}$ at $N_{\rm MST}$ is given by

$$\Lambda_{\rm MSR}(N_{\rm MST}) = \frac{\langle l_{\rm random}\rangle}{l_{\rm massive}} \pm \frac{\sigma_{l,\rm random}}{l_{\rm massive}}, \quad (8)$$

where $l_{\rm random}$ is the mean MST edge length of an ensemble of $N_{\rm MST}$ cores randomly chosen from the protocluster, and $l_{\rm massive}$ is the mean MST length of the top-$N_{\rm MST}$ MMCs. In our analyses, we performed 1000 Monte Carlo runs of choosing $N_{\rm MST}$ random cores to obtain a set of $l_{\rm random}$, calculating the mean value $\langle l_{\rm random}\rangle$ and its standard deviation $\sigma_{l,\rm random} = \sqrt{\langle (l_{\rm random} - \langle l_{\rm random}\rangle)^2\rangle}$. For each $N_{\rm MST}$, $\Lambda_{\rm MSR}$ is meant to measure how much the MST length of the top-$N_{\rm MST}$ MMCs deviate from the MST length of the entire protocluster. If the MST length of the top-$N_{\rm MST}$ ensemble is shorter than the MST length of the entire protocluster, it is suggested that massive cores have a more concentrated distribution.

By definition, $\Lambda \simeq 1$ means that the massive cores were distributed in the same way as the other cores (i.e., no mass segregation); $\Lambda > 1$ means that the massive cores were concentrated (i.e., mass segregation), and $\Lambda < 1$ means that the more massive cores were spread out relative to the other cores (i.e., inverse-mass segregation).

Figure 8 presents mass segregation ratio $\Lambda_{\rm MSR}$ versus the fraction of the selected core number to the total core number $f_{\rm MST} = N_{\rm MST}/N_{\rm MST,max}$, which is called $\Lambda_{\rm MSR}$ *plot* hereafter. We arrange the protoclusters in descending order of the maximum value of the mass segregation ratio $\Lambda_{\rm MSR,max}$. For example, the protocluster I16071 in the first panel has the highest $\Lambda_{\rm MSR,max}$ of 8.72($\pm$ 3.69), which implies strong mass segregation. In contrast, the protocluster I14382 has no mass segregation or even a weak inverse-mass segregation, as shown in the last panel.

There are three notable features that deserve additional explanations in the Λ plots:

1. $\Lambda_{\rm MSR}$ *peak at small* $f_{\rm MST}$. Protoclusters have a wide range of dense core mass while there are a small number of massive dense cores. When $f_{\rm MST}$ or $N_{\rm MST}$ are small, massive dense cores should account for a large proportion

[35] The uncertainty of median value $\sigma_{\rm med}$ is estimated from the median absolute deviation (MAD) as, $\sigma_{\rm med} \simeq 1.4826 \times$ MAD, based on the assumption of normality.

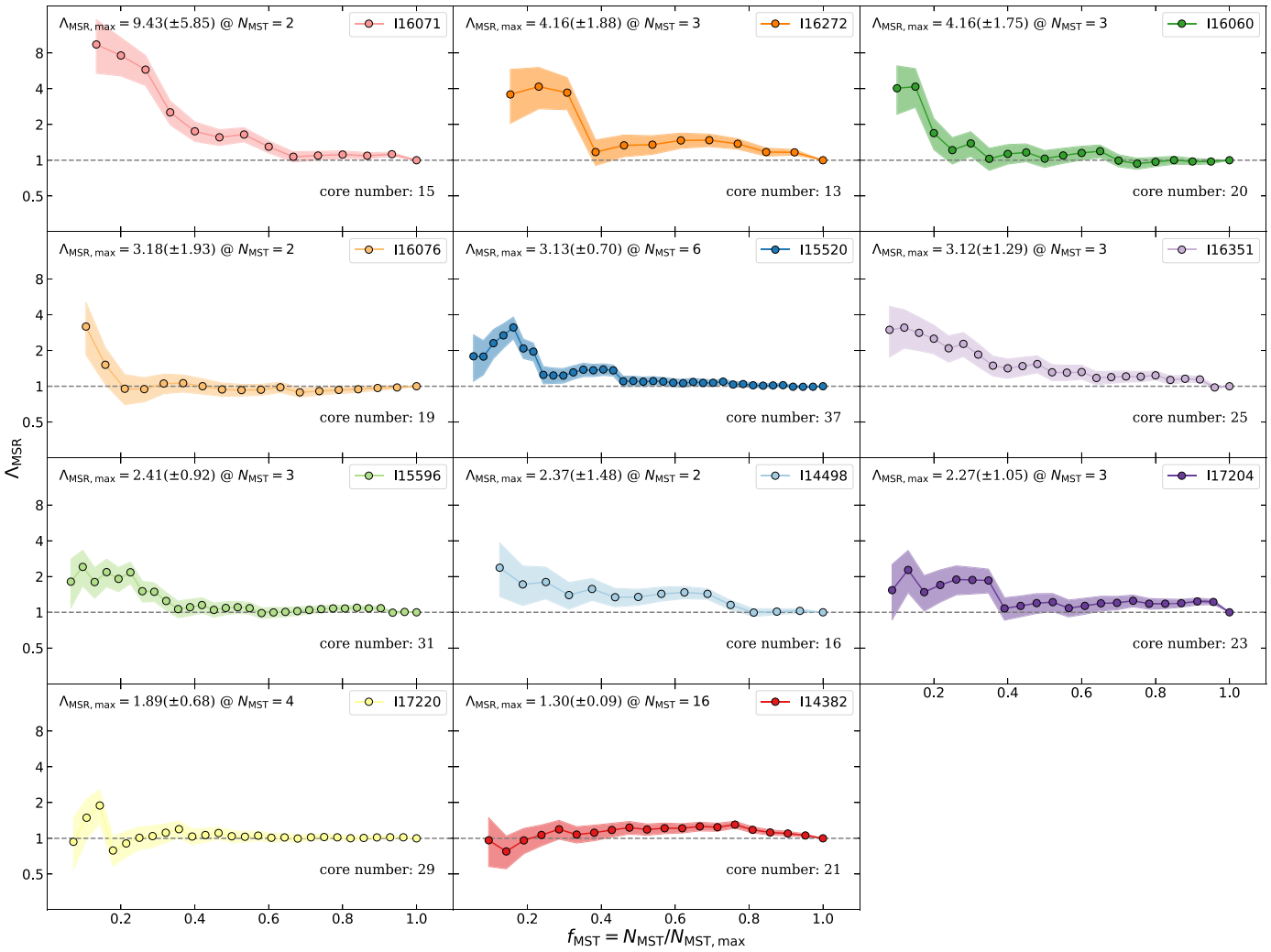

Figure 8. The mass segregation ratio $\Lambda_{\rm MSR}$ plots: $\Lambda_{\rm MSR}$ is presented as a function of the fraction of core number $f_{\rm MST} = N_{\rm MST}/N_{\rm MST,max}$ for the 11 protoclusters. The shaded regions in each panel represent the local uncertainties. The upper left corner of each panel shows the maximum value of $\Lambda_{\rm MSR,max}$ and its corresponding $N_{\rm MST}$. The $\Lambda_{\rm MSR}$ vs. $f_{\rm MST}$ panels are arranged in descending order of $\Lambda_{\rm MSR,max}$ for the corresponding protoclusters. The total core number is listed at the lower right corner.

in the ensemble, so $l_{\rm massive}$ should be significantly smaller than $\langle l_{\rm random}\rangle$ if mass segregation exists.

2. $\Lambda_{\rm MSR}$ *decrease with* $f_{\rm MST}$. When $N_{\rm MST}$ increase, $l_{\rm massive}$ will involve more low-mass cores so that the mass segregation trend, if it exists, will be washed out; furthermore, when $f_{\rm MST}$ is larger, the ensembles of cores used to compute both $l_{\rm massive}$ and $\langle l_{\rm random}\rangle$ are more similar to the entire core sample so that both quantities theoretically approach the same value, the MST length of the entire core sample.
3. *Diverse* $\Lambda_{\rm MSR}$ *profiles or diverse fractions of cores involved in mass segregation*. $\Lambda_{\rm MSR}$ drops with $f_{\rm MST}$ with different rates. The clump with the strongest mass segregation, I16071, has its $\Lambda_{\rm MSR}$ dropping toward 1 around $f_{\rm MST} \simeq 0.6$, while the clump with the second largest mass segregation, I16060, has its $\Lambda_{\rm MSR}$ rapidly dropping toward 1 around $f_{\rm MST} \simeq 0.2$. Such diversity is also true among the clumps with lower degrees of mass segregation (e.g., I15596 versus I14498). Therefore, it is of great interest to understand why the different protoclusters can show such different profiles of $\Lambda_{\rm MSR}$ plots in the future.

5.5.2. $\mathcal{I}_\Lambda^{\rm MSR}$: *Mass Segregation Integral*

$\Lambda_{\rm MSR}$ plots are difficult to compare with each other, because $\Lambda_{\rm MSR}$ by definition depends on $N_{\rm MST}$ or $f_{\rm MST}$. In other words, to fully characterize the degree of mass segregation of a protocluster, two main factors need to be taken into account: (1) $\Lambda_{\rm MSR,max}$, directly determines what the largest deviation from the random process is, according to the definition of Equation (8); (2) $N_{\rm MST,crit}$ or $f_{\rm MST,crit}$, which determines at what point the mass segregation ratio of cluster disappears for parameter $N_{\rm MST}$ or $f_{\rm MST}$.

Here, we propose mass segregation integral (MSI) $\mathcal{I}_\Lambda^{\rm MSR}$ to describe how a cluster is segregated,

$$\mathcal{I}_\Lambda^{\rm MSR} \equiv \frac{\sum_{i=2}^{N_{\rm MST,max}} \Lambda_{{\rm MSR},i}}{N_{\rm MST,max} - 1}, \tag{9}$$

where $\Lambda_{{\rm MST},i}$ is the mass segregation ratio at given $f_{\rm MST}$ or $N_{\rm MST}$. The MSI is meant to record every deviation from $\Lambda_{{\rm MSR},i} = 1$ (when there is no mass segregation) at each $N_{\rm MST}$.

In Section 6.6, we will examine the significant mass segregation observed in the ASSEMBLE protoclusters and explore its possible origins using the MSI. However, it is

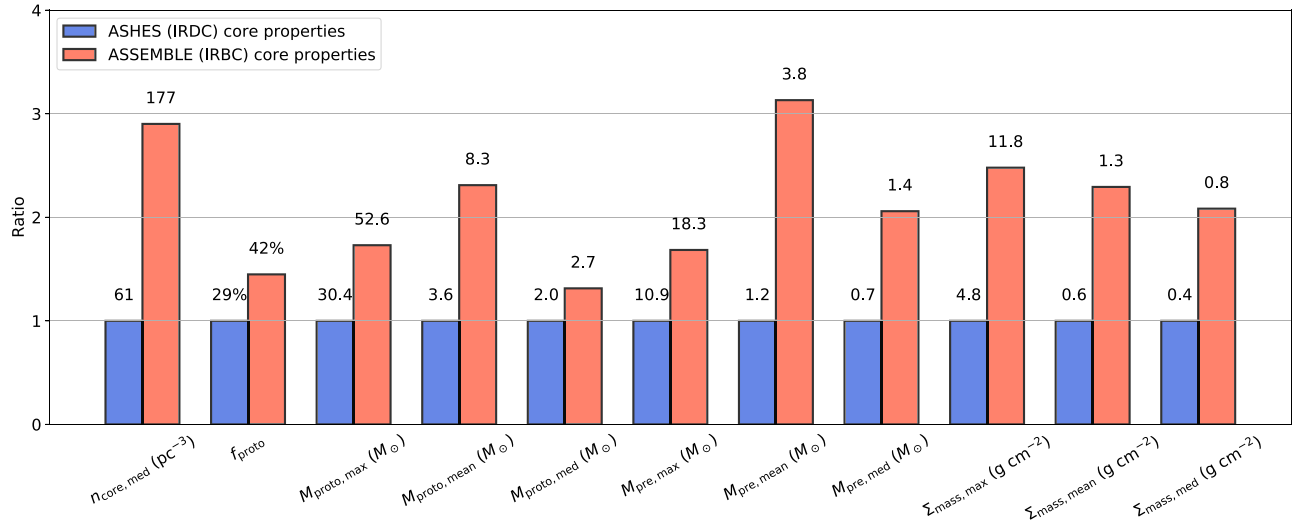

Figure 9. Comparison of parameters including median volume density of the core number $n_{core,med}$, number fraction of protostellar cores f_{proto}, the maximum, mean, and median value of mass of protostellar cores $M_{proto,max}/M_{proto,mean}/M_{proto,med}$, mass of prestellar cores $M_{pre,max}/M_{pre,mean}/M_{pre,med}$, and core surface mass density $\Sigma_{mass,max}/\Sigma_{mass,mean}/\Sigma_{mass,med}$. The IRDCs and IRBCs are in blue and red, respectively. The y value is normalized by the ratio to the value of IRDCs.

important to acknowledge limitations of using the MSI. First, the MSI collapses the Λ_{MSR} profile into a scalar value, thus disregarding the potentially various spatial distributions that can produce the same MSI value. Second, the physical and mathematical interpretations of the MSI are not yet fully understood, as it only records the deviations from the random distribution of cores within a protocluster. To enhance our understanding, future studies could establish a correlation between the MSI and the evolutionary timescale of a protocluster.

6. Evolution of Massive Protoclusters

The ASSEMBLE clumps have evolved to a late stage in the formation of massive protoclusters, with an L/M ratio ranging from 10 to 100 $L_\odot/M_\odot$. On the contrary, the ASHES clumps are in an early stage, with an L/M ratio between 0.1 and 1 $L_\odot/M_\odot$. Therefore, the ASSEMBLE and ASHES clumps can serve as mutual informative comparison groups, as the basis of our dynamic view of protocluster evolution. As introduced in Section 4.4, the statistics of the core parameters are summarized in Table 5, which can be directly compared to Table 5 in Sanhueza et al. (2019). To highlight the quiescent and active nature of ASHES and ASSEMBLE clumps, respectively, the samples also have the second names IRDCs and infrared bright clouds (IRBCs), respectively.

6.1. Core Growth and Mass Concentration

As shown in Figure 9, the IRBCs exhibit a median volume density of the core number $n_{core,med}$ of 177 per pc^3, which is approximately 3 times greater than the 61 per pc^3 observed in the IRDCs. We consider the potential effects from the different mass sensitivities between the two projects. The slightly worse sensitivity ($\bar{\sigma} = 0.089\ M_\odot$) of the ASSEMBLE compared to the ASHES ($\bar{\sigma} = 0.078\ M_\odot$) shows that correcting for sensitivity would only increase the core density in IRBCs. To exclude the effects of different source extraction algorithms, we also perform the source extraction using `getsf` in the 12 m alone data of ASHES as it was done for the ASSEMBLE in Appendix B, only obtaining a much lower core number of 66, mostly due to two factors: (1) `getsf` tends to extract spherical cores but miss those irregular ones; (2) the 12 m alone data filter out large-scale structures that are previously identified by `astrodendro` algorithm. Therefore, correcting the effects from the array configuration and source extraction algorithm will only result in an even larger difference between the two sets of parameters mentioned above. In any circumstances, the core number densities in the ASSEMBLE clumps are considerably higher than those of IRDCs.

As demonstrated in simulations by Camacho et al. (2020), massive clumps accrete mass and increase density as they evolve, resulting in a decrease in the freefall timescale. Consequently, the dense cores formed by Jeans fragmentation collapse to form protostars more quickly, leading to a higher fraction of protostellar cores, f_{proto}. As shown in the second column of Figure 9, f_{proto} increases significantly from 29% in the early stage IRDCs to 42% in the late-stage IRBCs on average. The increasing trend of f_{proto} with respect to evolutionary stage has also been previously reported by Sanhueza et al. (2019) and is consistent with the fragmentation reported by Palau et al. (2014, 2015, 2021) in more evolved IRBCs; because in these works, most of the cores are protostellar (given their higher masses and compactness compared to the ASSEMBLE sample).

Furthermore, we provide several pieces of evidence for the growth of the core mass from IRDC to IRBC in columns (3)–(11) in Figure 9. Parameters, including the maximum, mean, and median mass of protostellar cores $M_{proto,max}$, $M_{proto,mean}$, and $M_{proto,med}$, respectively; those of prestellar cores $M_{pre,max}$, $M_{pre,mean}$, and $M_{pre,med}$, respectively; and those of surface mass density of total core population $\Sigma_{mass,mean}$, $\Sigma_{mass,max}$, and $\Sigma_{mass,med}$, respectively, all exhibit systematic increments from IRDCs to IRBCs. These mass or surface density increments have also been observed in another comparative work between IRDCs and IRBCs with hub-filament structures (Liu et al. 2023), where gas inflow is thought to be responsible for the hierarchical and multiscale mass accretion (Galván-Madrid et al. 2010; Liu et al. 2022a, 2022b, 2023; Xu et al. 2023a; Yang et al. 2023). Very recent statistical studies of dense cores in the Dragon IRDC

(Kong et al. 2021), the Orion Molecular Cloud (OMC; Takemura et al. 2023), the ASHES IRDC sample (Li et al. 2023), and the ALMA-initial mass function protoclusters (Nony et al. 2023; Pouteau et al. 2023) also suggest that the protostellar cores are considerably more massive than the starless cores, suggesting that cores grow with time. If the missing flux in the ASSEMBLE sample were recovered, the effect would be to increase the core mass and surface density, strengthening the arguments above.

6.2. Nurture *but Not* Nature: *A Dynamic View of the* M_{max} *versus* M_{clump} *Relation*

As discussed in Section 5.1, a positive correlation between $M_{\rm max}$ and $M_{\rm clump}$ is observed within the ASSEMBLE protoclusters, suggesting a close relationship between the natal clump and the MMC through multiscale gas accretion (Xu et al. 2023a). On the contrary, the ASHES pilots, represented by the gray data points in the right panel of Figure 5, exhibit no significant $M_{\rm max}$ versus $M_{\rm clump}$ correlation. The Spearman rank correlation coefficient is calculated to be 0.04, with a corresponding p-value of 0.90. This lack of correlation aligns with the concept of thermal Jeans fragmentation (e.g., Palau et al. 2015, 2018; Sanhueza et al. 2019), where the clump's thermal Jeans mass is primarily determined by its dust temperature within a narrow range of 10–20 K (Morii et al. 2023), rather than the turbulence whose energy is governed by the clump's gravity assuming energy equipartition (Palau et al. 2015). This finding supports the notion that early stage cores are characterized by dominant initial fragmentation rather than gravitational accretion. In that case, no correlation is naturally expected between the mass of the natal clump and the mass of the core resulting from fragmentation. Therefore, we propose a dynamic picture of the clump–core connection.

1. At the beginning, initial Jeans fragmentation produces a set of dense cores whose mass is not associated with clump-scale gravitational potential (mass) and turbulence.
2. As a massive clump evolves, multiscale continuous gas accretion helps build up the connection between clump and core scales, for example the mass correlation that we have observed.

6.3. *Implications of the* M_{max} *versus* $M_{cluster}$ *Relation*

The relation $M_{\star\,\rm max}$ versus $M_{\star\rm cluster}$, which describes the relationship between the mass of the most massive star ($M_{\star\,\rm max}$) and the total mass of the star cluster ($M_{\star\rm cluster}$), has been previously established both analytically (Weidner & Kroupa 2004) and observationally (e.g., Testi et al. 1999; Weidner & Kroupa 2006). This relation highlights the systematic variation of the typical upper mass limit with the overall mass of the star cluster. It suggests that the formation of stars within cloud cores is primarily influenced by growth processes occurring in an environment with limited resources. This finding underscores the significance of resource availability in shaping the stellar population within star clusters.

Protoclusters provide a retrospective glimpse toward the early version of star clusters. Throughout this paper, we refer $M_{\rm cluster}$ as the sum of all the core masses in a protocluster. Note that $M_{\rm cluster}$ is different than the total mass of a stellar cluster ($M_{\star\rm cluster}$). As shown in the left panel of Figure 5, $M_{\rm cluster}$ are plotted with the masses of the MMCs $M_{\rm max}$ in Figure 10.

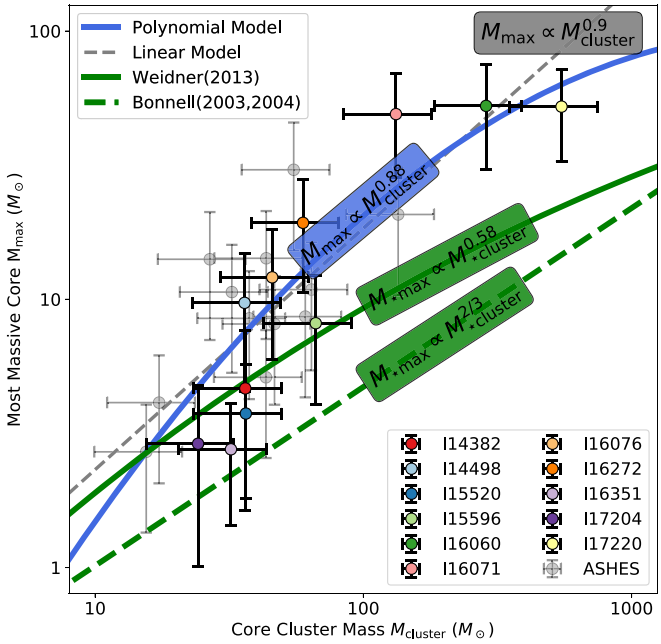

Figure 10. The $M_{\rm max}$ vs. $M_{\rm cluster}$ relation. The ASSEMBLE and the ASHES Pilots are plotted with errorbars. The blue solid line shows the second-order polynomial fitting of all the data points. $M_{\rm max} \propto M_{\rm cluster}^{0.88}$ in the blue box indicates the power-law index in the first-order approximation. The gray dashed line shows the linear regression of all the data points. The green solid line shows the polynomial fitting of the data from 139 star clusters (Weidner et al. 2013), with first-order approximation of $M_{\star\,\rm max} \propto M_{\star\rm cluster}^{0.58}$ in the upper green box. The dashed green line as well as the lower green box shows the relation from smoothed particle hydrodynamics simulation (Bonnell et al. 2003, 2004).

Contrary to what has been found in the $M_{\rm max}$ *versus* $M_{\rm clump}$ plane (shown in the right panel of Figure 5), both the ASHES and the ASSEMBLE protoclusters have positive correlation between $M_{\rm max}$ and $M_{\rm cluster}$. A second-order polynomial model $y = -0.27x^2 + 1.96x - 1.51$ fits the $M_{\rm max}$ versus $M_{\rm cluster}$ relation best, with a correlation coefficient of 0.68, which is shown in the blue solid line. Besides, we also present the relation in stellar clusters given by observations (Weidner et al. 2013) in the green solid line and smoothed particle hydrodynamics simulations (Bonnell et al. 2003, 2004) in the green dashed line.

In order to facilitate a direct comparison between the stellar cluster and protostellar cluster, we have also performed a first-order approximation of the polynomial models represented by the blue and green solid lines. By utilizing the mean value theorem, the first-order power-law index can be estimated by considering the average derivatives within the given value range. The estimation of the power-law index is indicated in the blue and green boxes, which are overlaid on the respective solid lines. As shown in the gray dashed line, we directly perform the linear regression to derive a power-law index of 0.9, validating our first-order approximation of 0.88.

Despite uncertainties, the slope of the $\log M_{\rm max}$ versus $\log M_{\rm cluster}$ relation is notably steeper compared to that of the $\log M_{\star\,\rm max}$ versus $\log M_{\star\rm cluster}$ relation. To reconcile this disparity within the context of protocluster evolution, we take into account the influence of multiple star systems on massive star formation. As depicted in Figure 1 by Offner et al. (2023), the probability of events involving multiplicity is nearly 100%.

In other words, it is highly likely that massive cores give rise to the formation of more than one massive star. Hence, it is natural to expect that the slope of the $\log M_{\rm max}$ versus $\log M_{\rm cluster}$ relation can evolve into a shallower version, akin to what is observed in the $\log M_{\star\,\rm max}$ versus $\log M_{\star\rm cluster}$ relation.

Another noteworthy finding is the similarity in the total cluster mass distribution between the ASHES and ASSEMBLE protoclusters, particularly within the range of 10–100 $M_\odot$, when excluding four outliers with $M_{\rm cluster} > 100\ M_\odot$. Since these outliers also exhibit higher clump masses (refer to Figure 5), a more fundamental question arises: Why do these protoclusters with different evolutionary stages consistently maintain a mass proportion of cluster to clump ($M_{\rm cluster}/M_{\rm clump}$) between 1% and 10%? This proportion can be regarded as the dense gas fraction (DGF), which is often closely associated with star formation efficiency (Ge et al. 2023). Consequently, it is of great significance to investigate the evolution of DGF in relation to massive star-forming clumps (Xu et al. 2023a).

6.4. Gravitational Contraction: Protoclusters Evolve to Greater Compactness

As shown in Figure 6, the core separation distribution of the ASSEMBLE protoclusters have two prominent features. One is a peak at 0.035 pc as discussed in Section 5.3, systematically smaller than what has been found in the ASHES, meaning that the spatial distribution becomes tighter and more compact as the protocluster evolves. The other one is an extended tail at 0.06–0.3 pc, numerically consistent with what has been found in the ASHES protoclusters of 0.06–0.24 pc (refer to green or gray histograms of Figure 6), which are assumed to be the residuals of the initial fragmentation at the early stage. In this section, we discuss how gravity leads to the tightening process of protoclusters and complete the dynamic picture of fragmentation and gravitational contraction.

We can make a simple semiquantitative calculation. Given that the thermal Jeans fragmentation is observed to dominate at the early stage of massive star formation (Palau et al. 2014, 2015; Sanhueza et al. 2019; Palau et al. 2021; Li et al. 2022; Morii et al. 2023), we assume the initial condition that the cores could have initially fragmented on Jeans length scales of ∼0.14 pc (mean Jeans lengths in the ASHES; Sanhueza et al. 2019). If dense cores are moving toward the center of the clump by gravity, then the velocity of the cores should be freefall velocity as $v_{\rm ff} = \frac{R_{\rm cl}}{t_{\rm ff}}$. Adopting the typically observed massive clump size and density of $R_{\rm cl} = 1$ pc, and $n_{\rm H_2} = 10^4$ cm^{-3}, the core separation will be tighter by

$$\begin{aligned}\Delta l_{\rm gc} &= v_{\rm ff}\, t_{\rm life} = \frac{R_{\rm cl}}{t_{\rm ff}} \times t_{\rm life} \\ &= 0.088 \left(\frac{R_{\rm cl}}{1\ {\rm pc}}\right)\left(\frac{n_{\rm H_2}}{10^4\ {\rm cm^{-3}}}\right)^{0.5}\left(\frac{t_{\rm life}}{5\times10^4\ {\rm yr}}\right){\rm pc}, \end{aligned} \quad (10)$$

where $t_{\rm life} \sim 0.2$–1×10^5 yr (Motte et al. 2018) are the freefall timescale and the statistical lifetime of massive starless clumps, respectively. Therefore, the core separation should tighten by $\Delta l_{\rm gc} \simeq 0.04$–$0.18$ pc in the protoclusters by gravitational contraction, numerically consistent with the shift from extended tail (0.06–0.3 pc) to the observed separation (peaked at 0.035 pc).

The simple gravitational contraction model fits the observations well, indicating ongoing bulk motions from the global gravitational collapse of massive clumps (Beuther et al. 2018; Vázquez-Semadeni et al. 2019). But note that we are still unable to rule out another possibility that the closer separation is due to hierarchical fragmentation to produce a series of condensations inside a massive core.

6.5. Evolution of the $\mathcal{Q}$ Parameter

As clumps evolve over time, the primordial distribution of cores dissolves due to dynamical relaxation, leading to a more radially concentrated structure as predicted by simulations (Guszejnov et al. 2022). Consequently, the more-evolved clumps are predicted to have higher $\mathcal{Q}$ values. In the 12 ASHES Pilots, Sanhueza et al. (2019) used the fraction of protostellar cores $f_{\rm proto}$ to gauge the evolutionary stage. Due to the narrow parameter space of similar evolutionary properties such as dust temperature $T_{\rm dust}$ (10–15 K) and luminosity-to-mass ratio L/M (0.1–1 $L_\odot/M_\odot$), only a weak correlation between $\mathcal{Q}$ and $f_{\rm proto}$ was found (Sanhueza et al. 2019). However, Dib & Henning (2019) found that the most active star-forming region W43 has a higher $\mathcal{Q}$ value compared to more quiescent regions (L1495 in the Taurus, Aquila, and Corona Australis). These studies inspire a larger sample with wide evolutionary stages to shed more light on the interplay between star formation in clouds and the spatial distribution of dense cores.

The combination of the ASSEMBLE and ASHES clumps provides a systematic sample with a wide dynamic range of evolutionary stage (L/M of 0.1–100 $L_\odot/M_\odot$ and $T_{\rm dust}$ of 10–35 K). To make the comparison between two samples more directly, we have simulated the mock 0.87 mm continuum data with only the 12 m array configuration (see details in Appendix C). Following the same procedure of core extraction, we have an updated ASHES core catalog used for the MST algorithm (see more details in Appendix B). The $\mathcal{Q}$ parameters for the ASHES sample range from 0.40 to 0.75, with a median value of 0.61(0.13). The Mann–Whitney U test between the $\mathcal{Q}$ parameters of the ASSEMBLE and ASHES samples has a p-value of 0.03 (<0.05), showing the two samples have significantly different $\mathcal{Q}$ parameters.

The linear regressions between the $\mathcal{Q}$ parameters and the evolutionary indicators (L/M and $T_{\rm dust}$) are performed in the log versus log space and shown in Figure 11. The positive correlations, between both L/M and $T_{\rm dust}$, indicate that the $\mathcal{Q}$ parameters evolve with time. The correlations are confirmed to be statistically significant by the high Pearson correlation coefficients $\rho_{\rm p}$ of 0.61 and 0.60 with p-values of 0.0024 and 0.0034 for L/M and $T_{\rm dust}$, respectively. Moreover, the Spearman correlation coefficients $\rho_{\rm s}$ are 0.57 and 0.53 with p-values of 0.0056 and 0.0088. Statistically, it is tentatively evident that the $\mathcal{Q}$ parameter of the protostellar clusters should increase in later evolutionary stages, indicating more subclustering distribution at an early stage but more centrally condensed structure when the cluster evolves, which agrees with the results and predictions in Sanhueza et al. (2019).

6.6. Origin of Mass Segregation

Compared to the early stage clusters previously reported by Sanhueza et al. (2019), Morii et al. (2023), we have identified three main differences in mass segregation according to the Λ-plots in Figure 8. First, evident mass segregation (with

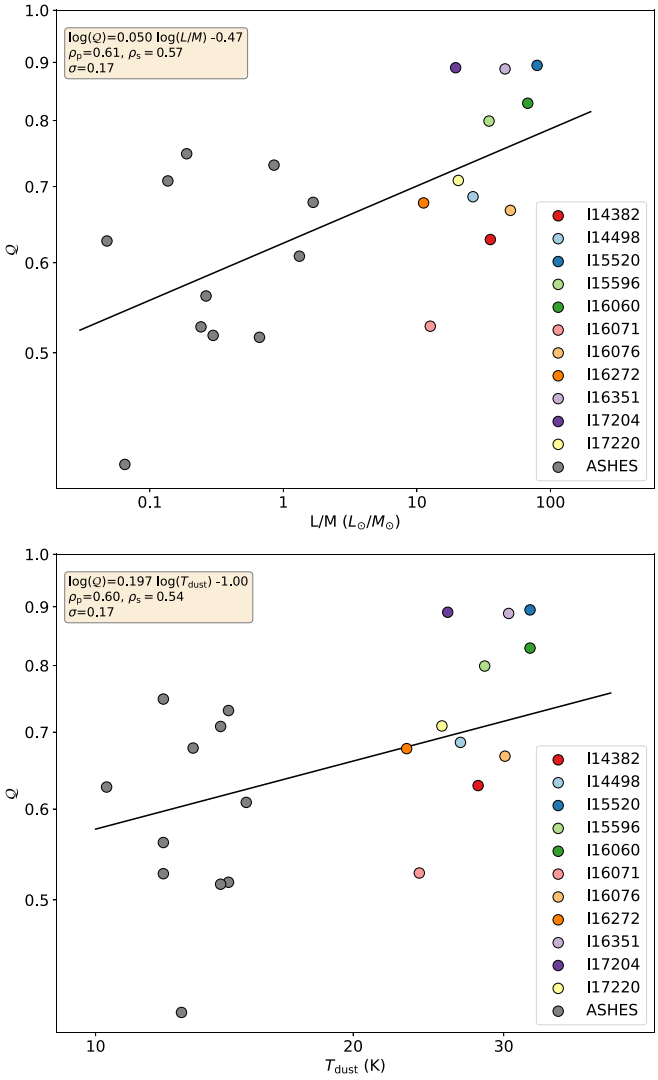

Figure 11. The $\mathcal{Q}$ values vs. luminosity-to-mass ratio L/M (left) and temperature $T_{\rm dust}$ (right). The linear regression results including the fitting model, Pearson correlation coefficient $\rho_{\rm p}$, Spearman rank correlation coefficient $\rho_{\rm s}$, and scatter σ are shown on the upper left. ASHES clumps are shown in the $\mathcal{Q}$ vs. $T_{\rm dust}$ panel with gray points.

$\Lambda_{\rm MSR} \gtrsim 3$) was found in 73% (8 out of 11) ASSEMBLE protoclusters, which is >5 times more than it was identified in the ASHES sample (13%, 5 out of 39; Morii et al. 2023). Second, the mass segregation ratios we observed were significantly higher, with some clusters exhibiting values as large as ∼9. Finally, we observed $\Lambda_{\rm MSR} > 1$ even for larger core numbers ($\gtrsim$10) in certain protoclusters such as I16351, I15520, and I15596.

Using the MSI introduced in Section 5.5.2, we present a direct comparison between the ASSEMBLE and the ASHES Totals, as illustrated in Figure 12. The Mann–Whitney U test reveals significant differences between the ASSEMBLE and the ASHES protoclusters, as indicated by the green histogram of p-values. To establish a reference sample for statistical analysis, we simulate 100 clusters with mean MSI of $\mu = 0$ (no mass segregation), and with the random perturbation of $\sigma = 0.1$ (assumed to be the same as typical uncertainties when calculating the MSI) in MSI. The Mann–Whitney U test rejects the null hypothesis that the MSI of the ASSEMBLE

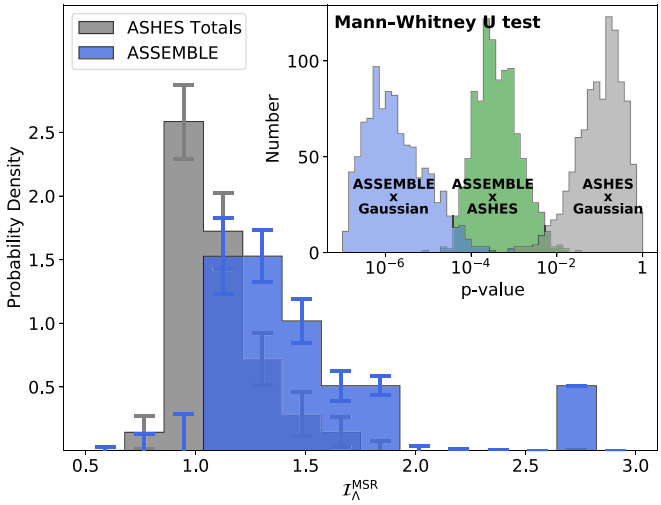

Figure 12. The mass segregation integrals $\mathcal{I}_\Lambda^{\rm MSR}$ of the ASSEMBLE and the ASHES are shown as blue and gray histograms with errorbars. The p-values of the Mann–Whitney U test between ASSEMBLE, ASHES, and Gaussian distributions ($\mu = 1$, $\sigma = 0.1$) are shown in three colors, which are attached in the upper right corner.

protoclusters follows the random perturbation (p-value $\ll$ 0.01), highlighting the presence of evident mass segregation. In contrast, the null hypothesis cannot be confidently rejected for the ASHES protoclusters, with mean and median p-values of 0.18 and 0.12, respectively. Thus, the ASSEMBLE protoclusters exhibit robust evidences of mass segregation, whereas the mass segregation in the ASHES protoclusters is weak to moderate.

In the context of protocluster evolution, the degree of mass segregation increases unambiguously from the ASHES clusters to the ASSEMBLE clusters (this work). Therefore, the natural question is the origin of mass segregation. Here, we test whether mass segregation can result from the canonical dynamical relaxation by a two-body relaxation.

To analyze the dynamics of the cluster, we adopt the formulation of Reinoso et al. (2020), who extended the framework of Spitzer (1987) to include the effect of a gas potential. The crossing time of the cluster is then given as

$$t_{\rm cross} = \frac{R}{V_{\rm vir}}, \tag{11}$$

with velocity under the virial equilibrium $V_{\rm vir}$,

$$V_{\rm vir} = \sqrt{\frac{GM_{\rm core}}{R}}\,(1+q), \tag{12}$$

and $q = M_{\rm gas}/M_{\rm core}$. Here, $M_{\rm gas}$ and $M_{\rm cluster}$ are ambient low-density gas mass and total dense core cluster mass, respectively. R is radius of cluster. The relaxation time is then given as

$$t_{\rm relax} = 0.138\frac{N(1+q)^4}{\ln(\gamma N)}t_{\rm cross}, \tag{13}$$

where N is the number of core in a cluster, and γ is a constant of proportionality in the term of virial velocity. The γ value is between 0.42 and 0.38 for the polytropes of index of the cluster system between 3 and 5, and the $\gamma = 0.4$ provides a reasonably

good approximation for most systems (Spitzer 1969). So we use $\gamma = 0.4$ here.

We take the ASHES sample as the initial condition for our protocluster analyses, assuming a typical value of $R = 0.5$ pc, $N = 25$, and $M_{\rm cluster} = 100\ M_\odot$ (i.e., the mass of the protoclusters). For the gas mass, we used two different methods. The first method is calculating the total gas mass based on an average volume density of 5×10^4 cm^{-3} and a radius of $R = 0.5$ pc (from Table 1 in Morii et al. 2023), resulting in a gas mass of approximately $\sim 200\ M_\odot$. The second method considers that the ALMA recovered flux only comes from the dense cores that we identified, and the missing flux should come from diffuse gas, both of which are covered by the ATLASGAL emission. As shown in Section 4.2, the flux ratio of ALMA to ATLASGAL has a mean value of ∼20%. Therefore, the total gas mass should be 4 times larger than the total mass of the core cluster, giving a value of $400\ M_\odot$. Two independent methods yield $M_{\rm gas}$ within the range of 200–400 $M_\odot$. By considering both methods, we derived the q value of 2–4. Taking all factors into account, we found that the typical relaxation time of a protocluster was as long as 70–500 Myr, which is much longer than the typical lifetime of massive star formation (several Myr). Considering the short formation timescale of massive stars, the mass segregation is unlikely to be caused by dynamical relaxation (Zhang et al. 2022), as is the case for more evolved stellar clusters. In the context of a stellar cluster, such mass segregation should be considered *primordial*, although it has already evolved from its initial stage.

If the observed mass segregation is not induced by traditional dynamical processes by cores and/or stars themselves, what could be its origin? We propose that this could be naturally due to the gravitational concentration of the entire clump or gas accretion toward the center. The ALMA observations of IRDCs have already revealed a large number of sub-Jeans-mass cores during the initial fragmentation (Sanhueza et al. 2019; Morii et al. 2023). In the late stage, the MMCs are always located at the centers of the clumps or of their gravitational potentials. Our work supports the predictions of numerical simulations where members near the center of the gravitational potential will become the MMCs during the evolution due to their privileged location in the forming cluster (Bonnell & Davies 1998; Bonnell & Bate 2006).

7. Conclusion

ASSEMBLE is aimed at a comprehensive examination of the mass assembly process of massive star formation in a dynamic view, including fragmentation and accretion, and their relevance to theories. To this end, the survey employed ALMA 12 m mosaicked observations to capture both continuum and spectral line emissions in 11 massive ($M_{\rm clump} \gtrsim 10^3\ M_\odot$) and luminous ($L_{\rm bol} \gtrsim 10^4\ L_\odot$) clumps protoclusters with blue profiles. This paper releases the continuum data, characterizes the core physical properties, and presents the analyses of the evolution of the protostellar clusters, while outlining the conclusions drawn from the analysis as follows:

1. With a high angular resolution of ∼0.″8–1.″2, the 870 μm dust continuum emission reveals fragmentation with diverse morphologies. Applying the `getsf` algorithm to the continuum data, we identified a total of 248 cores across the 11 massive protoclusters, with the number of cores per clump ranging from 15 to 37.
2. We classified the cores on the basis of molecular outflows and line identification. Of the 248 cores, 142 were classified as prestellar core candidates, while 106 were identified as protostellar cores. To estimate the temperature, we used the rotational temperature derived from the multitransition lines of H_2CS and CH_3OCHO. If neither of the two lines are detected, we used the clump-averaged temperature for the prestellar core candidates. The properties of H_2CS lines in the ASSEMBLE sample will be discussed in a forthcoming article.
3. Compared to early stage ASHES protoclusters, the more evolved ASSEMBLE protoclusters show systematic increases in the average and maximum mass as well as in the surface density of both protostellar and prestellar cores. These increases indicate ongoing mass accretion onto these dense cores, which aligns with the gas accretion process observed in these massive clumps with blue profiles.
4. The mass of the MMC $M_{\rm max}$ correlates with the mass of the clump $M_{\rm clump}$ as $M_{\rm max} \propto M_{\rm clump}^{0.75}$, with a Spearman correlation coefficient of 0.73. The sublinear correlation indicates a coevolution between the clump and MMC potentially by multiscale gas accretion. In contrast, the correlation is not observed in the early stage ASHES protoclusters, consistent with the idea that early stage cores are characterized by dominant initial fragmentation rather than clump-scale gravitational accretion.
5. The correlation between the mass of MMC $M_{\rm max}$ and the mass of protoclusters $M_{\rm cluster}$ is almost linear with a power index of ∼0.9 in the first-order approximation. Despite uncertainties, the slope of the $\log M_{\rm max}$ versus $\log M_{\rm cluster}$ relation is steeper compared to that of the $\log M_{\star\,\rm max}$ versus $\log M_{\star\rm cluster}$ relation found in star clusters, which can be reconciled by an increasing trend of stellar multiplicity with mass.
6. The most massive prestellar cores found in our study have an average mass of $18.6\ M_\odot$, which is approximately 2 times larger than that found in the ASHES Pilots. Furthermore, the median and mean masses of the prestellar cores in the protoclusters are ∼2–3 times higher than those in the IRDCs. This suggests that prestellar cores are becoming more massive as a result of the continued mass accumulation within the natal clump and that high-mass prestellar cores can potentially survive in protostellar clusters. Thus, we recommend a systematic search for high-mass prestellar cores in massive protoclusters.
7. Using the MST algorithm, the cores within each cluster are connected by edges. The core separations in the ASSEMBLE sample are systematically smaller than those in the ASHES sample, indicating that the cluster becomes tighter with closer separations during its evolution. The $\mathcal{Q}$ parameters are observed to be positively correlated with both luminosity to mass ratio L/M and dust temperature $T_{\rm dust}$, indicating a more subclustered distribution at an early stage, but a more centrally condensed structure as the cluster evolves.
8. According to the mass segregation ratio (Λ) plots and the MSI that we defined in this paper, mass segregation is commonly found (8 out of 11) and clearly evident in the ASSEMBLE protoclusters. The MSI of the ASHES sample shows an insignificant difference from the random

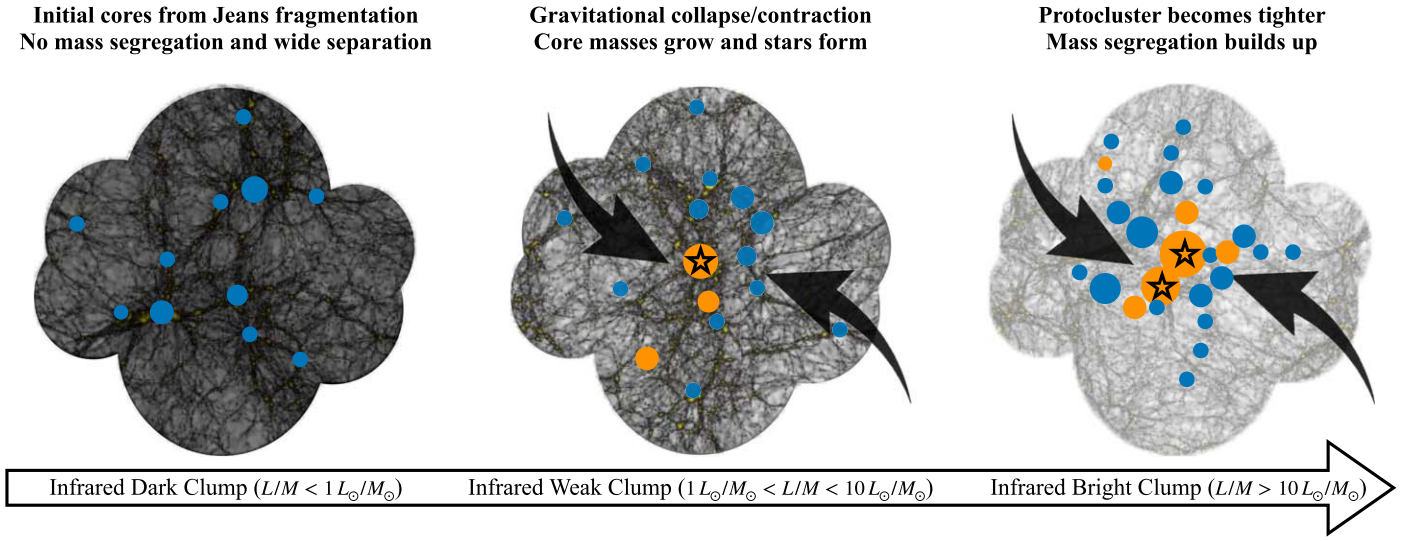

Figure 13. The cartoon of protocluster evolution from infrared dark, to infrared weak, and to infrared bright. The black filamentary structures connect the dense cores at the early stage (Morii et al. 2023) and transfer gas inwards (Xu et al. 2023a), and then fade away as the protocluster evolves (Zhou et al. 2022). The black arrows indicate inflow gas streams. ● = prestellar cores; ● = protostellar cores; ☆ = OB stars.

spatial distribution without mass segregation, indicating weak or no mass segregation in the initial stage. It was further proposed that the mass segregation should arise from gas accretion processes and gravitational concentration, as opposed to arising from dynamical interactions between point masses when the gas has already gone from the systems.

Leveraging the results and discussions presented above, we are proposing a comprehensive dynamic perspective on protocluster evolution as shown in Figure 13. At the initial stage, the protocluster originates from thermal Jeans fragmentation in infrared dark ($L/M < 1L_\odot/M_\odot$) clumps, with wide separation and no mass segregation. Subsequently, filamentary structures, especially the hub-filament system (Morii et al. 2023), act as "conveying belts" and facilitate a mass transfer toward the cores, by which the connection between the clump and the cores is gradually established (Xu et al. 2023a). Concurrently, protostars form from dense cores, leading to the heating of gas and dust within the clump, transitioning it into an infrared weak state ($1L_\odot/M_\odot < L/M < 10L_\odot/M_\odot$). Due to the effects of persistent global gravitational collapse and contraction, the protocluster becomes even tighter with narrower core separations, and the mass segregation builds up in the late stage ($L/M > 10L_\odot/M_\odot$).

The ASSEMBLE project not only provides valuable insights into the mass segregation and clustering properties of massive protoclusters but also can be used to investigate outflows (Baug et al. 2021), chemistry, and core-scale infall motion. When combined with Band 3/6 data from the ATOMS project (PI: Tie Liu; see the survey description in Liu et al. 2020), the ASSEMBLE project's data can facilitate more kinematic analyses, further illuminating how gas is transferred inward and how efficient accretion is at the clump scale and in a dynamic view. In this paper, our analyses and their statistical significance are mainly limited by sample size. As the ASSEMBLE project aims to expand its sample to include a wider range of parameters such as evolutionary stage (L/M) and clump mass ($M_{\rm clump}$), even more statistically significant results are expected.

Acknowledgments

We thank the anonymous referee for the constructive comments. This work has been supported by the National Science Foundation of China (11973013, 12033005, 11721303, 12073061, 12122307, 12203011, 12103045) and the National Key R&D Program of China (Nos. 2022YFA1603100, 2019YFA0405100). F.W.X. and K.W. acknowledge support from the China Manned Space Project (CMS-CSST-2021-A09, CMS-CSST-2021-B06), the High-Performance Computing Platform of Peking University, and the China-Chile Joint Research Fund (CCJRF No. 2211). CCJRF is provided by the Chinese Academy of Sciences South America Center for Astronomy (CASSACA) and established by the National Astronomical Observatories, Chinese Academy of Sciences (NAOC) and the Chilean Astronomy Society (SOCHIAS) to support Chinese–Chilean collaborations in astronomy. T.L. acknowledges the support by the international partnership program of Chinese academy of sciences through grant No. 114231KYSB20200009, and Shanghai Pujiang Program 20PJ1415500. M.Y.T. acknowledges the support by Yunnan provincial Department of Science and Technology through grant No. 202101BA070001-261. K.M. is financially supported by Grants-in-Aid for the Japan Society for the Promotion of Science (JSPS) Fellows (KAKENHI No. 22J21529). P.S. was partially supported by a Grant-in-Aid for Scientific Research (KAKENHI Nos. JP22H01271 and JP23H01221) of JSPS. H.L.L. is supported by Yunnan Fundamental Research Project (grant No. 202301AT070118). A.P., G.C.G., and E.V.S. acknowledge financial support from the UNAM-PAPIIT IG100223 grant. A.P. acknowledges financial support from the Sistema Nacional de Investigadores of CONAHCyT, and from the CONAHCyT project No. 86372 of the "Ciencia de Frontera 2019" program, entitled "Citlalcóatl: A multiscale study at the new frontier of the formation and early evolution of stars and planetary systems," México. G.C.G. acknowledges support from UNAM-PAPIIT IN10382 grant. A.S., G.G., and L.B. gratefully acknowledge support of ANID through the BASAL project FB210003. A.S. also gratefully acknowledges support from the Fondecyt Regular (project code 1220610). This work is sponsored (in part) by the Chinese Academy of Sciences (CAS), through a grant to the CAS South America Center for Astronomy (CASSACA) in Santiago, Chile.

M.J. acknowledges the support of the Academy of Finland grant No. 348342. C.W.L. is supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF- 2019R1A2C1010851), and by the Korea Astronomy and Space Science Institute grant funded by the Korea government (MSIT; project No. 2023-1-84000). K.T. was supported by JSPS KAKENHI (grant No. JP20H05645). This paper uses the following ALMA data: ADS/JAO. ALMA#2017.1.00545.S and 2019.1.00685.S, and ADS/JAO. ALMA#2015.1.01539.S, 2017.1.00716.S, and 2018.1.00192.S. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), Molonglo Observatory Synthesis Telescope (MOST) and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ. Data analysis was in part carried out on the High-Performance Computing Platform of Peking University, and the open-use data analysis computer system at the Astronomy Data Center (ADC) of the National Astronomical Observatory of Japan. The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation.

Facility: ALMA.

Software: Astropy (Astropy Collaboration et al. 2013, 2018), CASA (McMullin et al. 2007), getsf (Men'shchikov 2021).

Appendix A
Radio Counterpart and Environment

The radio counterpart and the environment atlas of the ASSEMBLE protoclusters are shown in Figure 14. The 1.28 GHz MeerKAT images (Padmanabh et al. 2023; S. Goedhart et al. 2023, in preparation) of the ASSEMBLE protoclusters are shown with yellow contours, with logarithmically spaced levels starting from 5σ to the peak flux. The background gray color maps show the ATLASGAL 870 μm continuum emission. The overlaid black contours show the ALMA 870 μm continuum emission as the right panels in Figure 2.

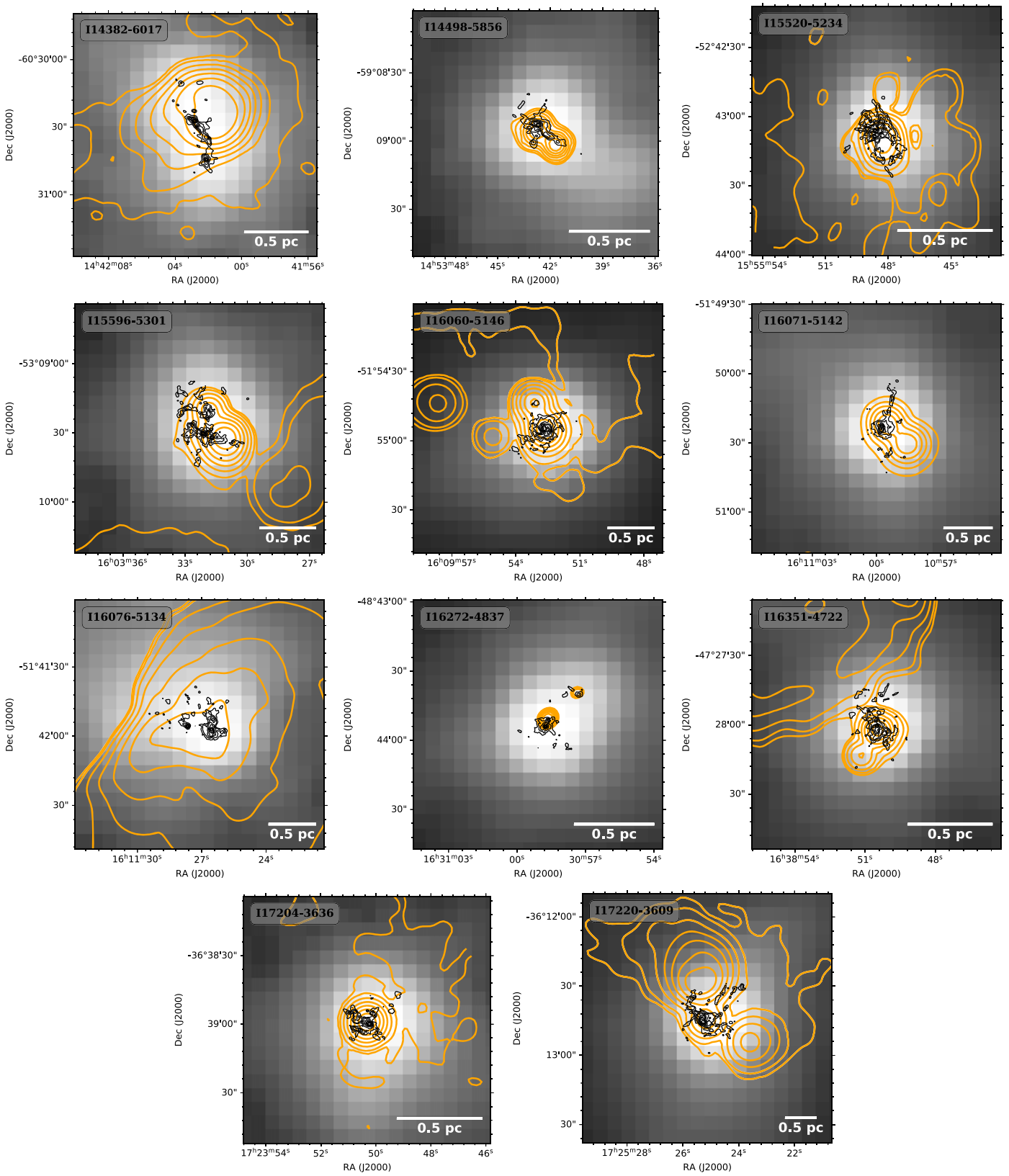

Figure 14. The 1.28 GHz radio emissions of the ASSEMBLE protoclusters are shown with yellow contours, with logarithmically spaced levels starting from 5σ to the peak flux. The background gray color maps show the ATLASGAL 870 μm continuum emission. The overlaid black contours show the ALMA 870 μm continuum emission as the right panels in Figure 2.

Appendix B
Minimum Spanning Tree Methods

MST, first developed for astrophysical applications by Barrow et al. (1985), has been applied to simulations (e.g., Wu et al. 2017) and to observations (Wang et al. 2016; Toth et al. 2017; Wang & Ge 2021; Ge & Wang 2022). In this paper, we use Prim's algorithm to find out the edges to form the tree including every node with the minimum sum of weights to form the MST. The Prim's algorithm starts with the single source node and later explores all the nodes adjacent to the source node with all the connecting edges. During the exploration, we choose the edges with the minimum weight and those that cannot cause a cycle. The edge weight is set to be the length between two vertices (Prim 1957). Therefore, MST determines a set of straight lines connecting a set of nodes (cores) that minimizes the sum of the lengths. Figures 15 displays the MST results of the 11 ASSEMBLE protoclusters.

We also adopt the MST algorithm for the 12 core clusters in the ASHES pilot survey. To keep consistency when comparing with the ASSEMBLE results, the core catalog is updated using the same source extraction algorithm on the mock 0.87 mm continuum data with the same array configuration and mosaicked coverage (see Appendix C). Compared to the original work by Sanhueza et al. (2019) who used the `astrodendro` for the 12 m+ACA+TP combined data, our new core catalog is focused on the dense and concentrated structures.

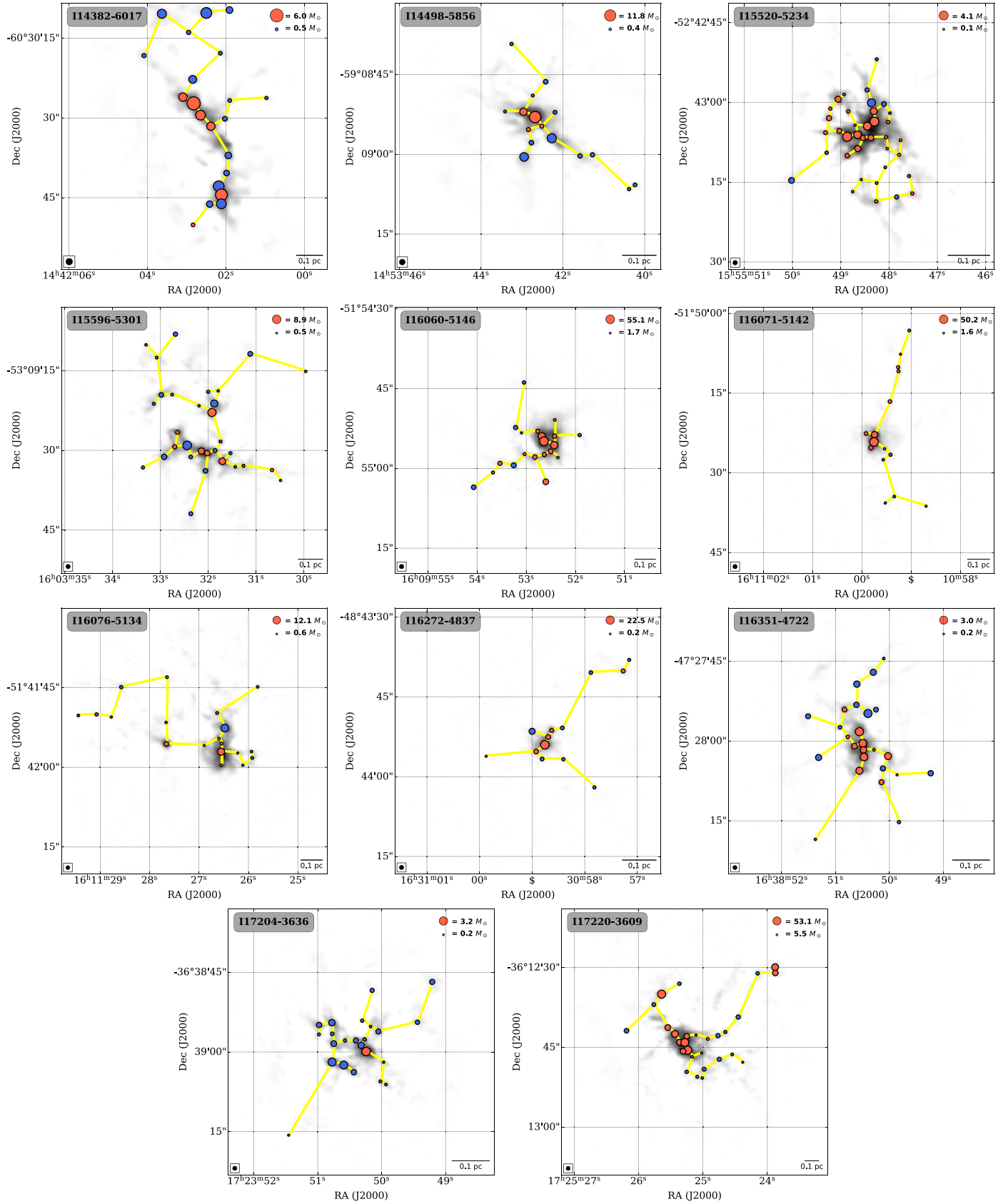

Figure 15. MST results of ASSEMBLE protoclusters. Prestellar and protostellar cores are assigned with blue and red colors, with the size normalized by the square root of core mass $\sqrt{M_{\rm core}}$ and stretched within the mass range. Yellow segments connect the cores to minimize the sum of the lengths of segments. The beam size is shown on the lower left, and the scale bar is on the right.

Appendix C
Mock Band 7 Continuum Images for ASHES Clumps

To compare the ASSEMBLE and ASHES results more directly, we simulate the Band 7 (0.87 mm) observations following the ASSEMBLE project with the ASHES Band 6 (1.3 mm) continuum data as input, to derive the mock continuum data. As the ASHES clumps are in their early stages, the free–free contamination especially at high-frequency bands (0.87 and 1.3 mm) is negligible. In other words, the continuum emission mainly comes from dust graybody emission. Therefore, the fluxes of 1.3 and 0.87 mm should follow the scaling relation as

$$\Gamma(T_{\rm dust}) \equiv \frac{F_{1.3\ {\rm mm}}}{F_{0.87\ {\rm mm}}} = \frac{\kappa_{1.3\ {\rm mm}}}{\kappa_{0.87\ {\rm mm}}} \frac{B_{1.3\ {\rm mm}}(T_{\rm dust})}{B_{0.87\ {\rm mm}}(T_{\rm dust})}, \quad \text{(C1)}$$

where κ_ν is the opacity at certain frequency, $\kappa_{1.3\ {\rm mm}}$=0.9 cm^2 g^{-1} (Sanhueza et al. 2019), and $\kappa_{0.87\ {\rm mm}} = 1.89$ cm^2 g^{-1} (Xu et al. 2023a). $B_\nu(T_{\rm dust})$ is the Planck function at corresponding frequency ν and dust temperature $T_{\rm dust}$.

As the ASHES clumps are in their early stages and lack prominent central heating sources that induce temperature gradients, we assume that dust the temperature of the entire clump is uniform, i.e., the clump-averaged $T_{\rm dust}$ as listed in column (7) of Sanhueza et al. (2019). Therefore, we simply adopt the same $T_{\rm dust}$ and then the same $\Gamma(T_{\rm dust})$ in one field to convert the flux density from the 1.3 to 0.87 mm continuum data.

Afterwards, the flux-converted images are then cropped into fields with the same shape and size as ASSEMBLE data, to assure the same field of view where sources are extracted. The mock field is placed with both the major and minor axes aligned with those of the ASHES field with its center toward the densest part of the cluster. The 12 mock images are used as the basic input for source extraction algorithm `getsf` and the MST algorithm (Appendix B).

ORCID iDs

Fengwei Xu https://orcid.org/0000-0001-5950-1932
Ke Wang https://orcid.org/0000-0002-7237-3856
Tie Liu https://orcid.org/0000-0002-5286-2564
Mengyao Tang https://orcid.org/0000-0001-9160-2944
Neal J. Evans II https://orcid.org/0000-0001-5175-1777
Aina Palau https://orcid.org/0000-0002-9569-9234
Kaho Morii https://orcid.org/0000-0002-6752-6061
Jinhua He https://orcid.org/0000-0002-3938-4393
Patricio Sanhueza https://orcid.org/0000-0002-7125-7685
Hong-Li Liu https://orcid.org/0000-0003-3343-9645
Amelia Stutz https://orcid.org/0000-0003-2300-8200
Qizhou Zhang https://orcid.org/0000-0003-2384-6589
Xi Chen https://orcid.org/0000-0002-5435-925X
Pak Shing Li https://orcid.org/0000-0001-8077-7095
Gilberto C. Gómez https://orcid.org/0000-0003-4714-0636
Enrique Vázquez-Semadeni https://orcid.org/0000-0002-1424-3543
Shanghuo Li https://orcid.org/0000-0003-1275-5251
Xiaofeng Mai https://orcid.org/0000-0001-7573-0145
Xing Lu https://orcid.org/0000-0003-2619-9305
Meizhu Liu https://orcid.org/0000-0002-5789-7504
Li Chen https://orcid.org/0009-0009-8154-4205
Chuanshou Li https://orcid.org/0000-0001-5710-6509
Hongqiong Shi https://orcid.org/0000-0001-8277-1367
Zhiyuan Ren https://orcid.org/0000-0003-4659-1742
Di Li https://orcid.org/0000-0003-3010-7661
Guido Garay https://orcid.org/0000-0003-1649-7958
Leonardo Bronfman https://orcid.org/0000-0002-9574-8454
Lokesh Dewangan https://orcid.org/0000-0001-6725-0483
Mika Juvela https://orcid.org/0000-0002-5809-4834
Chang Won Lee https://orcid.org/0000-0002-3179-6334
S. Zhang https://orcid.org/0000-0002-9836-0279
Nannan Yue https://orcid.org/0000-0003-0355-6875
Yifei Ge https://orcid.org/0000-0002-8727-0868
Wenyu Jiao https://orcid.org/0000-0001-9822-7817
Qiuyi Luo https://orcid.org/0000-0003-4506-3171
Ken'ichi Tatematsu https://orcid.org/0000-0002-8149-8546
James O. Chibueze https://orcid.org/0000-0002-9875-7436
Keyun Su https://orcid.org/0009-0003-2243-7983
Shenglan Sun https://orcid.org/0000-0002-9796-1507
I. Ristorcelli https://orcid.org/0000-0002-1469-6323
L. Viktor Toth https://orcid.org/0000-0002-5310-4212

References

Allison, R. J., Goodwin, S. P., Parker, R. J., et al. 2009, MNRAS, 395, 1449
Anderson, M., Peretto, N., Ragan, S. E., et al. 2021, MNRAS, 508, 2964
Avison, A., Peretto, N., Fuller, G. A., et al. 2015, A&A, 577, A30
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F., & Zamora-Avilés, M. A. 2011, MNRAS, 411, 65
Ballesteros-Paredes, J., Vázquez-Semadeni, E., Palau, A., & Klessen, R. S. 2018, MNRAS, 479, 2112
Barnes, A. T., Liu, J., Zhang, Q., et al. 2023, A&A, 675, A53
Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, MNRAS, 216, 17
Baug, T., Wang, K., Liu, T., et al. 2020, ApJ, 890, 44
Baug, T., Wang, K., Liu, T., et al. 2021, MNRAS, 507, 4316
Beuther, H., Linz, H., Tackenberg, J., et al. 2013, A&A, 553, A115
Beuther, H., Mottram, J. C., Ahmadi, A., et al. 2018, A&A, 617, A100
Bonnell, I. A., & Bate, M. R. 2006, MNRAS, 370, 488
Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 2001, MNRAS, 323, 785
Bonnell, I. A., Bate, M. R., & Vine, S. G. 2003, MNRAS, 343, 413
Bonnell, I. A., & Davies, M. B. 1998, MNRAS, 295, 691
Bonnell, I. A., Vine, S. G., & Bate, M. R. 2004, MNRAS, 349, 735
Bronfman, L., Nyman, L. A., & May, J. 1996, A&AS, 115, 81
Camacho, V., Vázquez-Semadeni, E., Palau, A., Busquet, G., & Zamora-Avilés, M. 2020, ApJ, 903, 46
Cartwright, A., & Whitworth, A. P. 2004, MNRAS, 348, 589
Chambers, E. T., Jackson, J. M., Rathborne, J. M., & Simon, R. 2009, ApJS, 181, 360
Chira, R.-A., Smith, R. J., Klessen, R. S., Stutz, A. M., & Shetty, R. 2014, MNRAS, 444, 874
Churchwell, E., Walmsley, C. M., & Cesaroni, R. 1990, A&AS, 83, 119
Contreras, Y., Sanhueza, P., Jackson, J. M., et al. 2018, ApJ, 861, 14
Contreras, Y., Schuller, F., Urquhart, J. S., et al. 2013, A&A, 549, A45
Cyganowski, C. J., Brogan, C. L., Hunter, T. R., et al. 2014, ApJL, 796, L2
Cyganowski, C. J., Whitney, B. A., Holden, E., et al. 2008, AJ, 136, 2391
Dewangan, L. K., Ojha, D. K., Sharma, S., et al. 2020, ApJ, 903, 13
Dewangan, L. K., Zinchenko, I. I., Zemlyanukha, P. M., et al. 2022, ApJ, 925, 41
Dib, S., & Henning, T. 2019, A&A, 629, A135
Elia, D., Merello, M., Molinari, S., et al. 2021, MNRAS, 504, 2742
Elia, D., Molinari, S., Schisano, E., et al. 2017, MNRAS, 471, 100
Ellingsen, S. P., Shabala, S. S., & Kurtz, S. E. 2005, MNRAS, 357, 1003
Evans, N. J., Kim, J.-G., & Ostriker, E. C. 2022, ApJL, 929, L18
Faúndez, S., Bronfman, L., Garay, G., et al. 2004, A&A, 426, 97
Galván-Madrid, R., Zhang, Q., Keto, E., et al. 2010, ApJ, 725, 17
Ge, Y., & Wang, K. 2022, ApJS, 259, 36
Ge, Y., Wang, K., Duarte-Cabral, A., et al. 2023, A&A, 675, A119
Gieser, C., Semenov, D., Beuther, H., et al. 2019, A&A, 631, A142
Gómez, G. C., & Vázquez-Semadeni, E. 2014, ApJ, 791, 124

Guszejnov, D., Markey, C., Offner, S. S. R., et al. 2022, MNRAS, 515, 167
Harju, J., Walmsley, C. M., & Wouterloot, J. G. A. 1993, A&AS, 98, 51
He, C.-C., & Ricotti, M. 2023, MNRAS, 522, 5374
He, Y.-X., Zhou, J.-J., Esimbek, J., et al. 2015, MNRAS, 450, 1926
Heitsch, F., Hartmann, L. W., Slyz, A. D., Devriendt, J. E. G., & Burkert, A. 2008, ApJ, 674, 316
Huang, B., Wang, K., Girart, J. M., et al. 2023, ApJ, 949, 46
Kong, S., Arce, H. G., Shirley, Y., & Glasgow, C. 2021, ApJ, 912, 156
Kong, S., Tan, J. C., Caselli, P., et al. 2017, ApJ, 834, 193
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, ApJ, 656, 959
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2011, ApJ, 740, 74
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Li, S., Sanhueza, P., Lee, C. W., et al. 2022, ApJ, 926, 165
Li, S., Sanhueza, P., Zhang, Q., et al. 2023, ApJ, 949, 109
Li, S., Zhang, Q., Pillai, T., et al. 2019, ApJ, 886, 130
Liu, H.-L., Tej, A., Liu, T., et al. 2022a, MNRAS, 510, 5009
Liu, H.-L., Tej, A., Liu, T., et al. 2022b, MNRAS, 511, 4480
Liu, H.-L., Tej, A., Liu, T., et al. 2023, MNRAS, 522, 3719
Liu, M., Tan, J. C., Cheng, Y., & Kong, S. 2018, ApJ, 862, 105
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, MNRAS, 496, 2790
Liu, T., Kim, K.-T., Yoo, H., et al. 2016a, ApJ, 829, 59
Liu, T., Zhang, Q., Kim, K.-T., et al. 2016b, ApJ, 824, 31
Liu, T., Lacy, J., Li, P. S., et al. 2017, ApJ, 849, 25
Longmore, S. N., Burton, M. G., Barnes, P. J., et al. 2007, MNRAS, 379, 535
Longmore, S. N., Kruijssen, J. M. D., Bastian, N., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 291
Longmore, S. N., Pillai, T., Keto, E., Zhang, Q., & Qiu, K. 2011, ApJ, 726, 97
Louvet, F. 2018, SF2A-2018: Proc. of the Annual Meeting of the French Society of Astronomy and Astrophysics, ed. P. Di Matteo et al. (Paris: Société Française d'Astronomie et d'Astrophysique), 311
Lu, X., Zhang, Q., Liu, H. B., et al. 2018, ApJ, 855, 9
Mann, H. B., & Whitney, D. R. 1947, Ann. Math. Sci., 18, 50
Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425
McKee, C. F., & Tan, J. C. 2003, ApJ, 585, 850
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127
Men'shchikov, A. 2021, A&A, 649, A89
Moeckel, N., & Bonnell, I. A. 2009, MNRAS, 396, 1864
Molet, J., Brouillet, N., Nony, T., et al. 2019, A&A, 626, A132
Molinari, S., Brand, J., Cesaroni, R., & Palla, F. 1996, A&A, 308, 573
Molinari, S., Swinyard, B., Bally, J., et al. 2010, PASP, 122, 314
Morii, K., Sanhueza, P., Nakamura, F., et al. 2021, ApJ, 923, 147
Morii, K., Sanhueza, P., Nakamura, F., et al. 2023, ApJ, 950, 148
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Myers, A. T., Klein, R. I., Krumholz, M. R., & McKee, C. F. 2014, MNRAS, 439, 3420
Myers, A. T., McKee, C. F., Cunningham, A. J., Klein, R. I., & Krumholz, M. R. 2013, ApJ, 766, 97
Naranjo-Romero, R., Vázquez-Semadeni, E., & Loughnane, R. M. 2022, MNRAS, 512, 4715
Nony, T., Galván-Madrid, R., Motte, F., et al. 2023, A&A, 674, A75
Nony, T., Louvet, F., Motte, F., et al. 2018, A&A, 618, L5
Offner, S. S. R., Klein, R. I., McKee, C. F., & Krumholz, M. R. 2009, ApJ, 703, 131
Offner, S. S. R., Moe, M., Kratter, K. M., et al. 2023, in ASP Conf. Ser. 534, Protostars and Planets VII, ed. S. Inutsuka et al. (San Francisco, CA: ASP), 275
Olguin, F. A., Sanhueza, P., Ginsburg, A., et al. 2022, ApJ, 929, 68
Olguin, F. A., Sanhueza, P., Guzmán, A. E., et al. 2021, ApJ, 909, 199
Ossenkopf, V., & Henning, T. 1994, A&A, 291, 943
Padmanabh, P. V., Barr, E. D., Sridhar, S. S., et al. 2023, MNRAS, 524, 1291
Padoan, P., Pan, L., Juvela, M., Haugbølle, T., & Nordlund, Å. 2020, ApJ, 900, 82
Palau, A., Ballesteros-Paredes, J., Vázquez-Semadeni, E., et al. 2015, MNRAS, 453, 3785
Palau, A., Estalella, R., Girart, J. M., et al. 2014, ApJ, 785, 42
Palau, A., Zapata, L. A., Román-Zúñiga, C. G., et al. 2018, ApJ, 855, 24
Palau, A., Zhang, Q., Girart, J. M., et al. 2021, ApJ, 912, 159
Parker, R. J., & Goodwin, S. P. 2015, MNRAS, 449, 3381
Pelkonen, V. M., Padoan, P., Haugbølle, T., & Nordlund, Å. 2021, MNRAS, 504, 1219
Peretto, N., & Fuller, G. A. 2009, A&A, 505, 405
Peretto, N., Fuller, G. A., André, P., et al. 2014, A&A, 561, A83
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Pillai, T., Kauffmann, J., Zhang, Q., et al. 2019, A&A, 622, A54
Pillai, T., Wyrowski, F., Carey, S. J., & Menten, K. M. 2006, A&A, 450, 569
Pouteau, Y., Motte, F., Nony, T., et al. 2023, A&A, 674, A76
Prim, R. C. 1957, BSTJ, 36, 1389
Qin, S.-L., Liu, T., Liu, X., et al. 2022, MNRAS, 511, 3463
Ragan, S. E., Bergin, E. A., & Wilner, D. 2011, ApJ, 736, 163
Ragan, S. E., Heitsch, F., Bergin, E. A., & Wilner, D. 2012, ApJ, 746, 174
Rathborne, J. M., Jackson, J. M., & Simon, R. 2006, ApJ, 641, 389
Rathborne, J. M., Simon, R., & Jackson, J. M. 2007, ApJ, 662, 1082
Redaelli, E., Bovino, S., Sanhueza, P., et al. 2022, ApJ, 936, 169
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131
Reinoso, B., Schleicher, D. R. G., Fellhauer, M., Leigh, N. W. C., & Klessen, R. S. 2020, A&A, 639, A92
Rosolowsky, E., Dunham, M. K., Ginsburg, A., et al. 2010, ApJS, 188, 123
Sanhueza, P., Contreras, Y., Wu, B., et al. 2019, ApJ, 886, 102
Sanhueza, P., Garay, G., Bronfman, L., et al. 2010, ApJ, 715, 18
Sanhueza, P., Girart, J. M., Padovani, M., et al. 2021, ApJL, 915, L10
Sanhueza, P., Jackson, J. M., Foster, J. B., et al. 2012, ApJ, 756, 60
Sanhueza, P., Jackson, J. M., Zhang, Q., et al. 2017, ApJ, 841, 97
Schneider, N., Csengeri, T., Bontemps, S., et al. 2010, A&A, 520, A49
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415
Spitzer, L. 1987, Dynamical Evolution of Globular Clusters (Princeton, NJ: Princeton Univ. Press)
Spitzer, L. J. 1969, ApJL, 158, L139
Sridharan, T. K., Beuther, H., Schilke, P., Menten, K. M., & Wyrowski, F. 2002, ApJ, 566, 931
Svoboda, B. E., Shirley, Y. L., Traficante, A., et al. 2019, ApJ, 886, 36
Takemura, H., Nakamura, F., Arce, H. G., et al. 2023, ApJS, 264, 35
Tan, J. C., Beltrán, M. T., Caselli, P., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al., 149 (Tucson, AZ: Univ. Arizona Press)
Tan, J. C., Kong, S., Butler, M. J., Caselli, P., & Fontani, F. 2013, ApJ, 779, 96
Tang, M., Palau, A., Zapata, L. A., & Qin, S.-L. 2022, A&A, 657, A30
Testi, L., Palla, F., & Natta, A. 1999, A&A, 342, 515
Toth, L. V., Zahorecz, S., Marton, G., et al. 2017, in IAU Symp. 316, Formation, Evolution, and Survival of Massive Star Clusters, ed. C. Charbonnel & A. Nota (Cambridge: Cambridge Univ. Press), 133
Traficante, A., Jones, B. M., Avison, A., et al. 2023, MNRAS, 520, 2306
Urquhart, J. S., Csengeri, T., Wyrowski, F., et al. 2014, A&A, 568, A41
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Urquhart, J. S., Morgan, L. K., Figura, C. C., et al. 2011, MNRAS, 418, 1689
Vázquez-Semadeni, E., Gómez, G. C., & González-Samaniego, A. 2023, arXiv:2306.13846
Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., Ballesteros-Paredes, J., & Klessen, R. S. 2009, ApJ, 707, 1023
Vázquez-Semadeni, E., González-Samaniego, A., & Colín, P. 2017, MNRAS, 467, 1313
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Wang, K., & Ge, Y., 2021 MST: Minimum Spanning Tree algorithm for identifying large-scale filaments, Astrophysics Source Code Library, ascl:2102.002
Wang, K., Testi, L., Burkert, A., et al. 2016, ApJS, 226, 9
Wang, K., Zhang, Q., Testi, L., et al. 2014, MNRAS, 439, 3275
Wang, K., Zhang, Q., Wu, Y., Li, H.-b., & Zhang, H. 2012, ApJL, 745, L30
Wang, K., Zhang, Q., Wu, Y., & Zhang, H. 2011, ApJ, 735, 64
Wang, Y., Zhang, Q., Pillai, T., Wyrowski, F., & Wu, Y. 2008, ApJL, 672, L33
Wang, Y., Zhang, Q., Rathborne, J. M., Jackson, J., & Wu, Y. 2006, ApJL, 651, L125
Ward, J. L., & Kruijssen, J. M. D. 2018, MNRAS, 475, 5659
Weidner, C., & Kroupa, P. 2004, MNRAS, 348, 187
Weidner, C., & Kroupa, P. 2006, MNRAS, 365, 1333
Weidner, C., Kroupa, P., & Pflamm-Altenburg, J. 2013, MNRAS, 434, 84
Wu, B., Tan, J. C., Christie, D., et al. 2017, ApJ, 841, 88
Wu, Y., Zhang, Q., Yu, W., et al. 2006, A&A, 450, 607
Xie, J., Fuller, G. A., Li, D., et al. 2021, SCPMA, 64, 279511
Xu, F.-W., Wang, K., Liu, T., et al. 2023a, MNRAS, 520, 3259
Xu, F., Wang, K., He, Y., et al. 2023b, ApJS, 269, 38
Yang, D., Liu, H.-L., Tej, A., et al. 2023, ApJ, 953, 40
Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, ApJ, 852, 12
Yuan, J., Wu, Y., Ellingsen, S. P., et al. 2017, ApJS, 231, 11
Yue, Y.-H., Qin, S.-L., Liu, T., et al. 2021, RAA, 21, 014
Zhang, C., Zhu, F.-Y., Liu, T., et al. 2023, MNRAS, 520, 3245
Zhang, Q., & Wang, K. 2011, ApJ, 733, 26
Zhang, Q., Wang, K., Lu, X., & Jiménez-Serra, I. 2015, ApJ, 804, 141
Zhang, Q., Wang, Y., Pillai, T., & Rathborne, J. 2009, ApJ, 696, 268
Zhang, Y., Tanaka, K. E. I., Tan, J. C., et al. 2022, ApJ, 936, 68
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038